

UNIVERSIDADE FEDERAL FLUMINENSE  
ESCOLA DE ENGENHARIA  
LABORATÓRIO DE TECNOLOGIA, GESTÃO DE NEGÓCIOS E MEIO AMBIENTE  
DOUTORADO EM SISTEMAS DE GESTÃO SUSTENTÁVEIS

**AYRTON BENEDITO GAIA DO COUTO**

**AVALIAÇÃO DA SUSTENTABILIDADE E DO DESEMPENHO DE  
EMPRESAS DO SETOR ELÉTRICO BRASILEIRO: APLICAÇÃO DO APOIO  
MULTICRITÉRIO À DECISÃO E APRENDIZADO DE MÁQUINA**

Tese apresentada ao Programa de Doutorado em Sistemas de Gestão Sustentáveis da Universidade Federal Fluminense como requisito parcial para obtenção do Grau de Doutor em Sistemas de Gestão Sustentáveis. Área de Concentração: Sistemas de Gestão da Sustentabilidade. Linha de Pesquisa: Apoio à Decisão em Organizações Sustentáveis.

Orientador:

Prof. Luís Alberto Duncan Rangel, D.Sc.  
Universidade Federal Fluminense - UFF

Niterói, RJ  
2024

Ficha catalográfica automática - SDC/BEE  
Gerada com informações fornecidas pelo autor

G137a Gaia Do Couto, Ayrton Benedito  
Avaliação da sustentabilidade e do desempenho de empresas do setor elétrico brasileiro: aplicação do apoio multicritério à decisão e aprendizado de máquina / Ayrton Benedito Gaia Do Couto. - 2024.  
190 p.: il.

Orientador: Luís Alberto Duncan Rangel.  
Tese (doutorado)-Universidade Federal Fluminense, Escola de Engenharia, Niterói, 2024.

1. Modelo de avaliação. 2. Empresas de energia. 3. Sustentabilidade. 4. Apoio Multicritério à Decisão. 5. Produção intelectual. I. Duncan Rangel, Luís Alberto, orientador. II. Universidade Federal Fluminense. Escola de Engenharia. III. Título.

CDD - XXX

AYRTON BENEDITO GAIA DO COUTO

**AVALIAÇÃO DA SUSTENTABILIDADE E DO DESEMPENHO DE  
EMPRESAS DO SETOR ELÉTRICO BRASILEIRO: APLICAÇÃO DO APOIO  
MULTICRITÉRIO À DECISÃO E APRENDIZADO DE MÁQUINA**

Tese apresentada ao Programa de Doutorado em Sistemas de Gestão Sustentáveis da Universidade Federal Fluminense como requisito parcial para obtenção do Grau de Doutor em Sistemas de Gestão Sustentáveis. Área de Concentração: Sistemas de Gestão da Sustentabilidade. Linha de Pesquisa: Apoio à Decisão em Organizações Sustentáveis.

Aprovado em 05 de fevereiro de 2024.

**BANCA EXAMINADORA**

---

Prof. Luís Alberto Duncan Rangel, D.Sc. - Orientador  
Universidade Federal Fluminense - UFF

---

Prof. Gilson Brito Alves Lima, D.Sc.  
Universidade Federal Fluminense - UFF

---

Prof. Luiz Flavio Autran Monteiro Gomes, Ph.D.  
Universidade Federal do ABC - UFABC

---

Prof. Elton Felipe Sbruzzi, Ph.D.  
Instituto Tecnológico de Aeronáutica - ITA

---

Prof. Plácido Rogério Pinheiro, D.Sc.  
Universidade de Fortaleza - UNIFOR

Niterói, RJ  
2024

## Considerações, dedicatória e agradecimentos

“(…) ninguém pode prever eventos futuros com exatidão se não é capaz sequer de medir de forma precisa o atual estado do universo!”, sobre o princípio da incerteza de Werner Karl Heisenberg, alemão, físico, Prêmio Nobel de Física (HAWKING, 2015).

“(…) quase todos os nossos conhecimentos são apenas prováveis; e no pequeno número de coisas que podemos saber com certeza, mesmo nas ciências matemáticas, os principais meios para se chegar à verdade – a indução e a analogia – são fundados nas probabilidades”, Pierre-Simon Laplace, francês, matemático, astrônomo e físico (LAPLACE, 2010).

“Dê vida a seus dias, em vez de dar dias à sua vida”, Rita Levi-Montalcini, italiana, médica, Prêmio Nobel de Medicina (MORIN, 2021).

Dedico esta Tese de Doutorado aos meus pais *in memoriam*; à minha esposa, Giselle, e aos meus filhos Mariana e Guilherme, que com paciência e amor, souberam compreender e entender todo o meu trabalho necessário à construção desta tese.

E, agradeço aos Professores e funcionários do Programa de Doutorado em Sistemas de Gestão Sustentáveis, aos demais Professores da Banca Examinadora, e aos amigos, pelos incentivos recebidos.

## RESUMO

Os indicadores definidos pela ANEEL, Agência Nacional de Energia Elétrica, constantes no Manual de Contabilidade do Setor Elétrico, MCSE, constituem meios para avaliar as empresas de energia elétrica com base em padrões internacionais, como o GRI, *Global Reporting Initiative*, bem como, em necessidades específicas com o intuito de apreciar o desempenho daquelas empresas. Inicialmente, foram coletados sessenta e dois indicadores no período (ano-base) 2013-2021, distribuídos nas dimensões econômica, social, ambiental e de governança corporativa. Neste contexto, o modelo proposto de avaliação integrada das empresas permitiu realizar não somente um *benchmarking*, tendo em vista as várias formas existentes de construí-lo. Mas, aproveitar a ocasião como experiência de maravilhamento da retórica, ou, romper com o *imprinting* cultural, que impede de ver diferentemente do que se mostra. Trata-se de uma oportunidade para se perguntar: o que pode existir ou ser realizado além do já conhecido termo *benchmarking*? Assim, a contribuição desta Tese de Doutorado está na avaliação integrada da sustentabilidade e do desempenho das empresas concessionárias de distribuição de energia elétrica no Brasil, por meio de *benchmarking*, com Apoio Multicritério à Decisão e modelo preditivo para a classificação e gestão de riscos das empresas. Para tais tarefas, utilizou-se de teoria em Apoio Multicritério à Decisão, *Rough Sets Theory* e sua extensão, *Dominance Rough Set Approach*, as quais permitiram o tratamento matemático da imprecisão e incerteza de dados, e análise dos indicadores, com a identificação de padrões de dados e construção de matrizes de risco condicional. Ademais, foi utilizado Aprendizado de Máquina, subcampo da Inteligência Artificial, para a predição da classificação das empresas com base em dados históricos de indicadores, inferência de regras de decisão, simulações e método *ensemble bagging* para a reunião de algoritmos classificadores, constituindo-se em ferramenta para as empresas na gestão dos seus riscos e para se atingir metas. Como *softwares* de apoio à pesquisa, foram utilizados VBA/Microsoft Excel, jMAF, Weka e RStudio.

**PALAVRAS-CHAVE:** Sustentabilidade, Energia, Modelo de avaliação, Decisão Multicritério, Teoria dos Conjuntos Aproximativos, princípio de Dominância, Aprendizado de Máquina.

## ABSTRACT

The indicators defined by ANEEL, the National Electric Power Agency, constant in the Accounting Manual of the Electricity Sector, AMES, constitute the means to assess the electric power companies based on international standards, like the GRI, Global Reporting Initiative, as well as in specific necessities with the aim of appraising those companies' performances. Initially, sixty-two indicators from the period (year-base) 2013-2021 were collected, distributed around the economic, social, environmental and corporate governance dimensions. In this context, the proposed model of integrated assessment of the companies allowed not only to make a benchmarking, given that there are many ways of doing it. But should the opportunity be seized as an experience of rhetoric wonderment, or should the cultural imprinting that stops one from seeing things in differently than what is shown be broken with? It is an opportunity to ask: what could exist or be achieved beyond what is already known about the word benchmarking? Therefore, this Doctorate thesis' contribution is the integrated assessment of the sustainability and performance of electricity distribution concession companies in Brazil, through benchmarking, with Multi-criteria Decision Support and a predictive model for the classification and risk management of companies. For such tasks, the Rough Sets Theory and its extension, Dominance Rough Set Approach, were used, which allowed a mathematical treatment of data imprecision and uncertainty, and analysis of the indicators, with data patterns recognition and the construction of conditional risk matrixes. Furthermore, Machine Learning, a subset of Artificial Intelligence, was used for the prediction of companies' classification based on historic data of indicators, inferring of decision rules, simulation and ensemble bagging method for grouping classification algorithms, constituting tools for companies to use in their risk management and to reach target goals. As supporting software for this research, VBA/Microsoft Excel, jMAF, Weka and RStudio were used.

**KEYWORDS:** Sustainability, Energy, Assessment Model, Multicriteria Decision, Rough Sets Theory, Dominance Rough Set Approach, Machine Learning.

## LISTA DE FIGURAS

Figura 1 – A sustentabilidade em suas dimensões .....	17
Figura 2 – Modelo de Estrela Decisória para a solução de problemas.....	25
Figura 3– Classificação de Métodos de Apoio Multicritério à Decisão (MCDM) .....	29
Figura 4– Hierarquia do AM.....	39
Figura 5– Abordagens em AM, categoria supervisionado, tarefa de classificação.....	40
Figura 6 – Processo KDD, da seleção de dados até a obtenção de conhecimento .....	44
Figura 7– Relevância dos periódicos nas referências de Sartori et al. (2017), Eixo 1 .....	50
Figura 8 – Autores com maior participação nas referências Sartori et al. (2017), Eixo 1.....	50
Figura 9– Relevância dos periódicos nas referências dos artigos do portfólio bibliográfico final (PB5), Eixo 2.....	52
Figura 10– Autores com maior participação nas referências do portfólio bibliográfico final (PB5), Eixo 2.....	52
Figura 11– Linha de corte (tracejado, em vermelho) para o portfólio bibliográfico final (PB5), Eixo 3.....	58
Figura 12– Relevância dos periódicos no portfólio bibliográfico final (PB5), Eixo 3.....	64
Figura 13– Relevância dos periódicos nas referências do portfólio bibliográfico final (PB5), Eixo 3 .....	64
Figura 14– Relevância dos autores nas referências do portfólio bibliográfico final (PB5), Eixo 3 .....	65
Figura 15- Modelo conceitual para avaliação integrada da sustentabilidade e desempenho .....	72
Figura 16 – Ordem crescente de preferência forte para as classes diretamente (Dn) e indiretamente (In) relacionadas à sustentabilidade .....	74
Figura 17 – Matriz de Sustentabilidade para o modelo quartis-percentil 93% .....	75
Figura 18– Indicadores de sustentabilidade e desempenho, e dicionário de termos de indexação .....	89
Figura 19– Indicadores de governança corporativa e dicionário de termos de indexação .....	90
Figura 20– Preenchimento de valores ausentes pelas medianas horizontal (em azul) e vertical (em verde) .....	92
Figura 21– Indicadores dirSU, valores absolutos, ano 2021, dez primeiras empresas – parte 1. 94	
Figura 22– Indicadores dirSU, valores normalizados, ano 2021, dez primeiras empresas – parte 1 .....	96
Figura 23– Resultado da classificação para as dez primeiras empresas, período 2021-2013 .....	98
Figura 24– Processamento de tabela de decisão, indicadores e classe dirSU com apoio do software jMAF .....	101
Figura 25– Exportação da tabela de decisão dirSU como data frame para o package RoughSets em R .....	102
Figura 26– Probabilidade condicional #4, Quadro 24: probabilidade de uma empresa pertencer à região hachurada dada que a classe dirSU não seja D5 .....	106
Figura 27– Probabilidade condicional #5, Quadro 25: probabilidade de uma empresa pertencer à região hachurada dada que a classe indSU seja II .....	107
Figura 28– Regiões mais vantajosas (hachuradas em verde) na Matriz de Sustentabilidade vis-à-vis às probabilidades condicionais .....	107
Figura 29– Regiões menos vantajosas (hachuradas em vermelho) na Matriz de Sustentabilidade vis-à-vis às probabilidades condicionais .....	108
Figura 30– Método ensemble com “k” modelos classificadores, votação de classes e classe predita.....	111

Figura 31– Método ensemble bagging e algoritmos classificadores para predição de classe (votada) .....	115
Figura 32– Matriz de Sustentabilidade visando simular a ascensão ou descensão de classes, empresa E25, ano 2021, método ensemble bagging e algoritmos classificadores .....	117
Figura 33– Predição da classe D4 pelo algoritmo VC-DomLEM para a empresa E25, com base em dados históricos de indicadores dirSU e valores simulados .....	119
Figura 34– Resultado do benchmarking, dimensão geral, ano 2021; empresa E07 como benchmark .....	120
Figura 35– Posição relativa dos indicadores dirSU, empresa E07, ano 2021 .....	121
Figura 36– Posição relativa dos indicadores indSU, empresa E07, ano 2021 .....	121
Figura 37– Posição relativa dos indicadores dirSU, empresa E33, ano 2021 .....	122
Figura 38– Posição relativa dos indicadores indSU, empresa E33, ano 2021 .....	122
Figura 39– Resultado do benchmarking, dimensão econômica, ano 2021 .....	123
Figura 40– Resultado do benchmarking, dimensão social, ano 2021 .....	123
Figura 41– Resultado do benchmarking, dimensão ambiental, ano 2021 .....	124
Figura 42– Resultado do benchmarking, dimensão governança corporativa, ano 2021 .....	124
Figura 43– Resultado da evolução da empresa E27, dimensão geral, período 2013-2021 .....	125
Figura 44– Resultado da evolução da empresa E27, dimensão econômica, período 2013-2021 .....	126
Figura 45– Resultado da evolução da empresa E27, dimensão social, período 2013-2021 .....	126
Figura 46– Resultado da evolução da empresa E27, dimensão ambiental, período 2013-2021 .....	127
Figura 47– Resultado da evolução da empresa E27, dimensão governança corporativa, período 2013-2021 .....	127
Figura 48– Movimentos de migração de classes para a empresa E25, ano de 2021 .....	128
Figura 49– Simulação de descensão de classes, empresa E07, ano 2021 .....	134
Figura 50– Resultado do benchmarking e posições das empresas E07, E28 e E27 .....	137
Figura 51– Uso da técnica de <i>clustering</i> , algoritmo <i>FarthestFirst</i> , para identificar valor <i>outlier</i> .....	145



## LISTA DE QUADROS

Quadro 1– Exemplo de base de conhecimento .....	30
Quadro 2– Exemplo com seis lojas .....	33
Quadro 3– Resumo da Metodologia de Pesquisa.....	45
Quadro 4– Revisão sistemática da literatura .....	46
Quadro 5– Definição das palavras-chave e Eixos de pesquisa para a revisão sistemática da literatura .....	47
Quadro 6– Strings de pesquisa para a revisão sistemática da literatura .....	47
Quadro 7– Artigos que formam o portfólio bibliográfico final (PB5), Eixo 1.....	49
Quadro 8– Número de citações do portfólio bibliográfico final (PB5), Eixo 1 .....	49
Quadro 9– Palavras-chave pesquisadas no portfólio bibliográfico final (PB5), Eixo 1 .....	51
Quadro 10– Artigos que formam o portfólio bibliográfico final (PB5), Eixo 2.....	51
Quadro 11– Número de citações do portfólio bibliográfico (PB5), Eixo 2.....	52
Quadro 12– Palavras-chave pesquisadas no portfólio bibliográfico final (PB5), Eixo 2.....	53
Quadro 13– Artigos que formam o portfólio bibliográfico final (PB5), Eixo 3.....	54
Quadro 14– Número de citações do portfólio bibliográfico final (PB5), Eixo 3 .....	59
Quadro 15– Relação das palavras-chave pesquisadas no portfólio bibliográfico final (PB5), Eixo 3.....	65
Quadro 16- Complemento à revisão sistemática da literatura, período janeiro/2022-agosto/2023 .....	67
Quadro 17- Resultado da pesquisa de palavras-chave, Eixo 2.....	68
Quadro 18- Resultado da pesquisa de palavras-chave, Eixo 3.....	68
Quadro 19– Empresas concessionárias de distribuição de energia elétrica consideradas na pesquisa, em 2021 .....	81
Quadro 20– Indicadores ESA diretamente associados à sustentabilidade e desempenho (dirSU) .....	83
Quadro 21– Indicadores GC diretamente associados à sustentabilidade e desempenho (dirSU) 84	
Quadro 22– Indicadores ESA indiretamente associados à sustentabilidade e desempenho (indSU).....	86
Quadro 23– Indicadores GC indiretamente associados à sustentabilidade e desempenho (indSU) .....	87
Quadro 24– Matriz de Risco Condicional para indicadores e classes dirSU .....	105
Quadro 25– Matriz de Risco Condicional para indicadores e classes indSU.....	106
Quadro 26– Algoritmos classificadores indutores de regras de decisão .....	112
Quadro 27– Matriz de Confusão .....	114
Quadro 28– Simulação da ascensão (setas verdes) e descensão (setas vermelhas) de classe dirSU, empresa E25, classe D3, método <i>ensemble bagging</i> e algoritmos classificadores .....	118
Quadro 29– Simulação da ascensão (setas verdes) e descensão (setas vermelhas) de classe indSU, empresa E25, classe I3, método <i>ensemble bagging</i> e algoritmos classificadores .....	118
Quadro 30– Simulação de ascensão de classe de D3 para D4, empresa E25, ano 2021 .....	130
Quadro 31– Simulação de ascensão de classe de I3 para I4, empresa E25, ano 2021 .....	131
Quadro 32– Simulação de descensão de classe de D3 para D2, empresa E25, ano 2021 .....	132
Quadro 33– Simulação de descensão de classe de I3 para I2, empresa E25, ano 2021 .....	133
Quadro 34– Simulação de descensão de classe de D5 para D4, empresa E07, ano 2021 .....	135
Quadro 35– Simulação de descensão de classe de I5 para I4, empresa E07, ano 2021 .....	136

Quadro 36– Práticas de sustentabilidade ou desempenho da empresa E07 relacionadas aos indicadores dirSU em destaque, ano 2021 .....	138
Quadro 37– Práticas de sustentabilidade ou desempenho da empresa E07 relacionadas aos indicadores indSU em destaque, ano 2021.....	139
Quadro 38– Práticas de sustentabilidade ou desempenho da empresa E28 relacionadas aos indicadores dirSU em destaque, ano 2021 .....	140
Quadro 39– Práticas de sustentabilidade ou desempenho da empresa E28 relacionadas aos indicadores indSU em destaque, ano 2021.....	141
Quadro 40– Práticas de sustentabilidade ou desempenho da empresa E27 relacionadas aos indicadores dirSU em destaque, ano 2021 .....	142
Quadro 41– Práticas de sustentabilidade ou desempenho da empresa E27 relacionadas aos indicadores indSU em destaque, ano 2021.....	143
Quadro 42– Indicadores mais referenciados pelas regras de decisão e respectivos valores <i>outliers</i> .....	144

## LISTA DE TABELAS

Tabela 1 – Bases CAPES de pesquisa, período 2000-2021 (posição janeiro/2022), para a revisão sistemática da literatura.....	48
Tabela 2 – Número de publicações de cada portfólio bibliográfico e para cada Eixo de pesquisa .....	48
Tabela 3– Distribuição das classes dirSU e indSU, período 2013-2021 .....	99
Tabela 4– Indicadores dirSU referenciados em 70,1% do total de referências pelas regras de decisão, algoritmo VC-DomLEM.....	104
Tabela 5– Indicadores indSU referenciados em 70,4% do total de referências pelas regras de decisão, algoritmo VC-DomLEM.....	105
Tabela 6– Algoritmos de classificação indutores de regras de decisão e respectivas medidas de aferição.....	114

## LISTA DE SIGLAS

<b>ABNT</b>	Associação Brasileira de Normas Técnicas
<b>AM</b>	Aprendizado de Máquina
<b>AMD</b>	Apoio Multicritério à Decisão
<b>ANEEL</b>	Agência Nacional de Energia Elétrica
<b>CAPES</b>	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
<b>dirSu</b>	Indicador diretamente relacionado à Sustentabilidade
<b>DRSA</b>	<i>Dominance-based Rough Set Approach</i>
<b>DSR</b>	<i>Driving Forces-State-Response</i>
<b>Dn</b>	Classe diretamente associada à Sustentabilidade (n = 1, 2, 3, 4, 5)
<b>ESA</b>	Econômico, Social, Ambiental
<b>ESG</b>	<i>Environmental, Social, Governance</i>
<b>GC</b>	Governança Corporativa
<b>GPI</b>	<i>Genuine Progress Indicator</i>
<b>GRI</b>	<i>Global Reporting Initiative</i>
<b>IA</b>	Inteligência Artificial
<b>IBASE</b>	Instituto Brasileiro de Análises Sociais e Econômicas
<b>IDH</b>	Índice de Desenvolvimento Humano
<b>IDJS</b>	Índice Dow Jones de Sustentabilidade
<b>indSu</b>	Indicador indiretamente relacionado à Sustentabilidade
<b>ISE</b>	Índice de Sustentabilidade Empresarial
<b>ISO</b>	<i>International Organization for Standardization</i>
<b>In</b>	Classe indiretamente associada à Sustentabilidade (n = 1, 2, 3, 4, 5)
<b>jMAF</b>	<i>Dominance-Based Rough Set Data Analysis Framework</i>
<b>KDD</b>	<i>Knowledge Discovery in Databases</i>
<b>LaPES</b>	Laboratório de Pesquisa em Engenharia de <i>Software</i>
<b>MCSE</b>	Manual de Contabilidade do Setor Elétrico
<b>ODS</b>	Objetivo de Desenvolvimento Sustentável
<b>ONU</b>	Organização das Nações Unidas

<b>PB</b>	Portfólio Bibliográfico
<b>PDF</b>	<i>Portable Document Format</i>
<b>ProKnow-C</b>	<i>Knowledge Development Process – Constructivist</i>
<b>PSR</b>	<i>Pressure-State-Response</i>
<b>RSA</b>	Relatório de Responsabilidade Socioambiental e Econômico-Financeiro
<b>RST</b>	<i>Rough Sets Theory</i>
<b>StArt</b>	<i>State of the Art through Systematic Review</i>
<b>TBL</b>	<i>Triple Bottom Line</i>
<b>UFSCar</b>	Universidade Federal de São Carlos
<b>VBA</b>	<i>Visual Basic for Applications</i>
<b>Weka</b>	<i>Waikato Environment for Knowledge Analysis</i>

## SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>16</b>
1.1. CONTEXTUALIZAÇÃO DO TEMA	16
1.2. SITUAÇÃO PROBLEMA	17
1.3. OBJETIVO GERAL	19
1.4. OBJETIVOS ESPECÍFICOS	19
1.5. QUESTÃO DA PESQUISA	19
1.6. ADERÊNCIA AO PPSIG	20
1.7. INEDITISMO	20
1.8. ESTRUTURA DA TESE	21
<b>2. REFERENCIAL TEÓRICO</b>	<b>22</b>
2.1 APOIO MULTICRITÉRIO À DECISÃO	22
2.2. TEORIA DOS CONJUNTOS APROXIMATIVOS	29
2.3. <i>MACHINE LEARNING</i>	37
<b>3. METODOLOGIA DE PESQUISA</b>	<b>42</b>
3.1. CLASSIFICAÇÃO DA PESQUISA	42
3.2. INSTRUMENTOS DE COLETA DE DADOS	42
3.3. ANÁLISE E TRATAMENTO DE DADOS	43
3.4. REVISÃO SISTEMÁTICA DA LITERATURA	45
3.4.1. REVISÃO DA LITERATURA – EIXO 1	49
3.4.2. REVISÃO DA LITERATURA – EIXO 2	51
3.4.3. REVISÃO DA LITERATURA – EIXO 3	53
3.5. ATUALIZAÇÃO DA REVISÃO DA LITERATURA	66
3.6. ANÁLISE DOS RESULTADOS DA REVISÃO DA LITERATURA	68
<b>4. DESENVOLVIMENTO DO MODELO CONCEITUAL DE AVALIAÇÃO INTEGRADA DA SUSTENTABILIDADE E DESEMPENHO</b>	<b>72</b>
<b>5. APLICAÇÃO DO MODELO CONCEITUAL DE AVALIAÇÃO INTEGRADA DA SUSTENTABILIDADE E DESEMPENHO</b>	<b>80</b>
5.1. DEFINIÇÃO DOS INDICADORES	80
5.2. CLASSIFICAÇÃO DAS EMPRESAS	93
5.3. APOIO MULTICRITÉRIO À DECISÃO	99
5.3.1. ANÁLISE COM APOIO DO <i>SOFTWARE</i> jMAF	99
5.3.2. ANÁLISE COM APOIO DO <i>PACKAGE ROUGHSETS</i> EM R	102
5.3. PADRÕES E PREDIÇÃO	103
5.3.1. INDUÇÃO DE REGRAS COM APOIO DO <i>SOFTWARE</i> jMAF	103

5.3.2.	INDUÇÃO DE REGRAS COM APOIO DO <i>PACKAGE ROUGHSETS</i> EM R ....	108
5.5.3.	PREDIÇÃO DE CLASSES .....	110
5.4.	MATRIZ DE SUSTENTABILIDADE .....	120
5.4.1.	ANÁLISE DE SENSIBILIDADE .....	128
5.5.	VALIDAÇÃO .....	137
<b>6.</b>	<b>CONCLUSÕES .....</b>	<b>146</b>
<b>7.</b>	<b>APRESENTAÇÕES E PUBLICAÇÕES .....</b>	<b>149</b>
	<b>REFERÊNCIAS .....</b>	<b>150</b>
	<b>ANEXO I-a .....</b>	<b>165</b>
	<b>ANEXO I-b .....</b>	<b>166</b>
	<b>ANEXO II-a .....</b>	<b>167</b>
	<b>ANEXO II-b .....</b>	<b>168</b>
	<b>ANEXO III .....</b>	<b>169</b>
	<b>ANEXO IV .....</b>	<b>170</b>
	<b>ANEXO V .....</b>	<b>171</b>
	<b>ANEXO VI .....</b>	<b>172</b>
	<b>ANEXO VII .....</b>	<b>178</b>
	<b>ANEXO VIII .....</b>	<b>183</b>
	<b>ANEXO IX .....</b>	<b>186</b>

## 1. INTRODUÇÃO

### 1.1. CONTEXTUALIZAÇÃO DO TEMA

De acordo com Pereira Neto e Cândido (2020), o desenvolvimento sustentável e a sustentabilidade envolvem temáticas interdisciplinares e presentes nas discussões das diversas áreas da ciência, nas organizações públicas e privadas, não governamentais e na sociedade como um todo. Conforme *World Commission on Environment and Development* (WCED, 1987), o desenvolvimento sustentável é o desenvolvimento que atende às necessidades do presente sem comprometer a capacidade das gerações futuras de atender às suas próprias necessidades. Em essência, o desenvolvimento sustentável é um processo de mudança em que a exploração dos recursos, a direção dos investimentos, a orientação do desenvolvimento tecnológico e as mudanças institucionais estão todas em harmonia, e aumentam o potencial atual e futuro para atender às necessidades e aspirações humanas.

Segundo Paz e Kipper (2016), a sustentabilidade baseia-se na correlação das dimensões econômica, social e ambiental e, internacionalmente, são conhecidas como *Triple Bottom Line* (TBL) da sustentabilidade de um negócio. O conceito TBL questiona e pondera sobre a necessidade de as empresas basearem suas decisões estratégicas neste tripé, em que deverão manter a sustentabilidade econômica do seu negócio ao gerenciar empresas lucrativas e geradoras de valor, também a sustentabilidade social estimulando atividades ligadas à educação, cultura, lazer, bem estar e justiça social da comunidade onde a empresa está inserida. Isso mantendo cuidados ambientais como programas de reciclagem e preservação, dentre outros aspectos. Ademais, o TBL é um índice corporativo que incorpora também, as inter-relações entre as dimensões do desenvolvimento sustentável: eco-ambiental, eco-social, socioambiental e eco-sócio-ambiental (SILVEIRA, 2013), conforme Figura 1.

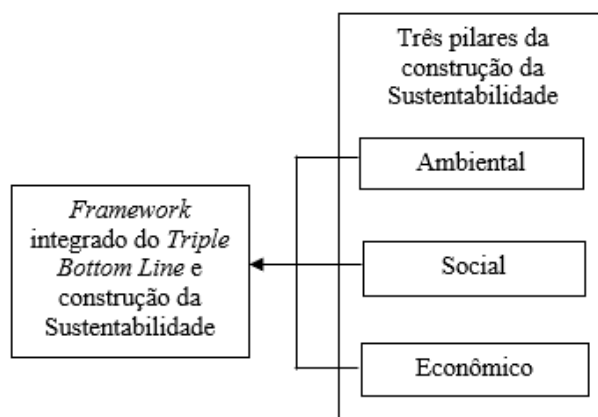
E, essas inter-relações tornam-se necessárias à compreensão de que as empresas não existem no vácuo: os negócios só podem prosperar se a sociedade prosperar, o que por sua vez, exige que salvaguadem os sistemas de suporte de vida da Terra. Com o tempo, e um mundo cada vez mais incerto, “focar no valor do sistema se tornará cada vez mais crítico para os negócios de sucesso” (ELKINGTON, 2020).

Ademais, em setembro de 2015, em Assembleia Geral das Nações Unidas, os 193 Estados Membros da ONU adotaram a Agenda 2030 para o Desenvolvimento



Sustentável, após processo global participativo de governos, sociedade civil, iniciativa privada e instituições de pesquisa.

Figura 1 – A sustentabilidade em suas dimensões



Fonte: adaptado de Goh *et al.* (2020).

A sua implementação teve início em janeiro de 2016, dando continuidade à Agenda de Desenvolvimento do Milênio no período 2000-2015, e ampliando seu escopo. A Agenda 2030 abrange o desenvolvimento econômico, a erradicação da pobreza, da miséria e da fome, a inclusão social, a sustentabilidade ambiental e a boa governança em todos os níveis, incluindo paz e segurança. Os compromissos estão presentes em 17 Objetivos de Desenvolvimento Sustentável (ODS) e 169 metas de ação global para alcance até 2030 (ODS BRASIL, 2022a, 2022b).

## 1.2. SITUAÇÃO PROBLEMA

Segundo Pereira Neto e Cândido (2020), como a sustentabilidade está relacionada não somente a aspectos ecológicos, mas também, se relaciona aos aspectos econômicos, políticos, culturais, sociais, temporais e espaciais, torna-se imprescindível a criação de instrumentos para medição, como os indicadores de sustentabilidade – “ferramentas” constituídas por uma ou mais variáveis, que podem ser relacionadas de várias maneiras. Conseqüentemente, estabelecer metas e criar instrumentos são etapas fundamentais para possibilitar a mensuração da sustentabilidade empresarial. Existem vários sistemas de indicadores de sustentabilidade empresarial utilizados no Brasil, tais como: o modelo internacional *Global Reporting Initiative* (GRI); Índice Dow Jones de Sustentabilidade

(IDJS); o modelo do Instituto Brasileiro de Análises Sociais e Econômicas (IBASE); o modelo Ethos; e o Índice de Sustentabilidade Empresarial (ISE).

Paz e Kipper (2016) também apontam vários modelos dentre outros, para medir a sustentabilidade nas organizações públicas e privadas dos mais diversos segmentos: *Pressure-State-Response* (PSR) (dimensão ambiental); *Driving Forces-State-Response* (DSR) (dimensões social, ambiental, institucional e econômica); *Genuine Progress Indicator* (GPI) (dimensões social e econômica); Banco Mundial (dimensões social, ambiental, econômica e cultural); Índice de Desenvolvimento Humano (IDH) (dimensões social, econômica, cultural e política); Barômetro da Sustentabilidade (dimensões social e ambiental); Painel de Sustentabilidade (dimensões social, ambiental, institucional e econômica); entre outros.

Segundo Lameira *et al.* (2013), em suas pesquisas com 205 empresas abertas brasileiras, no período entre 2005 e 2009, demonstraram que, os resultados foram favoráveis ao estabelecimento da associação e do discurso que defende que sustentabilidade e melhores desempenhos podem e devem estar integrados no processo de busca do crescimento sustentável. Ademais, a ISO 26000 que aborda as Diretrizes sobre responsabilidade social (ABNT, 2010), estabelece que, “Uma organização pode exercer sua influência sobre outros, tanto para fortalecer os impactos positivos no desenvolvimento sustentável como para minimizar os impactos negativos, ou ambos os casos”. Entre os métodos propostos para exercer influência, encontra-se “promover boas práticas”. A ISO 26000 ainda recomenda que, uma organização “implemente medidas de eficiência no uso de recursos para reduzir seu uso de energia, água e outros recursos, considerando indicadores de melhores práticas e outros padrões de referência”; e, como benefício da responsabilidade social, encontra-se “melhoria das práticas de gestão de risco da organização”.

Por conseguinte, a Agência Nacional de Energia Elétrica (ANEEL), por meio da Resolução Normativa nº 605, de 11/03/2014, aprovou o Manual de Contabilidade do Setor Elétrico (MCSE), em que ficou estabelecido dentre outros objetivos, contribuir para a otimização da performance socioambiental por meio da explicitação dos gastos oriundos do cumprimento da Política Nacional de Meio Ambiente, necessários à conformidade ambiental e sustentabilidade das concessões atribuídas pela União Federal, visando a elaboração do Relatório de Responsabilidade Socioambiental e Econômico-Financeiro (RSA) (ANEEL, 2014, 2015). O MCSE relaciona mais de duas centenas de indicadores quantitativos relativos às dimensões geral, governança corporativa,

econômico-financeira, social e setorial, e ambiental para empresas outorgadas (geração, transmissão e distribuição de energia elétrica) exigidos e/ou sugeridos pela ANEEL, além de indicadores de desempenho ambiental específicos para empresas de geração e transmissão e/ou distribuição (ANEEL, 2015). O MCSE ainda menciona indicadores qualitativos, tais como, aspectos de governança corporativa (ANEEL, 2014, 2015).

Neste contexto e após revisão sistemática da literatura, identificou-se a ausência de avaliação integrada da sustentabilidade e desempenho de empresas concessionárias de distribuição de energia elétrica no Brasil, que permita o *benchmarking* das empresas, a gestão efetiva dos riscos econômicos, sociais, ambientais e de governança corporativa, e o conhecimento e disseminação de boas práticas de sustentabilidade e desempenho.

### 1.3. OBJETIVO GERAL

Propor modelo de avaliação integrada da sustentabilidade e desempenho das empresas concessionárias de distribuição de energia elétrica, visando o *benchmarking*, a identificação de melhores práticas das empresas e a gestão de riscos.

### 1.4. OBJETIVOS ESPECÍFICOS

- Definir os indicadores de sustentabilidade e desempenho e classificar as empresas.
- Analisar os indicadores estabelecidos com emprego do Apoio Multicritério à Decisão, por meio de *Rough Sets Theory* e *Dominance-based Rough Set Approach*.
- Classificar as empresas com emprego de *Machine Learning*, regras de decisão “if ... then ...”, para a predição das classes de sustentabilidade e de desempenho com base em dados históricos de indicadores e para simulação de riscos.
- Elaborar *benchmarking* das empresas: Matriz Sustentabilidade visando a identificação e disseminação de conhecimento sobre melhores práticas.
- Validar modelo de avaliação.

### 1.5. QUESTÃO DA PESQUISA

Dada a situação problema e visando o objetivo geral e os específicos descritos anteriormente, apresenta-se a seguinte questão de pesquisa: “Como estender a análise sobre a sustentabilidade e desempenho – econômico, social, ambiental, governança corporativa, de empresas de energia a partir dos seus indicadores de sustentabilidade e

desempenho, visando o *benchmarking* das empresas e a identificação de melhores práticas?”.

#### 1.6. ADERÊNCIA AO PPSIG

A tese encontra-se aderente à linha de pesquisa do Programa de Pós-Graduação em Sistemas de Gestão Sustentáveis da Universidade Federal Fluminense (PPSIG/UFF) – “apoio à decisão em organizações sustentáveis”, vis-à-vis à revisão sistemática da literatura em bases multidisciplinares da CAPES.

Ademais, no contexto deste estudo, são reunidos saberes sobre: desenvolvimento sustentável e sustentabilidade; Apoio Multicritério à Decisão; Teoria dos Conjuntos Aproximativos/princípio de Dominância, teoria matemática para tratamento da imprecisão e incerteza de dados; *Machine Learning* ou Aprendizado de Máquina, subcampo da Inteligência Artificial; Tecnologia da Informação, por meio de Mineração de textos e de dados, e Recuperação de informação; Ciência de Dados; desempenho; gestão de empresas; gestão de riscos; e auditoria.

#### 1.7. INEDITISMO

Como consequência da revisão sistemática da literatura e da identificação de lacunas nos estudos reunidos (item 3.6), a contribuição desta pesquisa encontra-se na avaliação integrada da sustentabilidade e do desempenho – econômico, social, ambiental e governança corporativa, de empresas concessionárias de distribuição de energia elétrica no Brasil, por meio de *benchmarking*, com Apoio Multicritério à Decisão e modelo preditivo para a classificação e gestão de riscos das empresas. Neste caso, trata-se de *benchmarking* “colaborativo”, isto é, quando se compara frente a competidores visando o desenvolvimento de uma atmosfera de aprendizagem e compartilhamento de conhecimento, e iniciado por um agente de comum consenso (FERNANDEZ *et al.*, 2001; ANAND; KODALI, 2008). Permite também, a gestão de riscos nas ascensões e descensões das classificações das empresas por meio de Aprendizado de Máquina, com base em dados históricos de indicadores, inferência de regras de decisão e simulações.

E, ao gestor público ou privado, há a oportunidade de conhecer melhores práticas das empresas ao identificar aquelas consideradas *benchmark* e, utilizar este conhecimento como subsídio à elaboração e/ou aperfeiçoamento de políticas públicas de desenvolvimento sustentável para o setor de energia elétrica.

## 1.8. ESTRUTURA DA TESE

Este documento está dividido nos seguintes capítulos: 1, Introdução, com a contextualização do tema, descrição do problema e questão de pesquisa, objetivos do estudo e aderência ao PPSIG; 2, Referencial Teórico, sobre Apoio Multicritério à Decisão, Teoria dos Conjuntos Aproximativos/princípio de Dominância e *Machine Learning*; 3, Metodologia de Pesquisa, com a classificação da pesquisa, instrumentos para coleta de dados, revisão sistemática da literatura e bibliometria, análise dos resultados da revisão da literatura; 4, Desenvolvimento do modelo conceitual de avaliação integrada da sustentabilidade e desempenho, com a proposição de modelo conceitual em seis etapas; 5, Aplicação do modelo conceitual de avaliação integrada da sustentabilidade e desempenho, com a aplicação do modelo conceitual em seis etapas às empresas de distribuição de energia elétrica; 6, Conclusões, sobre os principais achados da pesquisa e perspectivas de estudos futuros; e, 7, Apresentações e publicações, sobre os estudos já publicados relacionados ao tema de pesquisa.

## 2. REFERENCIAL TEÓRICO

### 2.1 APOIO MULTICRITÉRIO À DECISÃO

O Apoio Multicritério à Decisão (AMD) pode ser definido como a atividade de um analista de decisões que, com base em modelos claramente apresentados, mas não necessariamente formalizados, ajuda na obtenção de elementos de resposta às questões de um agente de decisão no decorrer de um processo. Esses elementos tem como objetivo esclarecer cada decisão e, normalmente, recomendá-la ou simplesmente, favorecê-la. Trata-se de um comportamento natural que aumenta a coerência entre a avaliação do processo e os objetivos e o sistema de valores, a serviço dos quais o agente de decisão se posiciona (GOMES *et al.*, 2004; GOMES, 2007; GOMES; RANGEL, 2009a, 2009b). E, como decisão multicritério, aquela que envolve problemas relativos a um conjunto finito de objetos (alternativas, ações, soluções) que podem ser avaliadas por um conjunto finito de critérios (atributos, características, variáveis) (SLOWINSKI *et al.*, 2012).

Consequentemente, o ser humano é instado a tomar decisões, usando parâmetros quantitativos ou qualitativos, sendo os parâmetros quantitativos normalmente de mensuração mais fácil que os qualitativos. Este mesmo decisor, ou tomador de decisão ou agente de decisão (*decision maker* – DM), é o responsável por realizar a decisão; pode ser uma pessoa, um grupo, um comitê ou uma companhia. E, muitas vezes, tem de vislumbrar as consequências das decisões em um meio ambiente mutável e sujeito a condições que o decisor não pode controlar, com incertezas, imprecisão e/ou ambiguidade (GOMES; GOMES, 2019).

As decisões, normalmente, buscarão: minimizar perdas, maximizar ganhos e criar uma situação em que comparativamente, o decisor julgue que houve elevação (houve ganho) entre o estado da natureza em que se encontrava e o estado em que irá encontrar-se após implementar a decisão (GOMES; GOMES, 2019).

Segundo Gomes e Gomes (2019), a análise de cenários tem sido utilizada ao longo da história da humanidade, notadamente no campo militar, mas também como elemento fundamental para apoio à tomada de decisão civil. Tal análise caracteriza o que se denomina pensamento ou planejamento estratégico. A ideia primordial da análise de cenários é, após um estudo detalhado dos vários aspectos do problema de decisão que se pretende resolver, efetuar a construção de diferentes contextos (cenários) alternativos passíveis de materialização.

Consequentemente, torna-se necessário estruturar um possível inter-relacionamento entre estratégias de implementação de alternativas, de tal modo que, à medida que a realidade seja desvendada, tanto os cenários elaborados previamente como as estratégias sejam reavaliados (GOMES; GOMES, 2019).

Cenários formam uma ferramenta de planejamento poderosa, principalmente porque o futuro é imprevisível. Usar cenários é ensaiar para o futuro antes de o futuro chegar. Contudo, o nosso conhecimento apresenta “limites”. Assim, a incerteza é onde nos movemos, não só na ação, mas também no conhecimento. E a condição humana é marcada por duas grandes incertezas: a incerteza cognitiva – é preciso aceitar pensar com certa incerteza, porque pensamos de forma parcelar (vemos apenas fenômenos separados e somos incapazes de compreender suas relações); e a incerteza histórica – ligada ao caráter caótico da história humana (MORIN; VIVERET, 2013; GOMES; GOMES, 2019).

Pode-se mesmo dizer rigorosamente que, quase todos os nossos conhecimentos são apenas prováveis; e no pequeno número de coisas que podemos saber com certeza, mesmo nas ciências matemáticas, os principais meios para se chegar à verdade – a indução e a analogia – são fundados nas probabilidades (LAPLACE, 2010).

E, sem incerteza, não haveria esperança, ética e liberdade de escolha. É apenas porque não se sabe o que o futuro reserva é que se pode ter esperança. É só porque não se sabe exatamente os resultados futuros das escolhas, que uma escolha pode ser livre e, pode representar um verdadeiro dilema ético. Além disso, há muita incerteza no mundo, e uma das escolhas mais básicas é se aceita essa incerteza como um fato ou tenta-se fugir dela. Assim, o *insight* chave é aplicar estatística e a teoria da probabilidade sempre que puder – se não para calcular soluções numéricas, então para estruturar e orientar o processo de tomada de decisão (HASTIE; DAWES, 2010).

Portanto, entender a natureza de cada problema sob diversos ângulos – adquirir visão multidimensional, definindo as causas do problema é, assim, elemento-chave para uma boa solução do mesmo (GOMES; GOMES, 2019). E o ser humano é, ao mesmo tempo, biológico, psíquico, social, afetivo e racional; a sociedade comporta as dimensões histórica, econômica, sociológica e religiosa. Consequentemente, o conhecimento pertinente deve reconhecer o caráter multidimensional e nele inserir estes dados: não apenas não se poderia isolar uma parte do todo, mas as partes umas das outras. Nossa realidade não é outra senão nossa ideia da realidade (MORIN, 2013).

E a arrogância epistêmica tem um efeito duplo: superestima-se o que se sabe e subestima-se a incerteza, comprimindo a gama de possíveis estados incertos (ou seja, reduzindo o espaço do desconhecido) (TALEB, 2010). Muitas vezes, utiliza-se da heurística para encontrar respostas adequadas, ainda que geralmente imperfeitas, para perguntas difíceis – uma pergunta-alvo é substituída por uma “pergunta heurística”, uma pergunta mais simples (KAHNEMAN, 2012).

De acordo com Gomes; Gomes (2019), a teoria da decisão pode ser prescritiva ou normativa, no sentido de pretender ajudar as pessoas a tomarem decisões melhores, em face de suas preferências básicas. Parte-se do pressuposto de que os indivíduos são capazes de expressar suas preferências básicas e são racionais quando enfrentam situações de decisão simples. Consequentemente, torna-se possível a resolução de problemas de decisão mais complexos. E o ser humano tem uma capacidade cognitiva limitada para compreender todos os sistemas no seu entorno e processar todas as informações que recebe. Assim, define-se teoria da decisão como o conjunto de procedimentos e métodos de análise que procuram assegurar a coerência, a eficácia e a eficiência das decisões tomadas em função das informações disponíveis, antevendo cenários possíveis e que trata de escolhas entre alternativas. Pode usar ferramentas matemáticas ou não.

As preferências são relações binárias entre dois objetos ou alternativas. Há quatro categorias principais de preferências: a) indiferença: quando há razões claras e positivas que justifiquem a equivalência entre duas alternativas; b) forte ou estrita: quando há razões claras e positivas que justifiquem preferência significativa em favor de uma das duas alternativas; c) fraca: quando existem razões claras e positivas que não impliquem preferência estrita em favor de uma das duas alternativas; d) incomparabilidade: quando não há razões claras e positivas que justifiquem uma das três situações anteriores (GOMES, 2007).

De acordo com Almeida *et al.* (2015) e Gomes; Gomes (2019), uma relação binária  $H$  sobre um conjunto  $A$  é uma partição do conjunto de pares ordenados de  $A \times A$ . Essa partição cria então dois subconjuntos de  $A \times A$ . O primeiro, identificado por  $B$ , é o subconjunto de pares ordenados que estão na relação  $H$ . O segundo, identificado por  $B^-$ , se for igual a  $\{A \times A\} - B$ , inclui os pares ordenados que não estão na relação  $H$ . Em face da importância das relações binárias para a modelagem das preferências de um agente de decisão, são apresentadas a seguir as propriedades clássicas de uma relação binária  $H$  sobre um conjunto  $A$ : i) reflexividade –  $\forall a \in A, a H a$  ou  $(a, a) \in B$ ; ii) irreflexividade –



$\forall a \in A$ , não  $[a H a]$  ou  $(a, a) \notin B$ ; iii) simetria –  $\forall a, b \in A$ ,  $a H b \Rightarrow b H a$  ou  $(a, b) \in B \Rightarrow (b, a) \in B$ ; iv) assimetria –  $\forall a, b \in A$ ,  $a H b \Rightarrow$  não  $[b H a]$  ou  $(a, b) \in B \Rightarrow (b, a) \notin B$ ; v) transitividade –  $\forall a, b, c \in A$ ,  $[a H b \text{ e } b H c] \Rightarrow a H c$  ou  $(a, b) \in B \text{ e } (b, c) \in B \Rightarrow (a, c) \in B$ .

Uma ferramenta que viabiliza a estruturação das ações para a solução de problemas, é o modelo de Estrela Decisória: escada de oito degraus que permite a subida até o último degrau (solução do problema), conforme Figura 2 (GOMES; GOMES, 2019).

Figura 2 – Modelo de Estrela Decisória para a solução de problemas



Fonte: Gomes; Gomes (2019).

O AMD consiste em um conjunto de métodos e técnicas para auxiliar ou apoiar pessoas e organizações a tomarem decisões, sob a influência de uma multiplicidade de critérios. O AMD tem como princípio buscar o estabelecimento de uma relação de preferências (subjetivas) entre as alternativas que estão sendo avaliadas/priorizadas/ordenadas sob a influência de vários critérios, no processo de decisão. E a abordagem multicritério tem como característica: i) processos decisórios complexos, e neles existem vários atores envolvidos que definem os aspectos relevantes do processo de decisão; ii) cada ator tem a sua subjetividade (juízo de valores); iii)

reconhece os limites da objetividade e considera as subjetividades dos atores; iv) tem como pressuposto que o problema não está claramente definido nem bem estruturado (GOMES; GOMES, 2019).

De acordo com Gomes; Gomes (2019), a busca da solução de um problema frequentemente ocorre em ambiente em que os critérios são conflitantes, onde o ganho de um critério pode causar perda em outro. Na escolha da solução final deve-se levar em conta o compromisso das diversas relações de troca intermediárias dos procedimentos adotados. E cabem as seguintes observações: i) uma alternativa é considerada dominada quando existe outra alternativa que a supera em um ou mais critérios e iguala-se aos demais; o princípio da dominância é usado para eliminar uma alternativa que seja claramente inferior a outra alternativa; ii) superioridade de Pareto: o bem-estar social associado a um estado A é superior ao de outro estado B, se, e somente se, existe em A pelo menos um indivíduo com bem-estar maior que em B e não existe em B outro indivíduo que possua um bem-estar superior em A; um estado é superior a outro se é possível aumentar o bem-estar de pelo menos um indivíduo sem prejudicar os demais; iii) ótimo de Pareto: o bem-estar de uma sociedade é máximo se não existe outro estado em que seja possível aumentar o bem-estar de um indivíduo sem diminuir o bem-estar dos demais indivíduos dessa sociedade; não existe como melhorar o bem-estar de um indivíduo sem prejudicar o bem-estar de pelo menos outro indivíduo.

E a solução eficiente, não dominada, fronteira de eficiência, eficiência de Pareto, ou ótimo de Pareto é aquela que pode ser obtida de forma que a alternativa escolhida atinja um valor amplo em todos os critérios e não possua um decréscimo simultâneo (um valor dominado por outra alternativa) em nível inaceitável em qualquer dos demais critérios que estão sendo utilizados no processo de avaliação das alternativas. O decisor pode não estar interessado somente em identificar a melhor alternativa, mas também em saber o quanto a melhor alternativa tem um valor global suficientemente bom para ser tomada como decisão (GOMES; GOMES, 2019).

Sob forma geral, uma função multicritério de decisão (FD) pode ser assim descrita (Gomes; Gomes, 2019):

$G$  = objetivo e/ou objetivos a serem alcançados;

Conjunto  $A = \{a_1, a_2, \dots\}$ , conjunto de alternativas que deverão ser avaliadas, comparadas ou analisadas e, se possível e necessário, priorizadas pelo decisor;

Conjunto  $C = \{c_i, \text{ em que } i \text{ varia de } 1 \text{ até } n\}$ , conjunto de critérios a serem usados para avaliação das alternativas;

Conjunto  $K$  = em que  $k_j$  representa o número de participantes do processo de decisão;

Conjunto  $U$  = em que  $u_i$  representa a informação individual;

Conjunto  $O = \{o_j, \text{ em que } j \text{ varia de } 1 \text{ até } m\}$ , conjunto de atributos;

Conjunto  $W = \{w_1, w_2, \dots, w_k\}$ , conjunto de pesos atribuídos aos critérios e/ou atributos, sendo  $w_k$  função da preferência do decisor por  $c_i$  ou  $o_j$ ;

$R$  = restrições a serem observadas;

$\Omega$  = conjunto de possíveis estados da natureza ou possíveis eventos;

$p(\Omega)$  = a probabilidade; em sua ausência, é substituída pela possibilidade, associada a um evento específico  $\Omega$ ;

$H(A, \Omega, p(\Omega))$  = benefícios esperados da alternativa e/ou consequências;

$FD = G \otimes [A \otimes R] \otimes [W \otimes C \otimes O] \otimes [\Omega \otimes H \otimes p(\Omega)] \otimes [K \otimes U]$ , em que o símbolo  $\otimes$  significa “associado a”.

A relação  $A \otimes R$  busca eliminar as alternativas que não superam as restrições. A relação  $W \otimes C \otimes O$  associa os critérios e atributos com os respectivos pesos, modela as preferências do decisor. A relação  $[\Omega \otimes H \otimes p(\Omega)]$  quantifica a incerteza e/ou a imprecisão do processo de decisão e a agregação de valor das alternativas.

Ainda conforme Gomes; Gomes (2019), no contexto do apoio à decisão, o resultado pretendido em determinado problema pode ser identificado entre quatro tipos de problemática de referência, descritos a seguir: i) problemática tipo “alfa” ( $P\alpha$ ) – tem como objetivo esclarecer a decisão pela escolha de um subconjunto tão restrito quanto possível, considerando a escolha final de uma única ação. Esse conjunto conterá as melhores ações ou satisfatórias. O resultado pretendido é, portanto, uma escolha ou um procedimento de seleção; ii) problemática tipo “beta” ( $P\beta$ ) – tem como objetivo esclarecer a decisão por uma triagem resultante da alocação de cada ação a uma categoria (ou classe). As diferentes categorias são definidas a priori com base em normas aplicáveis ao conjunto de ações. O resultado pretendido é, portanto, uma triagem ou um procedimento de classificação; iii) problemática tipo “gama” ( $P\gamma$ ) – tem como objetivo esclarecer a decisão por um arranjo obtido pelo reagrupamento de todas ou de parte (as mais satisfatórias) das ações em classes de equivalência. Essas classes são ordenadas de modo completo ou parcial, conforme as preferências. O resultado pretendido é, portanto, um arranjo ou um procedimento de ordenação; iv) problemática tipo “delta” ( $P\delta$ ) – tem como objetivo esclarecer a decisão por uma descrição, em linguagem apropriada, das ações e de suas

consequências. O resultado pretendido é, portanto, uma descrição ou um procedimento cognitivo.

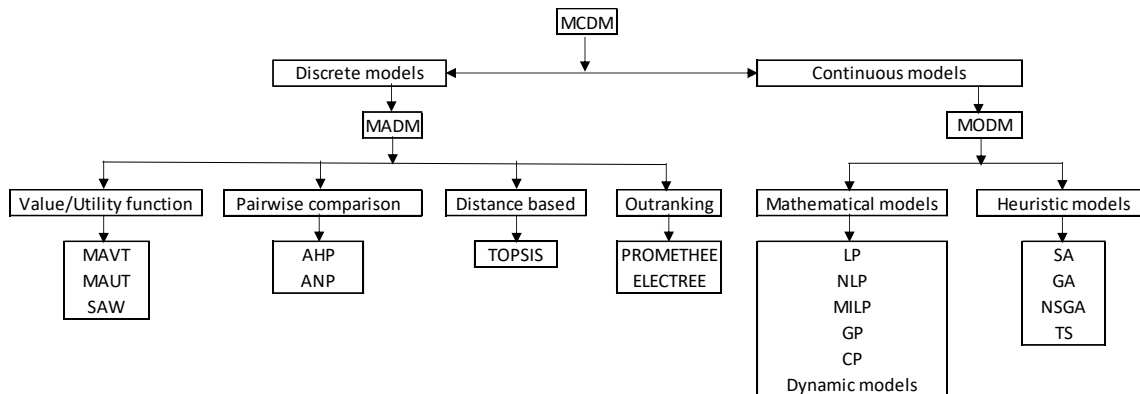
Segundo Gebre *et al.* (2021), MCDM, *Multi-Criteria Decision-Making*, é um termo genérico usado para todos os métodos úteis para resolver problemas de decisão complexos e conflitantes, em diferentes campos de aplicação. Permite a utilização de fatores de avaliação quantitativos e qualitativos. Os métodos MCDM podem ser classificados em duas grandes classes: MADM, *Multi-Attribute Decision-Making*, e MODM, *Multiple-Objective Decision-Making* – Figura 3.

Os métodos MADM são adequados para seleção de um número limitado de alternativas e classificação de preferências. A avaliação é baseada em alternativas de decisão predeterminadas em relação aos atributos ponderados, ou seja, o espaço de decisão é discreto. Alternativas são as diferentes escolhas ou preferências disponíveis para o tomador de decisão, e são consideradas finitas; devem ser selecionadas, priorizadas e, finalmente, classificadas de acordo com os critérios, decisões ou objetivos declarados. Existem diferentes tipos de métodos MADM (ou MCDM discreto) disponíveis, tais como: função de Valor/Utilidade, por exemplo, *Multi-Attribute Value Theory*, MAVT, *Multi-Attribute Utility Theory*, MAUT, e *Simple Additive Weighting*, SAW; comparação par a par, por exemplo, *Analytic Hierarchy Process*, AHP, e *Analytic Network Process*, ANP; com base em distância, como exemplo, *Technique for Order of Preference by Similarity to Ideal Solution*, TOPSIS; superação, como exemplos, *Preference Ranking Organization Method for Enrichment Evaluation*, PROMETHEE, e *Elimination and Choice Expressing Reality*, ELECTREE (GEBRE *et al.*, 2021; YALCIN *et al.*, 2022; TAHERDOOST; MADANCHIAN, 2023).

Os métodos MODM são preferencialmente utilizados para problemas de otimização contínua onde o número de alternativas é infinito, ou seja, o espaço de decisão é contínuo. As alternativas não são pré-determinadas mas, em vez disso, um conjunto de funções objetivo é otimizado e sujeito a um conjunto de restrições. Os métodos MODM são ainda agrupados em modelos de programação matemática e algoritmos heurísticos baseados em tempo computacional e solução. Por exemplo, os métodos matemáticos incluem programação linear (*linear programming*, LP), programação não linear (*non-linear programming*, NLP), programação linear inteira mista (*mixed integer linear programming*, MILP), programação de metas (*goal programming*, GP), programação de compromisso (*compromise programming*, CP) e programação dinâmica (*dynamic programming*). Os métodos heurísticos incluem aqueles como simulação de recozimento

(*simulating annealing*, SA), algoritmo genético (*genetic algorithm*, GA), algoritmo genético de classificação não dominado (*non-dominated sorting genetic algorithm*, NSGA) e *tabu search* (TS) (GEBRE *et al.*, 2021; YALCIN *et al.*, 2022; TAHERDOOST; MADANCHIAN, 2023).

Figura 3– Classificação de Métodos de Apoio Multicritério à Decisão (MCDM)



Fonte: Gebre *et al.* (2021).

Ademais, Roy (1968), Keeney; Raiffa (1976), Saaty (1980), Roy; Bouyssou (1991), Gomes; Lima (1992), Bana e Costa; Vansnick (1993), Figueira *et al.* (2005), Almeida (2011, 2013), Almeida *et al.* (2015), Gomes; Gomes (2019) e Thakkar (2021) apresentam modelos de processos de decisão, métodos para a elaboração de cenários, técnicas qualitativas de auxílio à decisão, abordagens para decisões sob incerteza e decisões sob risco, e outros métodos quantitativos em apoio à decisão.

## 2.2. TEORIA DOS CONJUNTOS APROXIMATIVOS

De acordo com Pawlak (1982, 1991), Pawlak; Slowinski (1994), Pawlak *et al.* (1995) e Slowinski *et al.* (2012), a Teoria dos Conjuntos Aproximativos (*Rough Sets Theory*, RST) teve origem com Zdzislaw Pawlak, no início dos anos 1980, como ferramenta matemática para o tratamento da imprecisão e incerteza de dados. A abordagem que a RST proporciona é de fundamental importância para a Inteligência Artificial (IA) e as ciências cognitivas, especialmente nas áreas de Aprendizado de Máquina (*Machine Learning*), aquisição de conhecimento, análise de decisão, descoberta de conhecimento em bancos de dados, sistemas especialistas, sistemas de suporte à decisão, raciocínio indutivo e reconhecimento de padrões. A RST não compete com a Lógica Nebulosa (*Fuzzy*) (EKEL *et al.*, 2020), com a qual é frequentemente contrastada,

mas a complementa. De qualquer forma, essas teorias são abordagens independentes do conhecimento imperfeito. Uma das principais vantagens da RST é que ela não precisa de informações preliminares ou adicionais sobre os dados, como distribuição de probabilidade ou grau de pertinência na teoria dos conjuntos nebulosos.

A RST tem como conceito fundamental, a relação de “indiscernibilidade” (*indiscernibility relation*), isto é, relação que identifica objetos com a mesma propriedade. Assim, objetos que possuem as mesmas propriedades são “indiscerníveis” e, conseqüentemente, são tratados como idênticos ou similares (“grânulos”). Esses grânulos são conhecidos como “conjuntos elementares” (*elementary sets*) e constituem os “conceitos” (*concepts*) de conhecimento sobre algo de interesse (PAWLAK, 2000). A granularidade pode ser origem de inconsistências nas decisões, dada à ambigüidade para explicar e prescrever com base em informação inconsistente (PAWLAK; SLOWINSKI, 1994).

Em geral, os dados remetem a números, e números são comumente vistos como o resultado de medições. E, relacionados à precisão dos dados, há vieses, erros e aproximações. Assim, os dados que são utilizados para representar resultados são considerados “aproximados” (*rough*) e os instrumentos para produzir as descrições são imperfeitos, por exemplo, contar pessoas em uma estação de trem de forma individual (ROY, 1989).

Conforme Pawlak (1991), há um exemplo com conceitos que serão posteriormente definidos: dado um conjunto  $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ , de brinquedos. O conjunto é classificado de acordo com a cor (vermelha, azul, amarela), forma (quadrada, redonda, triangular) e tamanho (pequeno, grande), conforme Quadro 1 a seguir.

Quadro 1– Exemplo de base de conhecimento

cor	forma	tamanho
vermelha: $x_1, x_3, x_7$	redonda: $x_1, x_5$	grande: $x_2, x_7, x_8$
azul: $x_2, x_4$	quadrada: $x_2, x_6$	pequeno: $x_1, x_3, x_4, x_5, x_6$
amarela: $x_5, x_6, x_8$	triangular: $x_3, x_4, x_7, x_8$	

Fonte: Pawlak (1991).

São definidas três relações de equivalência,  $R_1$ ,  $R_2$  e  $R_3$ , para cor, forma e tamanho, respectivamente, tendo as seguintes classes de equivalência:

$$U/R_1 = \{\{x_1, x_3, x_7\}, \{x_2, x_4\}, \{x_5, x_6, x_8\}\},$$

$$U/R_2 = \{\{x_1, x_5\}, \{x_2, x_6\}, \{x_3, x_4, x_7, x_8\}\},$$

$$U/R_3 = \{\{x_2, x_7, x_8\}, \{x_1, x_3, x_4, x_5, x_6\}\},$$

as quais são conceitos (categorias) elementares na base de conhecimento  $K = (U, \{R_1, R_2, R_3\})$ . Por exemplo, as categorias básicas são interseções de categorias elementares:

a)  $\{x_2, x_4\} \cap \{x_2, x_6\} = \{x_2\}$ , é a categoria  $\{R_1, R_2\}$  básica “azul e quadrada”;

b)  $\{x_1, x_5\} \cap \{x_1, x_3, x_4, x_5, x_6\} = \{x_1, x_5\}$ , é a categoria  $\{R_2, R_3\}$  básica “redonda e pequena”;

c)  $\{x_2, x_4\} \cap \{x_1, x_5\} = \emptyset$ , pois a categoria “azul e redonda” não existe na base de conhecimento.

Conforme Pawlak (1991), o conhecimento fundamenta-se na habilidade em classificar objetos reais ou abstratos. Dado um conjunto finito  $U \neq \emptyset$  de objetos (universo), um subconjunto  $X \subseteq U$  é referenciado como um conceito ou categoria em  $U$  e, qualquer família de conceitos em  $U$  pode ser denominada como “conhecimento abstrato”, ou de forma abreviada, conhecimento sobre  $U$ . Ademais, não é usual tratar com uma única classificação, mas, com uma família de classificações básicas, como cor, temperatura, forma, etc. sobre  $U$ . Assim, relações de equivalência e classificações têm o mesmo significado, indistintamente.

A relação de equivalência também é conhecida como “relação de indiscernibilidade”; classes de equivalência são conhecidas como “conjuntos elementares”. Consequentemente, se  $R$  é uma relação de equivalência sobre  $U$ , então  $U/R$  significa a família de todas as classes de equivalência de  $R$  (GRZYMALA-BUSSE, 1988; PAWLAK, 1991; ZIARKO, 1993).

Se  $P \subseteq R$  e  $P \neq \emptyset$ , então  $\cap P$  (interseção de todas as relações de equivalência pertencentes a  $P$ ) é também uma relação de equivalência, é indicado por  $IND(P)$  e é conhecido como “relação de indiscernibilidade” sobre  $P$ . Assim,  $U/IND(P)$  representa a família de todas as classes de equivalência da relação de equivalência  $IND(P)$ , e significa o conhecimento associado com a família de relações de equivalência de  $P$  (PAWLAK, 1991).

Conceitua-se um Sistema de Informação como  $K = (U, A)$ , onde  $U$  é um conjunto finito e não-vazio de objetos;  $A$  é um conjunto finito e não-vazio de atributos, tal que,  $a: U \rightarrow V_a$  para todo  $a \in A$ ,  $V_a$  é o conjunto de valores que pode ser atribuído ao atributo  $a$ . De forma complementar, dado um sistema de informação  $(U, A)$  e para qualquer  $B \subseteq A$ ,

uma relação de indiscernibilidade (equivalência ou classificação, indistintamente)  $R_B$  é definida como (RIZA *et al.*, 2014), Equação (1):

$$R_B(x,y) = \{(x,y) \in U^2 | \forall a \in B, a(x) = a(y)\} \quad (1)$$

Se  $(x,y) \in R_B(x,y)$ , então  $x$  e  $y$  tem exatamente os mesmos valores para os atributos em  $B$ . Conforme Pawlak (1991), nem sempre é possível expressar exatamente algum conjunto de objetos com o conhecimento disponível. Assim, torna-se possível expressar um conjunto de objetos por outros dois subconjuntos: aproximações inferior (“lower”,  $\underline{R}$ ) e superior (“upper”,  $\bar{R}$ ), dado  $X \subseteq U$  e uma relação de equivalência  $R$ , Equação (2):

$$\underline{R}X = \cup \{Y \in U | R: Y \subseteq X\}; \bar{R}X = \cup \{Y \in U | R: Y \cap X \neq \emptyset\} \quad (2),$$

em que  $\underline{R}X$  é o subconjunto de elementos de  $U$ , que certamente podem ser classificados como elementos de  $X$ ;  $\bar{R}X$  é o subconjunto de elementos de  $U$ , que possivelmente podem ser classificados como elementos de  $X$ . Assim, um conjunto é “aproximativo” (“rough”) com relação a  $R$ , se e somente se,  $\underline{R}X \neq \bar{R}X$ .

De acordo com Pawlak (1991), do conceito de Sistema de Informação ( $K = (U, A)$ ), pode-se obter o conceito de Tabela de Decisão: quando  $A$  é formado por dois subconjuntos  $C$  e  $D$ , atributos de condição e de decisão, respectivamente, e ( $C, D \subset A$ ); ou,  $T = (U, A, C, D)$ . As classes de equivalência das relações  $R_C$  e  $R_D$  são conhecidas como classes de condição e decisão, respectivamente. Para cada  $x \in U$ , associa-se uma função  $d_x: A \rightarrow V$ , tal que,  $d_x(a) = a(x)$ , para todo  $a \in C \cup D$ ; a função  $d_x$  é conhecida como “regra de decisão” (em  $T$ ). Segundo Pawlak (1991) e Riza *et al.* (2014), define-se região “positiva-B” como o conjunto de objetos em  $U$  que certamente pode ser classificado usando os atributos condicionais  $B$ , Equação (3):

$$POS_B = \cup_{x \in U} \underline{R}X \quad (3)$$

Ademais, pode-se computar o grau de dependência do atributo de decisão em relação ao conjunto de atributos de condição  $B$ , Equação (4):

$$\gamma_B = |POS_B| / |U| \quad (4)$$

Assim, uma tabela de decisão  $A$  é considerada consistente se  $\gamma_A = 1$ .

De acordo com Riza *et al.* (2014), um superreduto (*superreduct*) é um conjunto de atributos  $B \subset A$  tal que  $R_B = R_A$ , onde  $R_B$  e  $R_A$  são relações de indiscernibilidade definidas por  $B$  e  $A$ , respectivamente. Se essa relação for também mínima, trata-se de um “reduto” (“reduct”). A interseção de todos os redutos é um “núcleo” (“core”), de um sistema de informação.



E, segundo Slowinski *et al.* (2012), há três tipos de atributos: a) indispensável, atributo incluído no núcleo; b) permutável, atributo incluído em algum reduto mas não no núcleo; c) redundante, atributo não incluído em qualquer reduto.

Conforme Pawlak (2000), há um exemplo de tabela de decisão, Quadro 2. Esta tabela é composta de seis lojas e quatro atributos (aspectos quantitativos ou qualitativos): E – autonomia dos vendedores; Q – qualidade da mercadoria; L – localização com trânsito intenso; P – resultado (lucro ou prejuízo). E, Q e L representam atributos de condição; P representa um atributo de decisão.

Quadro 2– Exemplo com seis lojas

Loja	E	Q	L	P
1	Alta	Boa	Não	Lucro
2	Média	Boa	Não	Prejuízo
3	Média	Boa	Não	Lucro
4	Sem	Média	Não	Prejuízo
5	Média	Média	Sim	Prejuízo
6	Alta	Média	Sim	Lucro

Fonte: Pawlak (2000).

Com base nos atributos E, Q e L, pode-se afirmar o seguinte: as lojas 1 e 6 realizaram lucro, as lojas 4 e 5 tiveram prejuízo e as lojas 2 e 3 não podem ser classificadas (em lucro ou prejuízo), porque são indiscerníveis em relação a esses atributos. Assim, empregando-se os atributos E, Q e L, deduz-se que: as lojas 1 e 6 certamente realizaram lucro, isto é, certamente pertencem ao conjunto {1,3,6}; enquanto as lojas 1, 2, 3 e 6 possivelmente realizaram lucro, isto é, possivelmente pertencem ao conjunto {1,3,6}. Os conjuntos {1,6} e {1,2,3,6} representam, respectivamente, as aproximações inferior e superior do conjunto {1,3,6}. O conjunto {2,3} representa a diferença entre as aproximações superior e inferior e caracteriza a “região de fronteira” do conjunto {1,3,6}. Possíveis regras extraídas da tabela de decisão (Quadro 2) (PAWLAK, 2000):

$$(E, \text{média}) \text{ and } (Q, \text{boa}) \rightarrow (P, \text{prejuízo})$$

$$(E, \text{sem}) \rightarrow (P, \text{prejuízo}) \quad (\text{ou se } E = \text{“sem” então “prejuízo”})$$

$$(E, \text{média}) \text{ and } (Q, \text{média}) \rightarrow (P, \text{prejuízo})$$

Conforme Pawlak (2002) e Greco *et al.* (2004), uma regra de decisão em S, sistema de informação, é uma expressão  $\Phi \rightarrow \Psi$ , leia-se “se  $\Phi$  então  $\Psi$ ”, onde  $\Phi \in \text{For}(C)$ ,  $\Psi \in \text{For}(D)$ , C e D são atributos de condição e decisão, respectivamente. São definidos três fatores:

a) *Strength* ou suporte, como definido na Equação (5):

$$\sigma_s = \text{supp}_s(\Phi, \Psi) / \text{card}(U) \quad (5),$$

$\text{supp}_s(\Phi, \Psi) = \text{card}(\|\Phi \wedge \Psi\|_s)$ , e representa a razão entre todos os fatos que podem ser classificados por uma determinada regra e todos os fatos na tabela;

b) *Certainty*, definido na Equação (6):

$$\text{cer}_s = \text{card}(\|\Phi \wedge \Psi\|_s) / \text{card}(\|\Phi\|_s) \quad (6),$$

e representa a estimativa de probabilidade condicional que  $\Psi$  seja verdadeiro em S dado que  $\Phi$  é verdadeiro em S; se  $0 < \text{cer}_s < 1$  para uma regra de decisão, diz-se que esta regra é incerta. Esse fator é amplamente utilizado em *Data Mining* e é conhecido como coeficiente de confiança;

c) *Coverage*, definido como na Equação (7):

$$\text{cov}_s = \text{card}(\|\Phi \wedge \Psi\|_s) / \text{card}(\|\Psi\|_s) \quad (7),$$

e representa a estimativa de probabilidade condicional que  $\Phi$  seja verdadeiro em S dado que  $\Psi$  é verdadeiro em S.

*Certainty* e *coverage* expressam quão exato é o conhecimento sobre a realidade considerada; eles não são assumidos como arbitrários, mas são calculados a partir dos dados e são objetivos. Ademais, a RST propicia uma outra forma de relacionar condições e decisões nas regras de decisão, sem se referir a probabilidades a priori ou posterior como na abordagem pelo Teorema de Bayes (PAWLAK, 2002).

De acordo com Slowinski *et al.* (2012), a RST ignora não apenas a ordem de preferência no conjunto de valores dos atributos mas também, a relação monotônica de avaliação de objetos em relação aos valores dos atributos de condição e a ordem de preferência dos valores dos atributos de decisão (classificação ou grau de preferência). Assim, Slowinski *et al.* (2012) apresentam o princípio de Dominância, *Dominance Rough Sets Approach*, DRSA: “objetos que possuem uma melhor avaliação ou que possuam no mínimo a mesma avaliação (classe de decisão), não podem ser associados a uma pior classe de decisão, considerados todos os critérios”. As relações de indiscernibilidade são então substituídas por relações de dominância nas aproximações da classe de decisão. Devido à ordem de preferência entre as classes de decisão, os conjuntos tornam-se aproximações e são conhecidos como uniões de classes de decisão: classes ascendentes (*upward*,  $Cl_{\bar{t}}^{\geq}$ ) e descendentes (*downward*,  $Cl_{\bar{t}}^{\leq}$ ).

Assim, visando manter a dependência monotônica entre critérios de condição e de decisão, as classes ascendentes e descendentes são definidas conforme Equações (8) e (9), respectivamente (SLOWINSKI *et al.*, 2012):

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s, t = 1, \dots, n \text{ (“at least” classe } Cl_t) \quad (8)$$

$$Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s, t = 1, \dots, n \text{ (“at most” classe } Cl_t) \quad (9)$$

Segundo Slowinski *et al.* (2012), diz-se que,  $x$  domina  $y$  com relação a  $P \subseteq C$ , denotado por  $x D_P y$ , se  $x S_q y$ , isto é,  $x$  é no mínimo tão bom quanto  $y$  com relação ao critério  $q$  para todo  $q \in P$ . A relação de dominância é reflexiva e transitiva. Dado que,  $P \subseteq C$  e  $x \in U$ , os “grânulos de conhecimento” usados nas aproximações DRSA são: um conjunto de objetos dominantes sobre  $x$ , denominado conjunto P-dominante:  $D_p^+(x) = \{y \in U : y D_p x\}$ ; e um conjunto de objetos dominados por  $x$ , denominado conjunto P-dominado:  $D_p^-(x) = \{y \in U : x D_p y\}$ . Usando os conjuntos  $D_p^+(x)$ , as aproximações P-inferior e P-superior de  $Cl_t^{\geq}$  são respectivamente:  $\underline{P}(Cl_t^{\geq}) = \{x \in U : D_p^+(x) \subseteq Cl_t^{\geq}\}$ ,  $\overline{P}(Cl_t^{\geq}) = \bigcup_{x \in Cl_t^{\geq}} D_p^+(x)$ , para  $t = 1, \dots, n$ . Analogamente, as aproximações P-inferior e P-superior de  $Cl_t^{\leq}$  são respectivamente:  $\underline{P}(Cl_t^{\leq}) = \{x \in U : D_p^-(x) \subseteq Cl_t^{\leq}\}$ ,  $\overline{P}(Cl_t^{\leq}) = \bigcup_{x \in Cl_t^{\leq}} D_p^-(x)$ , para  $t = 1, \dots, n$ . Os conjuntos P-fronteira (*P-boundaries*) de  $Cl_t^{\geq}$  e  $Cl_t^{\leq}$  são respectivamente,  $Bn_p(Cl_t^{\geq}) = \overline{P}(Cl_t^{\geq}) - \underline{P}(Cl_t^{\geq})$  e  $Bn_p(Cl_t^{\leq}) = \overline{P}(Cl_t^{\leq}) - \underline{P}(Cl_t^{\leq})$ , para  $t = 1, \dots, n$ .

Conforme Slowinski *et al.* (2012), as aproximações das uniões de classes ascendente e descendente podem servir para induzir regras de decisão “se ... então ...”. Como para dada união de classe ascendente ou descendente,  $Cl_t^{\geq}$  ou  $Cl_t^{\leq}$ ,  $s, t \in T$ , as regras induzidas sob a hipótese de que os objetos pertençam a aproximações inferiores  $\underline{P}(Cl_t^{\geq})$  ou a  $\underline{P}(Cl_t^{\leq})$  são positivas, e todas as outras são negativas, sugerem a atribuição de um objeto para “ao menos uma classe  $Cl_t$ ” ou para “ao máximo uma classe  $Cl_s$ ”, respectivamente. Essas regras são conhecidas como “regras de decisão certas” ( $D_{\leq}$  ou  $D_{\geq}$ ) porque atribuem os objetos para uniões de classes de decisão sem qualquer ambiguidade. Se os objetos pertencem a aproximações superiores, as regras são conhecidas como “regras de decisão possíveis”. Assim, os objetos poderiam pertencer para “ao menos uma classe  $Cl_t$ ” ou para “ao máximo uma classe  $Cl_s$ ”. E se os objetos pertencem à interseção  $\overline{P}(Cl_s^{\leq}) \cap \overline{P}(Cl_t^{\geq})$  ( $s < t$ ), as regras induzidas são conhecidas como “regras aproximadas”, isto é, os objetos estão entre as classes  $Cl_s$  e  $Cl_t$ . Assim, se para cada critério  $q \in C$ ,  $V_q$

$\subseteq \mathbb{R}$  ( $V_q$  é quantitativo) e que para cada  $x, y \in U$ ,  $f(x, q) \geq f(y, q)$  implica  $x S_q y$  ( $V_q$  possui uma ordem de preferência), as regras de decisão podem ser consideradas sob 5 tipos:

1- regras de decisão certas- $D_{\geq}$ :

se  $f(x, q_1) \geq r_{q_1}$  e  $f(x, q_2) \geq r_{q_2}$  e ...  $f(x, q_p) \geq r_{q_p}$ , então  $x \in Cl_t^{\geq}$

2- regras de decisão possíveis- $D_{\geq}$ :

se  $f(x, q_1) \geq r_{q_1}$  e  $f(x, q_2) \geq r_{q_2}$  e ...  $f(x, q_p) \geq r_{q_p}$ , então  $x$  possivelmente pertence a  $Cl_t^{\geq}$

3- regras de decisão certas- $D_{\leq}$ :

se  $f(x, q_1) \leq r_{q_1}$  e  $f(x, q_2) \leq r_{q_2}$  e ...  $f(x, q_p) \leq r_{q_p}$ , então  $x \in Cl_t^{\leq}$

4- regras de decisão possíveis- $D_{\leq}$ :

se  $f(x, q_1) \leq r_{q_1}$  e  $f(x, q_2) \leq r_{q_2}$  e ...  $f(x, q_p) \leq r_{q_p}$ , então  $x$  possivelmente pertence a  $Cl_t^{\leq}$

$P = \{q_1, \dots, q_p\} \subseteq C$ ,  $(r_{q_1}, \dots, r_{q_p}) \in V_{q_1} \times V_{q_2} \times \dots \times V_{q_p}$  e  $t \in T$ ;

5- regras de decisão aproximadas- $D_{\leq \geq}$ :

se  $f(x, q_1) \geq r_{q_1}$  e  $f(x, q_2) \geq r_{q_2}$  e ...  $f(x, q_k) \geq r_{q_k}$  e  $f(x, q_{k+1}) \leq r_{q_{k+1}}$  e  $f(x, q_p) \leq r_{q_p}$ , então  $x \in Cl_s \cup Cl_{s+1} \cup \dots \cup Cl_t$ .

As regras dos tipos 1 e 3 representam determinado conhecimento extraído dos dados, por exemplo, classificação ordinal; as regras dos tipos 2 e 4 representam conhecimentos possíveis; e regras do tipo 5 representam conhecimento duvidoso, isto é, são suportadas apenas por objetos inconsistentes.

Slowinski *et al.* (2012) definem a qualidade da aproximação ou grau de consistência de uma classificação ordinal  $Cl$  por um conjunto de atributos  $P$ , como a razão do número de objetos  $P$ -consistentes com o princípio de Dominância e o número de objetos em  $U$ , Equação (10):

$$\gamma_P(Cl) = \frac{|U - (\cup_{t=2, \dots, m} Bn_P(Cl_t^{\geq}))|}{|U|} = \frac{|U - (\cup_{t=1, \dots, m-1} Bn_P(Cl_t^{\leq}))|}{|U|} \quad (10)$$

Blaszczynski *et al.* (2013a) definem a acurácia ou precisão de aproximação da união de classes ordenadas,  $Cl_t^{\geq}$  e  $Cl_t^{\leq}$ , de um conjunto de atributos  $P$ , como a razão de objetos das aproximações  $P$ -inferior e  $P$ -superior, como nas Equações (11) e (12):

$$\alpha_P(Cl_t^{\geq}) = \frac{P(Cl_t^{\geq})}{\overline{P}(Cl_t^{\geq})} \quad (11)$$

$$\alpha_P(Cl_t^{\leq}) = \frac{P(Cl_t^{\leq})}{\overline{P}(Cl_t^{\leq})} \quad (12)$$

E, há estudos que tratam sobre modelos híbridos *Rough-Fuzzy*, *Fuzzy-Rough* e *Rough-Neural*, tais como: Greco *et al.* (2006, 2008); Jensen *et al.* (2008, 2009, 2010); Zhao *et al.* (2010); Jensen *et al.* (2011); Riza *et al.* (2014); Du; Hu (2017); Nowicki (2019); Riza *et al.* (2019).

### 2.3. MACHINE LEARNING

*Machine Learning* ou Aprendizado de Máquina (AM) refere-se ao subcampo da Inteligência Artificial (IA), estudo das faculdades mentais por meio de modelos computacionais, que visa projetar algoritmos, sequência finita de operações, permitindo que programas de computadores aprendam com dados progressos (RUSSELL; NORVIG, 2010). E, como instrumento, por exemplo, encontra-se a indução de regras: regras condição-ação, árvores de decisão ou estruturas similares de conhecimento (LANGLEY; SIMON, 1995; PAWLAK *et al.*, 1995).

Assim, programas de computador aprendem automaticamente, a reconhecer padrões complexos e tomar decisões inteligentes com base em dados. Por exemplo, um problema típico de AM é programar um computador para que ele possa automaticamente reconhecer códigos postais escritos à mão no correio depois de aprender com um conjunto de exemplos (HAN *et al.*, 2012).

Desde os anos 1970, houve uma expansão do uso da IA para a solução de problemas reais. No início, os problemas eram tratados pela IA por meio da aquisição de conhecimento de especialistas de um dado domínio, por exemplo, medicina, que era então codificado, frequentemente por regras lógicas, em um programa de computador. Esses programas eram conhecidos como Sistemas Especialistas ou Sistemas Baseados em Conhecimento. A aquisição do conhecimento de especialistas ocorria por meio de entrevistas que buscavam descobrir que regras eles utilizavam para tomar decisões. Esse processo possui várias limitações, como subjetividade e, muitas vezes, pouca cooperação por parte do especialista, por causa do receio de ser dispensado após repassar seu conhecimento (FACELI *et al.*, 2021).

Dada a crescente complexidade dos problemas a serem computacionalmente tratados, da velocidade e volume de dados gerados por diferentes setores, ferramentas computacionais mais sofisticadas e autônomas foram desenvolvidas, mais independentes da intervenção humana, para a aquisição de conhecimento. A maioria dessas ferramentas tem como base o AM, uma subárea da IA que faz parte de várias das tecnologias atualmente utilizadas (FACELI *et al.*, 2021).

Faceli *et al.* (2021) relacionam aplicações bem-sucedidas de técnicas de AM a problemas reais: interfaces que utilizam linguagem natural, escrita ou falada, reconhecimento facial, filtragem de *spam* em e-mails, detecção de fraude por bancos e operadoras de cartões de crédito, auxílio ao diagnóstico de doenças por meio da análise

de dados clínicos, de imagem ou dados genéticos, recomendação de produtos com base no perfil do consumidor e em seu histórico de consumo, comportamento inteligente em personagens de jogos, entre outros.

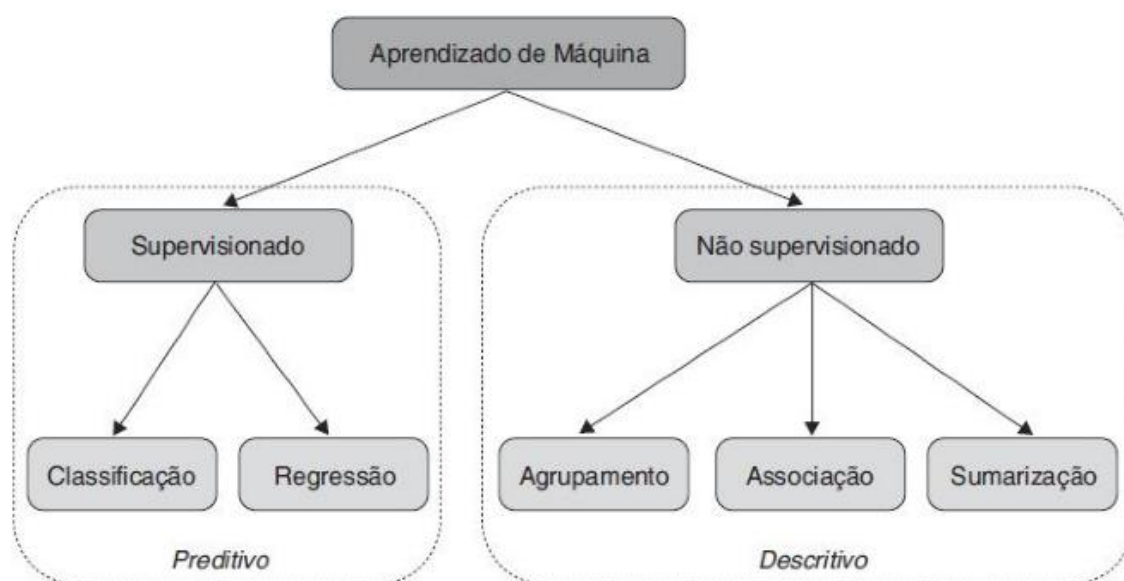
Aplicações baseadas em AM utilizam heurísticas que buscam por modelos capazes de representar o conhecimento presente em um conjunto de dados. Em geral, os conjuntos de dados são estruturados em formato tabular, uma matriz atributo-valor, em que cada linha representa um objeto, instância ou exemplo, e cada coluna representa um atributo, característica ou variável. Os atributos podem ser divididos em atributos preditivos, cujos valores descrevem características dos objetos, que formam um vetor de entrada, e atributo alvo, cujo valor rotula o objeto, com uma classe ou valor numérico. Essas denominações têm por origem o frequente uso dos valores dos atributos preditivos de um objeto para prever o valor de seu atributo alvo. Nem todos os conjuntos de dados possuem atributo alvo. Quando possuem, são chamados de conjuntos de dados rotulados (FACELI *et al.*, 2021).

Algoritmos de AM têm sido amplamente utilizados em diversas tarefas, que podem ser divididas em Preditivas e Descritivas. Em tarefas preditivas, algoritmos de AM são aplicados a conjuntos de dados de treinamento rotulados para induzir um modelo preditivo capaz de prever para um novo objeto representado pelos valores de seus atributos preditivos, o valor de seu atributo alvo. Modelos preditivos podem ser utilizados, por exemplo, para a partir de seus sintomas, prever o estado de saúde de um paciente. Nessas tarefas, em geral são utilizados algoritmos de AM que seguem o paradigma de aprendizado supervisionado. O termo supervisionado vem da simulação da presença de um supervisor externo, que conhece, por exemplo, o verdadeiro diagnóstico do novo paciente. Essa informação é usada para guiar o processo de aprendizado na extração de um modelo com boa capacidade preditiva. Em tarefas de descrição, ao invés de prever um valor, algoritmos de AM extraem padrões dos valores preditivos de um conjunto de dados. Como não fazem uso do conhecimento do supervisor externo, esses algoritmos usam o paradigma de aprendizado não supervisionado. Uma das principais tarefas descritivas, agrupamento de dados, procura grupos de objetos similares entre si no conjunto de dados. Outra tarefa descritiva é encontrar regras de associação, que associam valores de um subconjunto de atributos preditivos a valores de outro subconjunto (FACELI *et al.*, 2021; HAN *et al.*, 2012).

A Figura 4 ilustra hierarquicamente as categorias de AM e as tarefas associadas. No topo aparece AM indutivo, que aprende a realizar generalizações a partir de um

conjunto de dados. Tem-se em seguida, as categorias de algoritmos de aprendizado supervisionado, tarefas preditivas, e não supervisionado, tarefas descritivas. As tarefas preditivas se distinguem pelo valor de rótulo a ser predito: discreto, no caso de tarefas de classificação; e contínuo, no caso de tarefas de regressão. As tarefas descritivas são genericamente divididas em: agrupamento, que dividem os dados em grupos de acordo com sua similaridade; sumarização, que buscam uma descrição simples e compacta para um conjunto de dados; e associação, que procuram padrões frequentes de associações entre os atributos de um conjunto de dados (FACELI *et al.*, 2021).

Figura 4– Hierarquia do AM



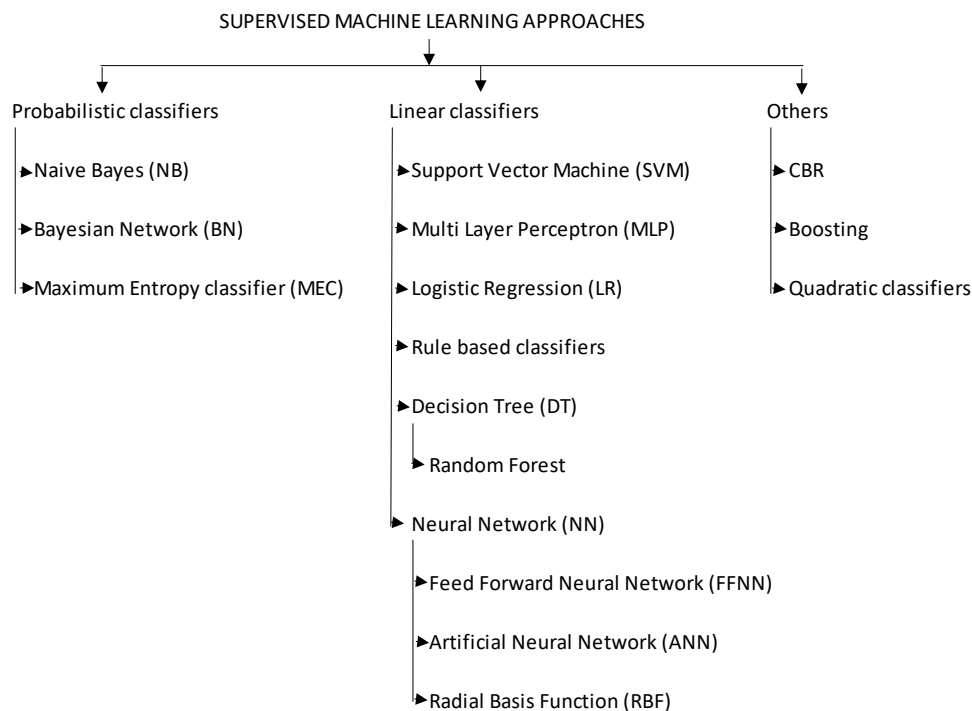
Fonte: Faceli *et al.* (2021).

Segundo Faceli *et al.* (2021), em AM, dispositivos computacionais são programados para aprender a partir de experiências passadas e, frequentemente, empregam um princípio de inferência denominado indução, que permite extrair conclusões genéricas a partir de um conjunto particular de exemplos. Deve ser observado que, apesar dessa divisão de modelos em preditivos e descritivos, um modelo preditivo pode gerar uma descrição de um conjunto de dados, e um modelo descritivo pode prover previsões após ser validado. E, um requisito importante para algoritmos de AM é que sejam capazes de lidar com dados imperfeitos. Muitos conjuntos de dados apresentam algum tipo de problema, como presença de ruídos, dados inconsistentes, dados ausentes e dados redundantes. Algoritmos de AM devem, idealmente, ser robustos a esses problemas, minimizando sua influência no processo indutivo. Entretanto, dependendo de

sua extensão, esses problemas podem prejudicar ou inviabilizar o aprendizado. Técnicas de pré-processamento permitem identificar e reduzir, ou até mesmo eliminar, esses problemas.

E, Saravanan; Sujatha (2018) apontam as abordagens em AM, na categoria supervisionado, tarefa de classificação, conforme Figura 5.

Figura 5– Abordagens em AM, categoria supervisionado, tarefa de classificação



Fonte: Saravanan; Sujatha (2018).

Os classificadores com base em regras de decisão modelam o espaço de dados para uma coleção de regras. O lado esquerdo indica o estado do conjunto de recursos representado na forma normal disjuntiva e o lado direito indica o rótulo da classe. Várias normas ou formas estão disponíveis para se criar regras; a fase de treinamento constrói as regras com base nestas normas. A medida de suporte indica o número real de instâncias no conjunto de dados de treinamento que são apropriadas à regra. A medida de confiança denota a possibilidade condicional de que o lado direito da regra é cumprido quando o lado esquerdo também é cumprido (SARAVANAN; SUJATHA, 2018).

E, de acordo com Faceli *et al.* (2021), há vários algoritmos para aprender regras de decisão. Por exemplo, o algoritmo da Cobertura. Esse algoritmo define o processo de aprendizado como um processo de procura: dados um conjunto de exemplos classificados e uma linguagem para representar generalizações dos exemplos, o algoritmo procede para



cada classe, a uma procura heurística. O algoritmo procura regras da seguinte forma: se  $\text{Atributo}_i = \text{valor}_j$  e  $\text{Atributo}_1 = \text{valor}_k \dots$  então  $\text{Classe}_z$ . A procura pode proceder a partir da regra mais geral, ou seja, uma regra sem parte condicional, para regras mais específicas, acrescentando condições; ou a partir de regras muito específicas para regras mais gerais, eliminando restrições. O processo de procura é guiado por uma função de avaliação das hipóteses. Essa função estima a qualidade das regras que são geradas durante o processo.

Especificamente, no contexto deste estudo, foi utilizada a abordagem Supervisionada de Aprendizado de Máquina, classificadores lineares, classificadores com base em regras de decisão. Esta decisão fundamenta-se principalmente, no aparato teórico que a RST apresenta sobre sistema de informação e regras de decisão que podem ser inferidas sobre uma tabela de decisão, além da possibilidade de tratamento matemático da imprecisão e incerteza de dados (GRECO *et al.*, 2004; NOWICKI, 2019; PAWLAK, 1991, 2002; SLOWINSKI *et al.*, 2012).

### 3. METODOLOGIA DE PESQUISA

A Metodologia de Pesquisa visa estudar, compreender e avaliar os vários métodos disponíveis para a realização de uma pesquisa acadêmica. A Metodologia, em um nível aplicado, examina, descreve e avalia métodos e técnicas de pesquisa que possibilitam a coleta e o processamento de informações, visando ao encaminhamento e à resolução de problemas e/ou questões de investigação. Assim, é a aplicação de procedimentos e técnicas que devem ser observados para a construção do conhecimento, com o propósito de comprovar sua validade e utilidade nos diversos âmbitos da sociedade. A presente proposta de pesquisa utilizou os critérios de classificação propostos por Prodanov; Freitas (2013).

#### 3.1. CLASSIFICAÇÃO DA PESQUISA

- natureza: aplicada, isto é, objetiva gerar conhecimentos para aplicação prática dirigidos à solução de problemas específicos;
- método científico: dedutivo (parte do geral e, a seguir, desce ao particular; a partir de princípios, leis ou teorias consideradas verdadeiras e indiscutíveis, prediz a ocorrência de casos particulares com base na lógica);
- objetivo: descritiva (fatos observados, registrados, analisados, classificados e interpretados);
- procedimento técnico: pesquisa documental/de campo;
- abordagem: quantitativa.

Esta classificação justifica-se por: i) natureza aplicada, proposição e aplicação de modelo para avaliação integrada da sustentabilidade e desempenho das empresas concessionárias de distribuição de energia elétrica; ii) método dedutivo, universo de empresas concessionárias de distribuição de energia elétrica; iii) objetivo descritiva, com base em fatos registrados, analisados, classificados e interpretados; iv) procedimento técnico de pesquisa documental; v) abordagem quantitativa com base em indicadores de sustentabilidade e desempenho das empresas concessionárias.

#### 3.2. INSTRUMENTOS DE COLETA DE DADOS

Foi utilizado o universo de empresas concessionárias de distribuição de energia elétrica no Brasil, por meio de acesso a várias fontes de dados (pesquisa documental;

automatizada/semi-automatizada) com triangulação (comparação entre dados oriundos de diferentes fontes).

Para o contexto desta pesquisa, a fonte de dados principal é o Relatório de Responsabilidade Socioambiental e Econômico-Financeiro (RSA) em formato PDF, que contém indicadores quantitativos e qualitativos relativos às dimensões geral, governança corporativa, econômico-financeira, social, setorial e ambiental, produzido por cada empresa concessionária de distribuição de energia elétrica, e disponibilizado anualmente pela Agência Nacional de Energia Elétrica (ANEEL) em seu *site* na internet (ANEEL, 2014, 2015, 2021, 2022a).

A coleta foi realizada pela captura de informações em formato “txt”, por técnicas e algoritmos de Recuperação de Informação (BAEZA-YATES; RIBEIRO-NETO, 2013), implementados em VBA (*Visual Basic for Applications*)/Microsoft Excel.

### 3.3. ANÁLISE E TRATAMENTO DE DADOS

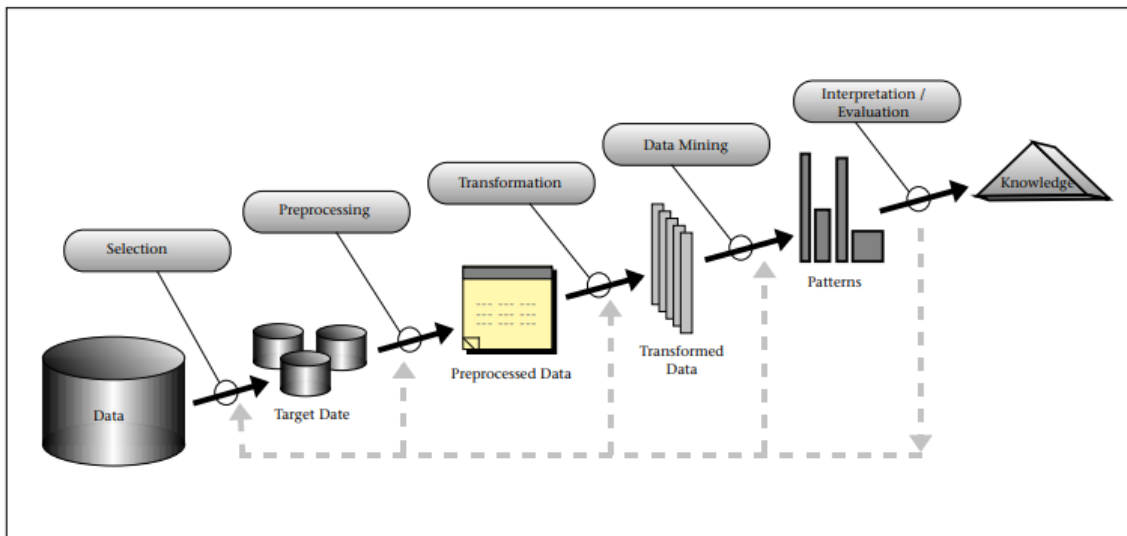
Foi utilizada a análise por “encadeamento lógico de evidências”: como se desenvolve a pesquisa, desde as questões iniciais até as conclusões.

E, de forma complementar, os dados extraídos foram então tratados visando identificar dados ausentes e inconsistentes: uso de técnicas de mineração de textos (HIGUCHI, 2001; HAN *et al.*, 2012; PROVOST; FAWCETT, 2013; ZHAI; MASSUNG, 2016; SILGE; ROBINSON, 2017; ZIZKA *et al.*, 2020); VBA/Microsoft Excel; e, Estatística.

Foram utilizados os seguintes *softwares* para análise de dados e inferência de regras de decisão: *package RoughSets*, v. 1.3-7, de 15/12/2019, em R (RIZA *et al.*, 2014, 2019); jMAF, *Dominance-Based Rough Set Data Analysis Framework*, v. 08-03-2019 (BLASZCZYNSKI *et al.*, 2013b), para análise sob a ótica da Teoria dos Conjuntos Aproximativos (*Rough Sets Theory*)/princípio de Dominância (*Dominance-based Rough Set Approach*, DRSA) (PAWLAK, 1991; COUTO; GOMES, 2010; SLOWINSKI *et al.*, 2012; COUTO; GOMES, 2020; KAROLAK *et al.*, 2020); e, Weka, *Waikato Environment for Knowledge Analysis*, v. 3.8.4, para mineração de dados (WITTEN; FRANK, 2005).

Foi utilizado o processo KDD, *Knowledge Discovery in Databases* (FAYYAD *et al.*, 1996a, 1996b) para seleção, pré-processamento, transformação, obtenção de padrões e descoberta de conhecimento sobre os indicadores das empresas concessionárias de distribuição de energia elétrica, conforme Figura 6.

Figura 6 – Processo KDD, da seleção de dados até a obtenção de conhecimento



Fonte: Fayyad *et al.* (1996a, 1996b).

E, o Quadro 3 resume as etapas da Metodologia de Pesquisa.

Quadro 3– Resumo da Metodologia de Pesquisa

Etapa	Objetivo(s)	Método(s)/Ferramenta(s)
1	Revisão sistemática da literatura: levantamento dos estudos nas bases de dados da CAPES, uso de 3 Eixos de pesquisa; verificação dos estudos aderentes à questão de pesquisa e ao programa; análise dos portfólios bibliográficos; identificação dos <i>gaps</i> nos estudos reunidos; demonstração do ineditismo do tema; elaboração da bibliometria	Método construtivista ProKnow-C; StArt; Microsoft Excel; VOSviewer
2	Coleta dos Relatórios de Responsabilidade Socioambiental e Econômico-Financeiro (RSA) das empresas, período 2013-2021; extração dos indicadores de sustentabilidade e desempenho	Processo KDD; VBA/Microsoft Excel; técnicas e algoritmos de Recuperação da Informação; mineração de textos
3	Tratamento de dados: identificação de valores mínimo e máximo, variações ao longo período 2013-2021, mediana, <i>outliers</i> , por indicador; verificação/conferência; padronização de unidades; identificação de inconsistências; preenchimento de valores ausentes com mediana, por empresa e anual; tratamento matemático de imprecisões e incertezas dos dados	Processo KDD; VBA/Microsoft Excel; Estatística; RST/DRSA; jMAF
4	Classificação das empresas; identificação de padrões de dados; inferência de regras de decisão	Processo KDD; RST/DRSA; VBA/Microsoft Excel; jMAF; <i>package RoughSets</i> em R; Weka
5	Elaboração de <i>benchmarking</i> ; análise de sensibilidade e risco	VBA/Microsoft Excel; RStudio
6	Predição da classificação de empresas; análise de risco	Algoritmos de classificação; método <i>ensemble bagging</i>
7	Validação: empresas de destaque no <i>benchmarking</i> e <i>outliers versus</i> boas práticas de sustentabilidade e desempenho/fatos relevantes	Coleta/verificação de boas práticas de sustentabilidade e desempenho/fatos relevantes descritos nos RSAs

Fonte: Autor.

### 3.4. REVISÃO SISTEMÁTICA DA LITERATURA

Para a revisão sistemática da literatura, isto é, a construção do conhecimento sobre o tema de pesquisa presente, foi utilizado o método ProKnow-C (*Knowledge Development Process – Constructivist*) (AFONSO *et al.*, 2011; ENSSLIN *et al.*, 2014); o *software* de apoio à construção de portfólio bibliográfico, StArt (*State of the Art through Systematic Review*), versão 2.3.4.2, do LaPES (Laboratório de Pesquisa em Engenharia de *Software*), Departamento de Computação, Universidade Federal de São Carlos (UFSCar) (LAPES, 2021); e o *software* de apoio à análise bibliométrica, VOSviewer, versão 1.6.17 (VAN ECK; WALTMAN, 2021).

O Quadro 4, constituído de oito fases e sub fases, reproduz os procedimentos utilizados pelo método ProKnow-C, visando construir um portfólio bibliográfico final (PB5) por meio de portfólios bibliográficos intermediários (PB, 0 a 4).

Quadro 4– Revisão sistemática da literatura

FASE 1	FASE 2	FASE 3	FASE 4	FASE 5	FASE 6	FASE 7	FASE 8
Definir palavras-chave para a pesquisa	Definir bases CAPES alinhadas à pesquisa	Realizar pesquisa em título, resumo e palavras-chave das publicações em cada base	Contabilizar nº publicações por Eixo e base visando montar o portfólio bibliográfico inicial (PB0)	Levantar nº citações do PB1 no Google Acadêmico	Selecionar publicações mais citadas (até 85% do total de citações)	Selecionar publicações menos citadas (15% restantes do total de citações)	Reunir publicações PB2, PB3 e PB4
Definir Eixos de pesquisa; montar expressões booleanas (“and”; “or”)	Definir filtros de pesquisa (período de tempo, assunto, tipo publicação, etc.)	Rever se palavras-chave pesquisadas estão aderentes às palavras-chave das publicações	Selecionar publicações aderentes à pesquisa (leitura do título); excluir publicações repetidas (PB1)	Classificar publicações em ordem decrescente de citações	Selecionar publicações aderentes à pesquisa (leitura do resumo) (PB2)	Selecionar publicações nos 2 últimos anos aderentes à pesquisa (leitura do resumo) (PB3);  Selecionar publicações restantes que possuem autores no PB2 e aderentes à pesquisa (leitura do resumo) (PB4)	Selecionar publicações aderentes à pesquisa (leitura integral) para montar o portfólio bibliográfico final (PB5)

Fonte: Autor; adaptado de Afonso *et al.* (2011); Ensslin *et al.* (2014).

Inicialmente, de acordo com a Fase 1, foram definidas as palavras-chave e os Eixos de pesquisa 1, 2 e 3 alinhados à questão de pesquisa, conforme Quadro 5.

Quadro 5– Definição das palavras-chave e Eixos de pesquisa para a revisão sistemática da literatura

Palavras-chave	Eixos		
	1	2	3
"multi-criteria" OR "multicriteria" OR "multi-objective" OR "multiobjective"		●	
"machine learning"		●	
"sustainability" OR "sustainable"	●	●	●
"performance"	●	●	●
"benchmarking" OR "benchmark"	●	●	
("energy companies" OR "electricity companies" OR "energy firms" OR "electricity firms" OR "energy industry" OR "power industry" OR "energy sector" OR "electricity sector" OR "electric sector" OR "energy enterprises")	●	●	●
("indicator" OR "indice" OR "index" OR "measurement" OR "assessment" OR "evaluation" OR "appraisal" OR "metric" OR "model" OR "framework" OR "template" OR "example")			●

Fonte: Autor.

A seguir, foram montadas as *strings* de pesquisa de acordo com os Eixos de pesquisa, Quadro 6.

Quadro 6– *Strings* de pesquisa para a revisão sistemática da literatura

Eixo	String de pesquisa	Tradução
1	("benchmarking" OR "benchmark") AND ("sustainability" OR "sustainable") AND "performance" AND ("energy companies" OR "electricity companies" OR "energy firms" OR "electricity firms" OR "energy industry" OR "power industry" OR "energy sector" OR "electricity sector" OR "electric sector" OR "energy enterprises")	<i>benchmarking</i> em sustentabilidade e desempenho para empresas de energia ou eletricidade
2	("benchmarking" OR "benchmark") AND ("sustainability" OR "sustainable") AND "performance" AND ("energy companies" OR "electricity companies" OR "energy firms" OR "electricity firms" OR "energy industry" OR "power industry" OR "energy sector" OR "electricity sector" OR "electric sector" OR "energy enterprises") AND (("multi-criteria" OR "multicriteria" OR "multi-objective" OR "multiobjective") OR "machine learning")	<i>benchmarking</i> em sustentabilidade e desempenho para empresas de energia ou eletricidade com apoio multicritério ou <i>machine learning</i>
3	("sustainability" OR "sustainable") AND "performance" AND ("indicator" OR "indice" OR "index" OR "measurement" OR "assessment" OR "evaluation" OR "appraisal" OR "metric" OR "model" OR "framework" OR "template" OR "example") AND ("energy companies" OR "electricity companies" OR "energy firms" OR "electricity firms" OR "energy industry" OR "power industry" OR "energy sector" OR "electricity sector" OR "electric sector" OR "energy enterprises")	avaliação ou modelo sobre sustentabilidade e desempenho para empresas de energia ou eletricidade

Fonte: Autor.

Conforme Fase 2, foram definidas onze bases multidisciplinares da CAPES para a realização da pesquisa da literatura, no período 2000-2021, com posição de janeiro/2022 (Tabela 1).

Tabela 1 – Bases CAPES de pesquisa, período 2000-2021 (posição janeiro/2022), para a revisão sistemática da literatura

Base	Nº artigos por Eixo		
	1	2	3
Scopus	12	2	301
Web of Science – Coleção Principal	7	2	154
Compendex	4	1	135
IEEE Xplore	0	0	0
Emerald Insight	0	0	2
Scielo Citation Index	0	0	1
Academic Search Premier (EBSCO)	5	1	68
ACM Digital Library	0	0	1
Biblioteca Digital de Teses e Dissertações: BDTD	1	0	21
Catálogo de Teses e Dissertações (CAPES)	7	0	64
Wiley Online Library	2	1	19
<b>Total</b>	<b>38</b>	<b>7</b>	<b>766</b>

Fonte: Autor.

A Tabela 2 resume o número de publicações com base nos procedimentos descritos nas Fases 4 a 8 no Quadro 4, em cada portfólio bibliográfico (PB; 0 a 5) e para cada Eixo de pesquisa.

Tabela 2 – Número de publicações de cada portfólio bibliográfico e para cada Eixo de pesquisa

Portfólio bibliográfico (PB)	Fase	Eixos		
		1	2	3
0	4: pesquisa (Eixo) nas bases	38	7	766
1	4: aderentes (leitura título); exclusão repetidos	3	2	132
2	6: mais citados (85%) de PB1; aderentes à pesquisa (leitura resumo)	2	-	23
3	7: menos citados (15%) de PB1, 2 últimos anos; aderentes à pesquisa (leitura resumo)	-	-	29
4	7: menos citados (15%) de PB1, autores em PB2; aderentes à pesquisa (leitura resumo)	-	-	1
5	8: PB2+PB3+PB4; aderentes à pesquisa (leitura integral)	2	2	53

Fonte: Autor.



A seguir, são apresentados os resultados da revisão sistemática da literatura para cada Eixo de pesquisa.

### 3.4.1. REVISÃO DA LITERATURA – EIXO 1

O Quadro 7 mostra o portfólio bibliográfico final (PB5).

Quadro 7– Artigos que formam o portfólio bibliográfico final (PB5), Eixo 1

1. (GOULART, 2013) Goulart, D. D. (2013), Avaliação de índices de eficiência e de produtividade de distribuidoras de energia elétrica no Brasil aplicando análise envoltória de dados (DEA), Biblioteca Digital de Teses e Dissertações - BDTD.
2. (SARTORI <i>et al.</i> , 2017) Sartori, S.; Witjes, S.; Campos, L.M.S. (2017), <i>Sustainability performance for Brazilian electricity power industry: An assessment integrating social, economic and environmental issues</i> , <i>Energy Policy</i> , 111, 41-51.

Fonte: Autor.

A seguir, apresenta-se a análise bibliométrica para o portfólio bibliográfico final (PB5): número de citações dos artigos, relevância dos periódicos e autores nas referências dos artigos – Quadro 8, Figuras 7 e 8, respectivamente.

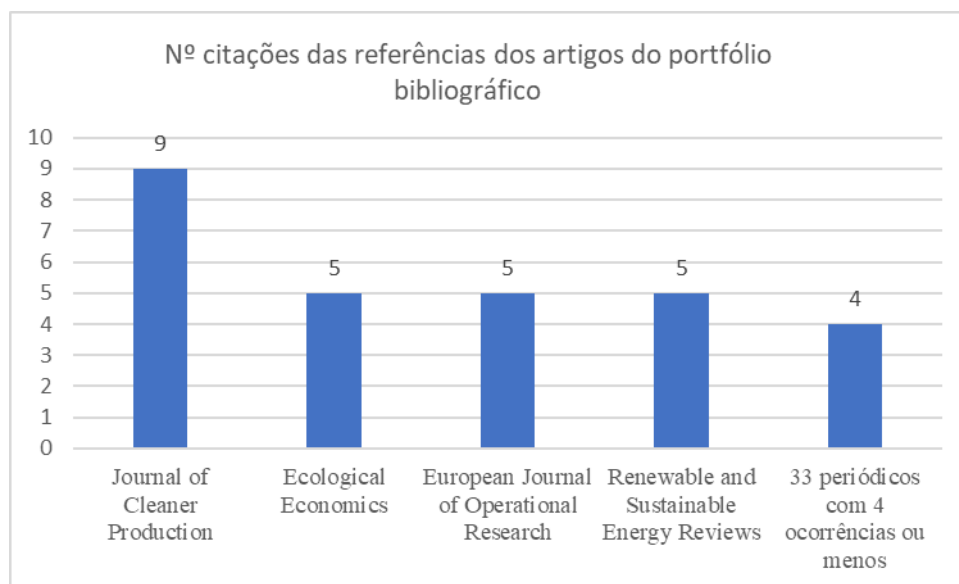
Quadro 8– Número de citações do portfólio bibliográfico final (PB5), Eixo 1

<b>Autor(es)</b>	<b>Artigo</b>	<b>Ano</b>	<b>Nº citações (Google Acadêmico)</b>
Sartori, S.; Witjes, S.; Campos, L.M.S.	<i>Sustainability performance for Brazilian electricity power industry: An assessment integrating social, economic and environmental issues</i>	2017	44
Goulart, D.D.	Avaliação de índices de eficiência e de produtividade de distribuidoras de energia elétrica no Brasil aplicando análise envoltória de dados (DEA)	2013	5

Fonte: Autor.

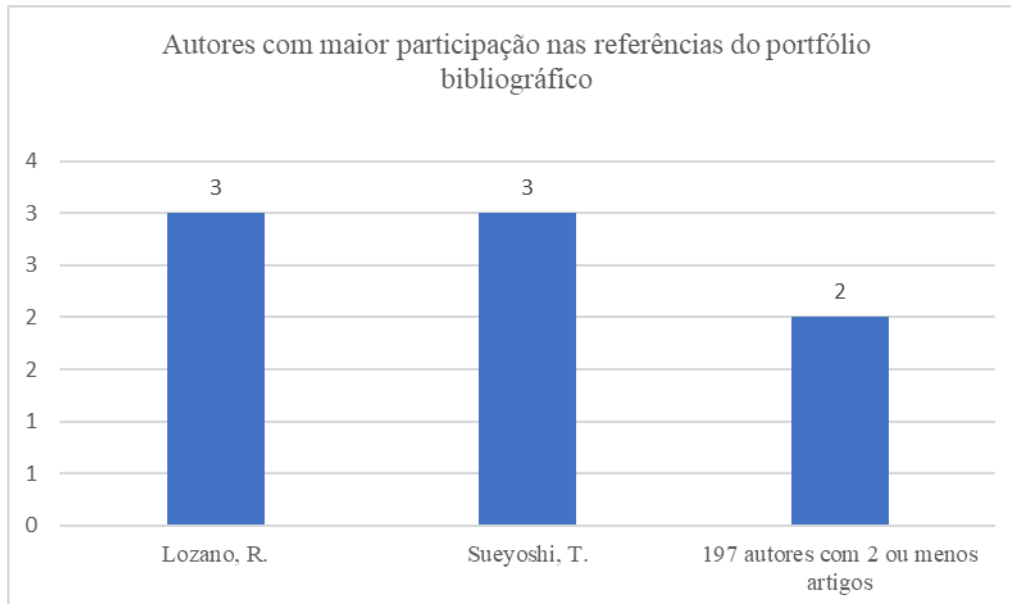
Não foi possível exportar as referências bibliográficas de Goulart (2013), da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD).

Figura 7– Relevância dos periódicos nas referências de Sartori et al. (2017), Eixo 1



Fonte: Autor.

Figura 8 – Autores com maior participação nas referências Sartori et al. (2017), Eixo 1



Fonte: Autor.

E, para complementar a análise bibliométrica, a frequência das palavras-chave pesquisadas (Quadro 9).

Quadro 9– Palavras-chave pesquisadas no portfólio bibliográfico final (PB5), Eixo 1

	Palavras-chave do algoritmo de pesquisa														
	benchmark	benchmarking	electric sector	electricity companies	electricity firms	electricity sector	energy companies	energy enterprises	energy firms	energy industry	energy sector	performance	power industry	sustainability	sustainable
<b>Total</b>	<b>1</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>8</b>	<b>1</b>	<b>9</b>	<b>0</b>
Título											1	1	1		
Resumo	1	1				1					1	7		7	
Palavras-chave														1	

Fonte: Autor.

Assim, conclui-se quanto ao Eixo 1 de pesquisa:

- a) A pesquisa revelou dois estudos sobre *benchmarking* em sustentabilidade e desempenho para empresas de energia ou eletricidade: Goulart (2013) e Sartori *et al.* (2017);
- b) *Journal of Cleaner Production* foi a publicação mais relevante nas referências das publicações do portfólio bibliográfico final (PB5);
- c) Lozano, R. e Sueyoshi, T. foram os autores mais citados nas referências em PB5.

### 3.4.2. REVISÃO DA LITERATURA – EIXO 2

O Quadro 10 mostra o portfólio bibliográfico final (PB5).

Quadro 10– Artigos que formam o portfólio bibliográfico final (PB5), Eixo 2

1. (ANDREOPOULOU; KOLIOUSKA, 2018) Andreopoulou, Z.; Koliouka, C. (2018), <i>Benchmarking internet promotion of renewable energy enterprises: Is sustainability present?</i> , <i>Sustainability</i> , 10, 4187, doi:10.3390/su10114187.
2. (HENDIANI <i>et al.</i> , 2020) Hendiani, S.; Sharifi, E.; Bagherpour, M.; Ghannadpour, S. F. (2020), <i>A multi-criteria sustainability assessment approach for energy systems using sustainability triple bottom line attributes and linguistic preferences</i> , <i>Environment, Development and Sustainability</i> , 22:7771–7805, <a href="https://doi.org/10.1007/s10668-019-00546-7">https://doi.org/10.1007/s10668-019-00546-7</a> .

Fonte: Autor.

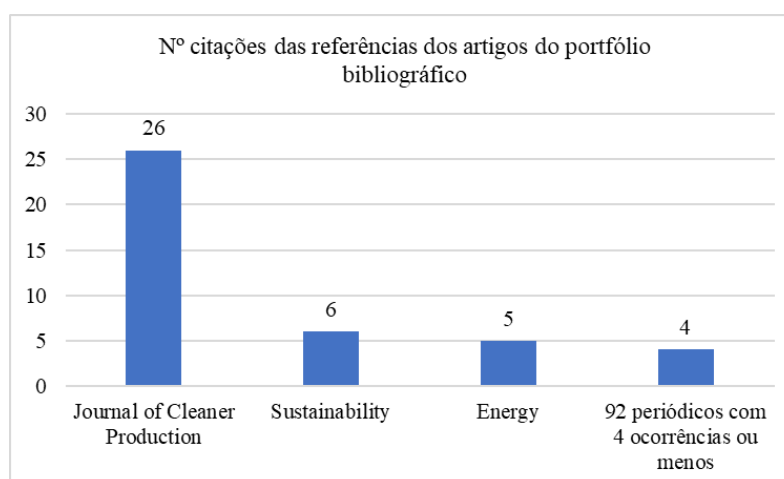
A seguir, apresenta-se a análise bibliométrica sobre o portfólio bibliográfico final (PB5): número de citações dos artigos, relevância dos periódicos e autores nas referências dos artigos – Quadro 11, Figuras 9 e 10, respectivamente.

Quadro 11– Número de citações do portfólio bibliográfico (PB5), Eixo 2

Autor(es)	Artigo	Ano	Nº citações (Google Acadêmico)
Hendiani, S.; Sharifi, E.; Bagherpour, M.; Ghannadpour, S. F.	<i>A multi-criteria sustainability assessment approach for energy systems using sustainability triple bottom line attributes and linguistic preferences</i>	2020	16
Andreopoulou, Z.; Koliouka, C.	<i>Benchmarking internet promotion of renewable energy enterprises: Is sustainability present?</i>	2018	4

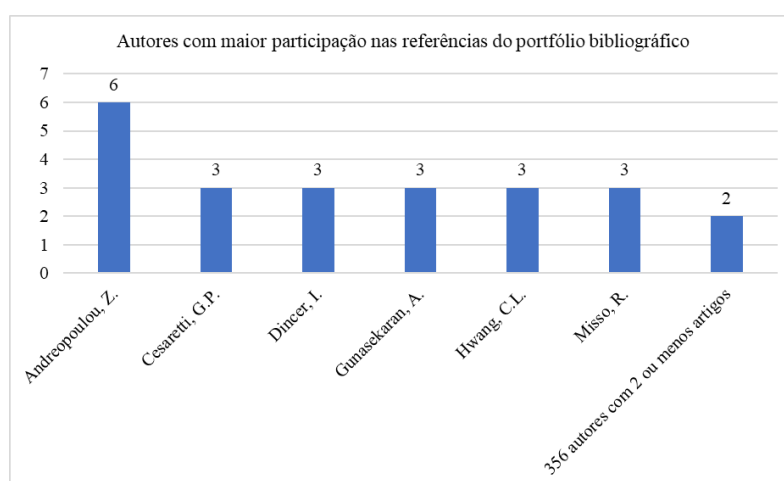
Fonte: Autor.

Figura 9– Relevância dos periódicos nas referências dos artigos do portfólio bibliográfico final (PB5), Eixo 2



Fonte: Autor.

Figura 10– Autores com maior participação nas referências do portfólio bibliográfico final (PB5), Eixo 2



Fonte: Autor.

E, para complementar a análise bibliométrica, a frequência das palavras-chave pesquisadas (Quadro 12).

Quadro 12– Palavras-chave pesquisadas no portfólio bibliográfico final (PB5), Eixo 2

	Palavras-chave do algoritmo de pesquisa																				
	benchmark	benchmarking	electric sector	electricity companies	electricity firms	electricity sector	energy companies	energy enterprises	energy firms	energy industry	energy sector	machine learning	multicriteria	multi-criteria	multiobjective	multi-objective	performance	power industry	sustainability	sustainable	
<b>Total</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>5</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>2</b>	<b>0</b>	<b>0</b>	<b>5</b>	<b>0</b>	<b>22</b>	<b>3</b>	
Título		1						1						1						3	
Resumo	1						1	3						1			5			17	3
Palavras-chave	1							1												2	

Fonte: Autor.

Assim, conclui-se em relação ao Eixo 2 de pesquisa:

- a) A pesquisa revelou somente dois estudos sobre *benchmarking* em sustentabilidade e desempenho para empresas de energia ou eletricidade com apoio multicritério: Andreopoulou; Koliouka (2018) e Hendiani *et al.* (2020). Porém, não houve menção a *machine learning* (Quadro 12);
- b) *Journal of Cleaner Production* foi a publicação mais relevante nas referências das publicações no portfólio bibliográfico final (PB5);
- c) Andreopoulou, Z. foi o autor mais citado nas referências em PB5.

### 3.4.3. REVISÃO DA LITERATURA – EIXO 3

O Quadro 13 ilustra o portfólio bibliográfico final (PB5).

## Quadro 13– Artigos que formam o portfólio bibliográfico final (PB5), Eixo 3

1. (AHMAD <i>et al.</i> , 2021a) Ahmad, M.; Jan, I.; Jabeen, G.; Alvarado, R. (2021), <i>Does energy-industry investment drive economic performance in regional China: Implications for sustainable development</i> , <i>Sustainable Production and Consumption</i> , 27, 176-192.
2. (AHMAD <i>et al.</i> , 2021b) Ahmad, T.; Zhang, D.; Huang, C.; Zhang, H.; Dai, N.; Song, Y.; Chen, H. (2021), <i>Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities</i> , <i>Journal of Cleaner Production</i> , <i>Artificial intelligence in sustainable</i> , 289, <a href="https://doi.org/10.1016/j.jclepro.2021.125834">https://doi.org/10.1016/j.jclepro.2021.125834</a> .
3. (BARROS <i>et al.</i> , 2020) Barros, J. J. C.; Coira, M. L.; Lopez, M. P. C.; Gochi, A. C.; Soares, I. (2020), <i>Probabilistic multicriteria environmental assessment of power plants: A global approach.</i> , <i>Applied Energy</i> , 260.
4. (CAIADO <i>et al.</i> , 2020) Caiado, R. G. G.; Heymann, M. C.; Silveira, C. L. R.; Meza, L. A.; Quelhas, O. L. G. (2020), <i>Measuring the eco-efficiency of Brazilian energy companies using DEA and directional distance function</i> , <i>IEEE LATIN AMERICA TRANSACTIONS</i> , 18, 11, 1844-1852.
5. (CHAMANDOUST <i>et al.</i> , 2020) Chamandoust, H.; Derakhshan, G.; Bahramara, S. (2020), <i>Multi-objective performance of smart hybrid energy system with Multi-optimal participation of customers in day-ahead energy market</i> , <i>Energy and Buildings</i> , 216.
6. (COLLA <i>et al.</i> , 2020) Colla, M.; Ioannou, A.; Falcone, G. (2020), <i>Critical review of competitiveness indicators for energy projects.</i> , <i>Renewable &amp; Sustainable Energy Reviews</i> , 125.
7. (DASHTPEYMA; GHODSI, 2019) Dashtpeyma, M.; Ghodsi, R. (2019), <i>Developing the resilient solar energy management system: a hybrid qualitative-quantitative approach</i> , <i>International Journal of Ambient Energy</i> , <a href="https://doi.org/10.1080/01430750.2019.1630301">https://doi.org/10.1080/01430750.2019.1630301</a> .
8. (EBRAHIMI; RAHMANI, 2019) Ebrahimi, M.; Rahmani, D. (2019), <i>A five-dimensional approach to sustainability for prioritizing energy production systems using a revised GRA method: A case study</i> , <i>Renewable Energy</i> , 135, 345-354.
9. (ERVURAL <i>et al.</i> , 2018a) Ervural, B. C.; Evren, R.; Delen, D. (2018), <i>A multi-objective decision-making approach for sustainable energy investment planning</i> , <i>Renewable Energy</i> , 126, 387-402.
10. (ERVURAL <i>et al.</i> , 2018b) Ervural, B. C.; Zaim, S.; Demirel, O. F.; Aydin, Z.; Delen, D. (2018), <i>An ANP and fuzzy TOPSIS-based SWOT analysis for Turkey's energy planning</i> , <i>Renewable and Sustainable Energy Reviews</i> , 82, 1538-1550.
11. (GARDAZI <i>et al.</i> , 2020) Gardazi, S.S.N.; Hassan, A.F.S.; Johari, J.B. (2020), <i>Board of Directors Attributes and Sustainability Performance in the Energy Industry</i> , <i>Journal of Asian Finance, Economics and Business</i> , 12, 317-328.
12. (HASSAN, 2019) Hassan, A. (2019), <i>Do renewable energy incentive policies improve the performance of energy firms? Evidence from OECD countries</i> , <i>OPEC Energy Review</i> , 43, 168-192.
13. (HERBOHN <i>et al.</i> , 2014) Herbohn, K.; Walker, J.; Loo, H.Y.M. (2014), <i>Corporate Social Responsibility: The Link Between Sustainability Disclosure and Sustainability Performance</i> , <i>Abacus</i> , 50, 422-459.
14. (IRFAN <i>et al.</i> , 2019) Irfan, M.; Zhao, Z.-Y.; Ahmad, M.; Batool, K.; Jan, A.; Mukeshimana, M.C. (2019), <i>Competitive assessment of Indian wind power industry: A five forces model</i> , <i>Journal of Renewable and Sustainable Energy</i> , 11.

Fonte: Autor.

## Quadro 13 (cont.) – Artigos que formam o portfólio bibliográfico final (PB5), Eixo 3

15. (JARVIS; SOVACOOOL, 2011) Jarvis, D.S.L.; Sovacool, B.K. (2011), <i>Conceptualizing and evaluating best practices in electricity and water regulatory governance</i> , <i>Energy</i> , 36, 4340-4352.
16. (KARNITIS <i>et al.</i> , 2021) Karnitis, E.; Bicevskis, J.; Karnitis, G. (2021), <i>Measuring the implementation of the agenda 2030 vision in its comprehensive sense: Methodology and tool</i> , <i>Energies</i> , 14, 856, <a href="https://doi.org/10.3390/en14040856">https://doi.org/10.3390/en14040856</a> .
17. (KOLOSOK <i>et al.</i> , 2020) Kolosok, S.; Pimonenko, T.; Yevdokymova, A.; Hajiyev, N. O.; Palienko, M.; Prasol, L. (2020), <i>ENERGY EFFICIENCY POLICY: IMPACT OF GREEN INNOVATIONS, MARKETING AND MANAGEMENT OF INNOVATIONS</i> , 50-60.
18. (KWAKKEL; PRUYT, 2013) Kwakkel, J.H.; Pruyt, E. (2013), <i>Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty</i> , <i>Technological Forecasting and Social Change</i> , 80, 419-431.
19. (LI; TAO, 2017) Li, M.-J.; Tao, W.-Q. (2017), <i>Review of methodologies and polices for evaluation of energy efficiency in high energy-consuming industry</i> , <i>Applied Energy</i> , 187, 203-215.
20. (LI <i>et al.</i> , 2016) Li, Y.; Shi, X.; Yao, L. (2016), <i>Evaluating energy security of resource-poor economies: A modified principle component analysis approach</i> , <i>Energy Economics</i> , 58, 211-221.
21. (LI; DONG, 2020) Li, Z.; Dong, Z. (2020), <i>Comprehensive performance evaluation of Chinese nuclear power listed companies based on HSE perspective</i> , <i>IOP Conference Series: Earth and Environmental Science</i> , 510.
22. (MILOJEVIC <i>et al.</i> , 2020) Milojevic, M.; Urbanski, M.; Terzic, I.; Prasolov, V. (2020), <i>Impact of non-financial factors on the effectiveness of audits in energy companies</i> , <i>Energies</i> , 13.
23. (PACHECO <i>et al.</i> , 2018) Pacheco, L. M.; Alves, M. F. R.; Liboni, L. B. (2018), <i>Green absorptive capacity: A mediation-moderation model of knowledge for innovation</i> , <i>Business Strategy and the Environment</i> , 27, 1502-1513.
24. (PAPAGEORGIU <i>et al.</i> , 2020) Papageorgiou, K.; Carvalho, G.; Papageorgiou, E.I.; Bochtis, D.; Stamoulis, G. (2020), <i>Decision-making process for photovoltaic solar energy sector development using fuzzy cognitive map technique</i> , <i>Energies</i> , 13.
25. (PATARI <i>et al.</i> , 2014) Patari, S.; Arminen, H.; Tuppara, A.; Jantunen, A. (2014), <i>Competitive and responsible? The relationship between corporate social and financial performance in the energy sector.</i> , <i>Renewable &amp; Sustainable Energy Reviews</i> , 37, 142 - 154.
26. (PATARI <i>et al.</i> , 2012) Patari, S.; Jantunen, A.; Kylaheiko, K.; Sandstrom, J. (2012), <i>Does Sustainable Development Foster Value Creation? Empirical Evidence from the Global Energy Industry</i> , <i>Corporate Social Responsibility and Environmental Management</i> , 19, 317-326.
27. (PHILLIS <i>et al.</i> , 2020) Phillis, A.; Grigoroudis, E.; Kouikoglou, V.S. (2020), <i>Assessing national energy sustainability using multiple criteria decision analysis</i> , <i>International Journal of Sustainable Development and World Ecology</i> , 1-18.

Fonte: Autor.

## Quadro 13 (cont.) – Artigos que formam o portfólio bibliográfico final (PB5), Eixo 3

28. (QI <i>et al.</i> , 2020) Qi, W.; Huang, Z.; Dinçer, H.; Korsakiene, R.; Yuksel, S. (2020), <i>Corporate governance-based strategic approach to sustainability in energy industry of emerging economies with a novel interval-valued intuitionistic fuzzy hybrid decision making model</i> , <i>Sustainability (Switzerland)</i> , 12.
29. (RAFAJ <i>et al.</i> , 2006) Rafaj, P. ; Barreto, L.; Kypreos, S. (2006), <i>Combining policy instruments for sustainable energy systems: An assessment with the GMM model</i> , <i>Environmental Modeling and Assessment</i> , 11, 277-295.
30. (RIGO <i>et al.</i> , 2020) Rigo, P.D.; Rediske, G.; Rosa, C.B.; Gastaldo, N.G.; Michels, L.; Junior, A.L.N.; Siluk, J.C.M. (2020), <i>Renewable energy problems: Exploring the methods to support the decision-making process</i> , <i>Sustainability (Switzerland)</i> , 12, 1-27.
31. (ROSA <i>et al.</i> , 2020) Rosa, C.B.; Siluk, J.C.M.; Rigo, P.D.; Rediske, G.; Burin, H.P.; Michels, L. (2020), <i>Organizational Performances of Distributed Generation in Brazil Electric Utilities: A Balanced Scorecard Perspective</i> , <i>Springer Proceedings in Mathematics and Statistics</i> , 337, 101-114.
32. (SAHABUDDIN; KHAN, 2021) Sahabuddin, M.; Khan, I. (2021), <i>Multi-criteria decision analysis methods for energy sector's sustainability assessment: Robustness analysis through criteria weight change</i> , <i>Sustainable Energy Technologies and Assessments</i> , 47, <a href="https://doi.org/10.1016/j.seta.2021.101380">https://doi.org/10.1016/j.seta.2021.101380</a> .
33. (SALMI <i>et al.</i> , 2019) Salmi, S.; Sukur, A.L.A.; Norlena, H. (2019), <i>The need of corporate social responsibility (CSR) implementation in energy industry: Proposition development</i> , <i>International Journal of Recent Technology and Engineering</i> , 8, 201-207.
34. (SARANGI <i>et al.</i> , 2019) Sarangi, G.K.; Mishra, A.; Chang, Y.; Taghizadeh-Hesary, F. (2019), <i>Indian electricity sector, energy security and sustainability: An empirical assessment</i> , <i>Energy Policy</i> , 135.
35. (SHAHBAZ <i>et al.</i> , 2020) Shahbaz, M.; Karaman, A.S.; Kilic, M.; Uyar, A. (2020), <i>Board attributes, CSR engagement, and corporate performance: What is the nexus in the energy sector?</i> , <i>Energy Policy</i> , 143.
36. (SI <i>et al.</i> , 2020) Si, K.; Xu, X.L.; Chen, H.H. (2020), <i>Examining the interactive endogeneity relationship between R&amp;D investment and financially sustainable performance: Comparison from different types of energy enterprises</i> , <i>Energies</i> , 13.
37. (SUEYOSHI; GOTO, 2014) Sueyoshi, T.; Goto, M. (2014), <i>Environmental assessment for corporate sustainability by resource utilization and technology innovation: DEA radial measurement on Japanese industrial sectors</i> , <i>Energy Economics</i> , 46, 295-307.
38. (SUEYOSHI; GOTO, 2017) Sueyoshi, T.; Goto, M. (2017), <i>World trend in energy: an extension to DEA applied to energy and environment</i> , <i>Journal of Economic Structures</i> , 6.
39. (SUEYOSHI <i>et al.</i> , 2021) Sueyoshi, T.; Ryu, Y.; Yun, J-Y (2021), <i>Covid-19 response and prospects of clean/sustainable energy transition in industrial nations: New environmental assessment</i> , <i>Energies</i> , 14, 1174, <a href="https://doi.org/10.3390/en14041174">https://doi.org/10.3390/en14041174</a> .
40. (SUN <i>et al.</i> , 2020) Sun, J.; Xu, S.; Li, G. (2020), <i>Analyzing sustainable power supply chain performance: Evidence from China's provincial regions</i> , <i>Journal of Enterprise Information Management</i> , DOI 10.1108/JEIM-09-2019-0296.

Fonte: Autor.



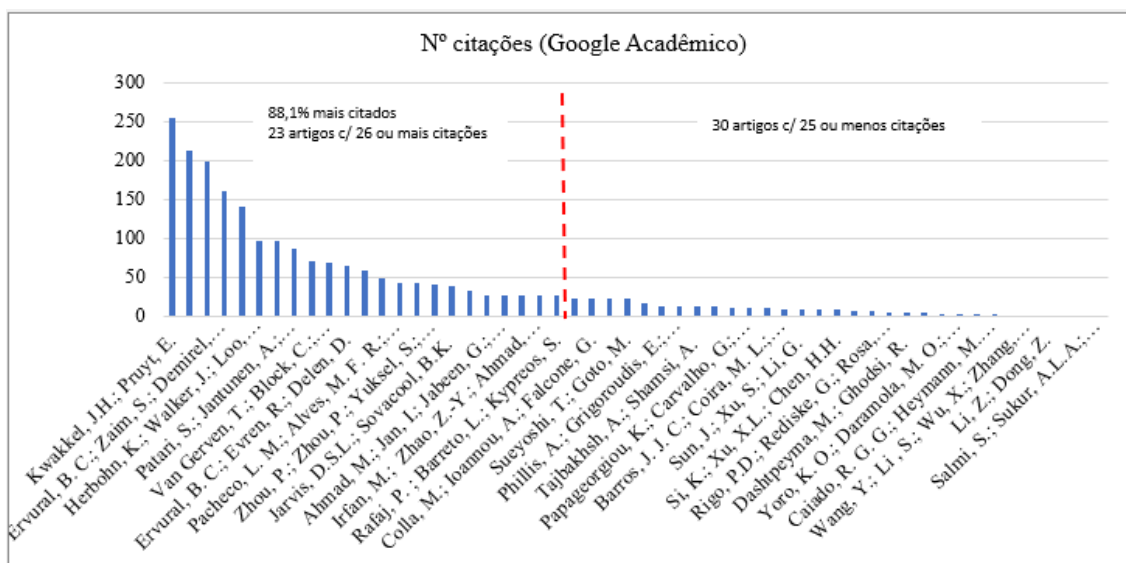
## Quadro 13 (cont.) – Artigos que formam o portfólio bibliográfico final (PB5), Eixo 3

41. (TAJBAKSH; SHAMSI, 2019) Tajbakhsh, A.; Shamsi, A. (2019), <i>A facility location problem for sustainability-conscious power generation decision makers</i> , <i>Journal of Environmental Management</i> , 230, 319 - 334.
42. (TALBOT; BOIRAL, 2018) Talbot, D.; Boiral, O. (2018), <i>GHG reporting and impression management: An assessment of sustainability reports from the energy sector</i> , <i>Journal of Business Ethics</i> , 147, 367-383.
43. (VAN GERVEN <i>et al.</i> , 2007) Van Gerven, T.; Block, C.; Geens, J.; Cornelis, G.; Vandecasteele, C. (2007), <i>Environmental response indicators for the industrial and energy sector in Flanders</i> , <i>Journal of Cleaner Production</i> , 15, 886-894.
44. (VIEIRA <i>et al.</i> , 2017) Vieira, R.; O'Dwyer, B.; Schneider, R. (2017), <i>Aligning Strategy and Performance Management Systems: The Case of the Wind-Farm Industry</i> , <i>ORGANIZATION &amp; ENVIRONMENT</i> , 30, 3-26.
45. (WANG <i>et al.</i> , 2014) Wang, D.; Li, S.; Sueyoshi, T. (2014), <i>DEA environmental assessment on U.S. industrial sectors: Investment for improvement in operational and environmental performance to attain corporate sustainability</i> , <i>Energy Economics</i> , 45, 254 - 267.
46. (WANG <i>et al.</i> , 2021) Wang, Y.; Li, S.; Wu, X.; Zhang, Y.; Li, B.; Gao, I. (2021), <i>Using sustainable performance prediction in data-scarce scenarios: A study of park-level integrated microgrid projects in Tianjin, China</i> , <i>Journal of Cleaner Production</i> , 304, <a href="https://doi.org/10.1016/j.jclepro.2021.127042">https://doi.org/10.1016/j.jclepro.2021.127042</a> .
47. (WANKE <i>et al.</i> , 2020) Wanke, P.; Tan, Y.; Antunes, J.; Hadi-Vencheh, A. (2020), <i>Business environment drivers and technical efficiency in the Chinese energy industry: A robust Bayesian stochastic frontier analysis</i> , <i>Computers and Industrial Engineering</i> , 144.
48. (XU; CHEN, 2020) Xu, X.-L.; Chen, H.H. (2020), <i>Exploring the relationships between environmental management and financial sustainability in the energy industry: Linear and nonlinear effects</i> , <i>Energy and Environment</i> , 31, 1281-1300.
49. (XU; LIU, 2019) Xu, X.-L.; Liu, C.K. (2019), <i>How to keep renewable energy enterprises to reach economic sustainable performance: From the views of intellectual capital and life cycle</i> , <i>Energy, Sustainability and Society</i> , 9, <a href="https://doi.org/10.1186/s13705-019-0187-2">https://doi.org/10.1186/s13705-019-0187-2</a> .
50. (YANG; ZHANG, 2020) Yang, L.; Zhang, Y. (2020), <i>An evaluation of the green performance of Chinese new energy enterprises from the perspective of social responsibility</i> , <i>Nature Environment and Pollution Technology</i> , 19, 1527-1535.
51. (YORO <i>et al.</i> , 2021) Yoro, K. O.; Daramola, M. O.; Sekoai, P. T.; Wilson, U. N.; Eterigho-Ikelegbe, O. (2021), <i>Update on current approaches, challenges, and prospects of modeling and simulation in renewable and sustainable energy systems</i> , <i>Renewable and Sustainable Energy Reviews</i> , 150, <a href="https://doi.org/10.1016/j.rser.2021.111506">https://doi.org/10.1016/j.rser.2021.111506</a> .
52. (YUSIFBAYLI; NASIBOV, 2020) Yusifbayli, N.; Nasibov, V. (2020), <i>Comparative analysis of Azerbaijan's energy sector efficiency trend at the current development stage</i> , <i>E3S Web of Conferences</i> , 209.
53. (ZHOU <i>et al.</i> , 2019) Zhou, P.; Zhou, P.; Yuksel, S.; Dincer, H.; Uluer, G. S. (2019), <i>Balanced scorecard-based evaluation of sustainable energy investment projects with IT2 fuzzy hybrid decision making approach</i> , <i>Energies</i> , 13.

Fonte: Autor.

A seguir, apresenta-se a análise bibliométrica sobre o portfólio bibliográfico final (PB5): os artigos e autores mais relevantes, com o uso de “linha de corte” de no mínimo 85% do total de citações no Google Acadêmico (23 artigos com 26 ou mais citações representam 88,1% do total de citações) – Figura 11; e a relação de artigos e autores em ordem decrescente de citações – Quadro 14.

Figura 11– Linha de corte (tracejado, em vermelho) para o portfólio bibliográfico final (PB5), Eixo 3



Fonte: Autor.

Quadro 14– Número de citações do portfólio bibliográfico final (PB5), Eixo 3

<b>Autor(es)</b>	<b>Artigo</b>	<b>Ano</b>	<b>Nº citações (Google Acadêmico)</b>
Kwakkel, J.H.; Pruyt, E.	<i>Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty</i>	2013	254
Li, M.-J.; Tao, W.-Q.	<i>Review of methodologies and polices for evaluation of energy efficiency in high energy-consuming industry</i>	2017	213
Ervural, B. C.; Zaim, S.; Demirel, O. F.; Aydin, Z.; Delen, D.	<i>An ANP and fuzzy TOPSIS-based SWOT analysis for Turkey's energy planning</i>	2018	198
Talbot, D.; Boiral, O.	<i>GHG reporting and impression management: An assessment of sustainability reports from the energy sector</i>	2018	160
Herbohn, K.; Walker, J.; Loo, H.Y.M.	<i>Corporate Social Responsibility: The Link Between Sustainability Disclosure and Sustainability Performance</i>	2014	141
Wang, D.; Li, S.; Sueyoshi, T.	<i>DEA environmental assessment on U.S. industrial sectors: Investment for improvement in operational and environmental performance to attain corporate sustainability</i>	2014	97
Patari, S.; Jantunen, A.; Kylaheiko, K.; Sandstrom, J.	<i>Does Sustainable Development Foster Value Creation? Empirical Evidence from the Global Energy Industry</i>	2012	97
Patari, S.; Arminen, H.; Tuppara, A.; Jantunen, A.	<i>Competitive and responsible? The relationship between corporate social and financial performance in the energy sector</i>	2014	87
Van Gerven, T.; Block, C.; Geens, J.; Cornelis, G.; Vandecasteele, C.	<i>Environmental response indicators for the industrial and energy sector in Flanders</i>	2007	70
Sueyoshi, T.; Goto, M.	<i>Environmental assessment for corporate sustainability by resource utilization and technology innovation: DEA radial measurement on Japanese industrial sectors</i>	2014	68
Ervural, B. C.; Evren, R.; Delen, D.	<i>A multi-objective decision-making approach for sustainable energy investment planning</i>	2018	65
Li, Y.; Shi, X.; Yao, L.	<i>Evaluating energy security of resource-poor economies: A modified principle component analysis approach</i>	2016	58

Fonte: Autor.

Quadro 14 (cont.) – Número de citações do portfólio bibliográfico final (PB5), Eixo 3

<b>Autor(es)</b>	<b>Artigo</b>	<b>Ano</b>	<b>Nº citações (Google Acadêmico)</b>
Pacheco, L. M.; Alves, M. F. R.; Liboni, L. B.	<i>Green absorptive capacity: A mediation-moderation model of knowledge for innovation</i>	2018	48
Ahmad, T.; Zhang, D.; Huang, C.; Zhang, H.; Dai, N.; Song, Y.; Chen, H.	<i>Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities</i>	2021	42
Zhou, P.; Zhou, P.; Yuksel, S.; Dincer, H.; Uluer, G. S.	<i>Balanced scorecard-based evaluation of sustainable energy investment projects with IT2 fuzzy hybrid decision making approach</i>	2019	42
Shahbaz, M.; Karaman, A.S.; Kilic, M.; Uyar, A.	<i>Board attributes, CSR engagement, and corporate performance: What is the nexus in the energy sector?</i>	2020	41
Jarvis, D.S.L.; Sovacool, B.K.	<i>Conceptualizing and evaluating best practices in electricity and water regulatory governance</i>	2011	38
Sarangi, G.K.; Mishra, A.; Chang, Y.; Taghizadeh- Hesary, F.	<i>Indian electricity sector, energy security and sustainability: An empirical assessment</i>	2019	33
Ahmad, M.; Jan, I.; Jabeen, G.; Alvarado, R.	<i>Does energy-industry investment drive economic performance in regional China: Implications for sustainable development</i>	2021	26
Qi, W.; Huang, Z.; Dinçer, H.; Korsakiene, R.; Yuksel, S.	<i>Corporate governance-based strategic approach to sustainability in energy industry of emerging economies with a novel interval-valued intuitionistic fuzzy hybrid decision making model</i>	2020	26
Irfan, M.; Zhao, Z.-Y.; Ahmad, M.; Batool, K.; Jan, A.; Mukeshimana, M.C.	<i>Competitive assessment of Indian wind power industry: A five forces model</i>	2019	26
Vieira, R.; O'Dwyer, B.; Schneider, R.	<i>Aligning Strategy and Performance Management Systems: The Case of the Wind-Farm Industry</i>	2017	26
Rafaj, P. ; Barreto, L.; Kypreos, S.	<i>Combining policy instruments for sustainable energy systems: An assessment with the GMM model</i>	2006	26

Fonte: Autor.

Quadro 14 (cont.) – Número de citações do portfólio bibliográfico final (PB5), Eixo 3

<b>Autor(es)</b>	<b>Artigo</b>	<b>Ano</b>	<b>Nº citações (Google Acadêmico)</b>
Chamandoust, H.; Derakhshan, G.; Bahramara, S.	<i>Multi-objective performance of smart hybrid energy system with Multi-optimal participation of customers in day-ahead energy market</i>	2020	23
Colla, M.; Ioannou, A.; Falcone, G.	<i>Critical review of competitiveness indicators for energy projects</i>	2020	23
Xu, X.-L.; Liu, C.K.	<i>How to keep renewable energy enterprises to reach economic sustainable performance: From the views of intellectual capital and life cycle</i>	2019	23
Sueyoshi, T.; Goto, M.	<i>World trend in energy: an extension to DEA applied to energy and environment</i>	2017	22
Ebrahimi, M.; Rahmani, D.	<i>A five-dimensional approach to sustainability for prioritizing energy production systems using a revised GRA method: A case study</i>	2019	17
Phillis, A.; Grigoroudis, E.; Kouikoglou, V.S.	<i>Assessing national energy sustainability using multiple criteria decision analysis</i>	2020	12
Kolosok, S.; Pimonenko, T.; Yevdokymova, A.; Hajiyev, N. O.; Palienko, M.; Prasol, L.	<i>ENERGY EFFICIENCY POLICY: IMPACT OF GREEN INNOVATIONS</i>	2020	12
Tajbakhsh, A.; Shamsi, A.	<i>A facility location problem for sustainability-conscious power generation decision makers</i>	2019	12
Hassan, A.	<i>Do renewable energy incentive policies improve the performance of energy firms? Evidence from OECD countries</i>	2019	12
Papageorgiou, K.; Carvalho, G.; Papageorgiou, E.I.; Bochtis, D.; Stamoulis, G.	<i>Decision-making process for photovoltaic solar energy sector development using fuzzy cognitive map technique</i>	2020	10
Xu, X.-L.; Chen, H.H.	<i>Exploring the relationships between environmental management and financial sustainability in the energy industry: Linear and nonlinear effects</i>	2020	10
Barros, J. J. C.; Coira, M. L.; Lopez, M. P. C.; Gochi, A. C.; Soares, I.	<i>Probabilistic multicriteria environmental assessment of power plants: A global approach.</i>	2020	10

Fonte: Autor.

Quadro 14 (cont.) – Número de citações do portfólio bibliográfico final (PB5), Eixo 3

<b>Autor(es)</b>	<b>Artigo</b>	<b>Ano</b>	<b>Nº citações (Google Acadêmico)</b>
Sueyoshi, T.; Ryu, Y.; Yun, J-Y	<i>Covid-19 response and prospects of clean/sustainable energy transition in industrial nations: New environmental assessment</i>	2021	9
Sun, J.; Xu, S.; Li, G.	<i>Analyzing sustainable power supply chain performance: Evidence from China's provincial regions</i>	2020	9
Wanke, P.; Tan, Y.; Antunes, J.; Hadi-Vencheh, A.	<i>Business environment drivers and technical efficiency in the Chinese energy industry: A robust Bayesian stochastic frontier analysis</i>	2020	8
Si, K.; Xu, X.L.; Chen, H.H.	<i>Examining the interactive endogeneity relationship between R&amp;D investment and financially sustainable performance: Comparison from different types of energy enterprises</i>	2020	8
Sahabuddin, M.; Khan, I.	<i>Multi-criteria decision analysis methods for energy sector's sustainability assessment: Robustness analysis through criteria weight change</i>	2021	6
Rigo, P.D.; Rediske, G.; Rosa, C.B.; Gastaldo, N.G.; Michels, L.; Junior, A.L.N.; Siluk, J.C.M.	<i>Renewable energy problems: Exploring the methods to support the decision-making process</i>	2020	6
Gardazi, S.S.N.; Hassan, A.F.S.; Johari, J.B.	<i>Board of Directors Attributes and Sustainability Performance in the Energy Industry</i>	2020	5
Dashtpeyma, M.; Ghodsí, R.	<i>Developing the resilient solar energy management system: a hybrid qualitative-quantitative approach</i>	2019	5
Milojevic, M.; Urbanski, M.; Terzic, I.; Prasolov, V.	<i>Impact of non-financial factors on the effectiveness of audits in energy companies</i>	2020	4
Yoro, K. O.; Daramola, M. O.; Sekoai, P. T.; Wilson, U. N.; Eterigho- Ikelegbe, O.	<i>Update on current approaches, challenges, and prospects of modeling and simulation in renewable and sustainable energy systems</i>	2021	2
Yusifbayli, N.; Nasibov, V	<i>Comparative analysis of Azerbaijan's energy sector efficiency trend at the current development stage</i>	2020	2

Fonte: Autor.

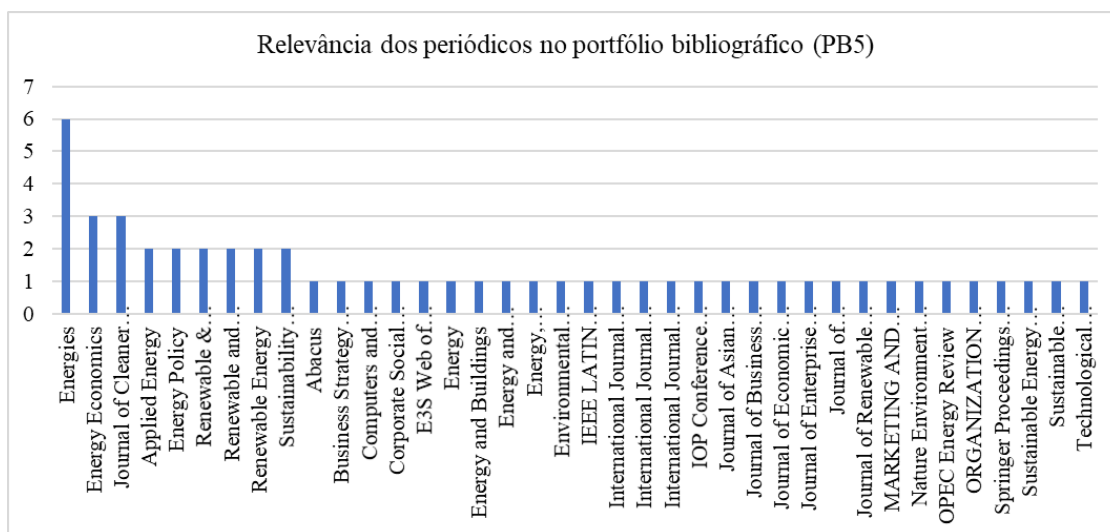
Quadro 14 (cont.) – Número de citações do portfólio bibliográfico final (PB5), Eixo 3

<b>Autor(es)</b>	<b>Artigo</b>	<b>Ano</b>	<b>Nº citações (Google Acadêmico)</b>
Caiado, R. G. G.; Heymann, M. C.; Silveira, C. L. R.; Meza, L. A.; Quelhas, O. L. G.	<i>Measuring the eco-efficiency of Brazilian energy companies using DEA and directional distance function</i>	2020	2
Karnitis, E.; Bicevskis, J.; Karnitis, G.	<i>Measuring the implementation of the agenda 2030 vision in its comprehensive sense: Methodology and tool</i>	2021	1
Wang, Y.; Li, S.; Wu, X.; Zhang, Y.; Li, B.; Gao, I.	<i>Using sustainable performance prediction in data-scarce scenarios: A study of park-level integrated microgrid projects in Tianjin, China</i>	2021	0
Yang, L.; Zhang, Y.	<i>An evaluation of the green performance of Chinese new energy enterprises from the perspective of social responsibility</i>	2020	0
Li, Z.; Dong, Z.	<i>Comprehensive performance evaluation of Chinese nuclear power listed companies based on HSE perspective</i>	2020	0
Rosa, C.B.; Siluk, J.C.M.; Rigo, P.D.; Rediske, G.; Burin, H.P.; Michels, L.	<i>Organizational Performances of Distributed Generation in Brazil Electric Utilities: A Balanced Scorecard Perspective</i>	2020	0
Salmi, S.; Sukur, A.L.A.; Norlena, H.	<i>The need of corporate social responsibility (CSR) implementation in energy industry: Proposition development</i>	2019	0

Fonte: Autor.

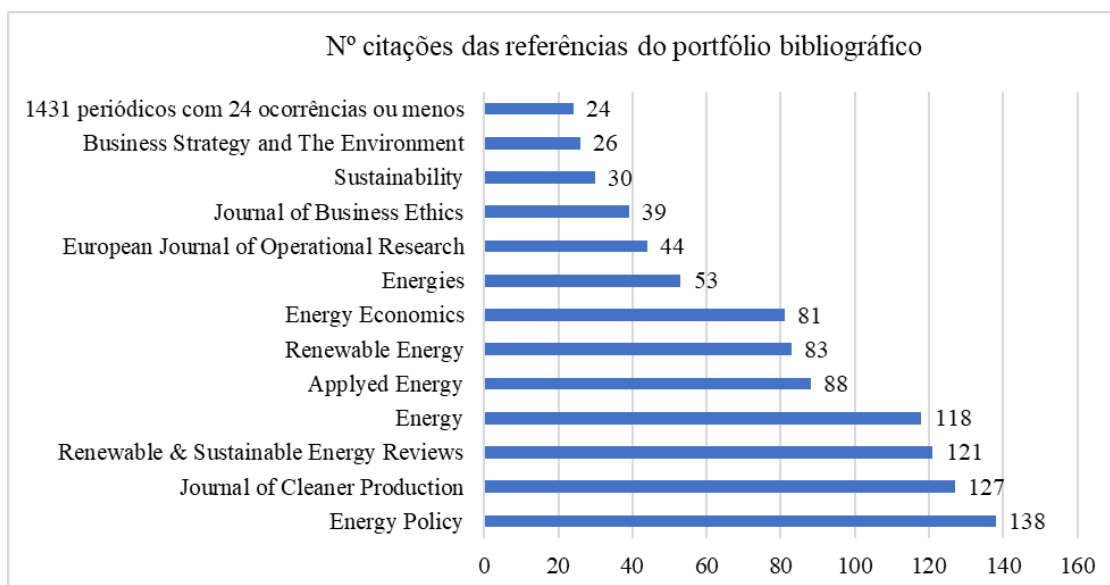
De forma complementar, a relevância dos periódicos no portfólio bibliográfico final (PB5) e referências – Figuras 12 e 13, respectivamente.

Figura 12– Relevância dos periódicos no portfólio bibliográfico final (PB5), Eixo 3



Fonte: Autor.

Figura 13– Relevância dos periódicos nas referências do portfólio bibliográfico final (PB5), Eixo 3

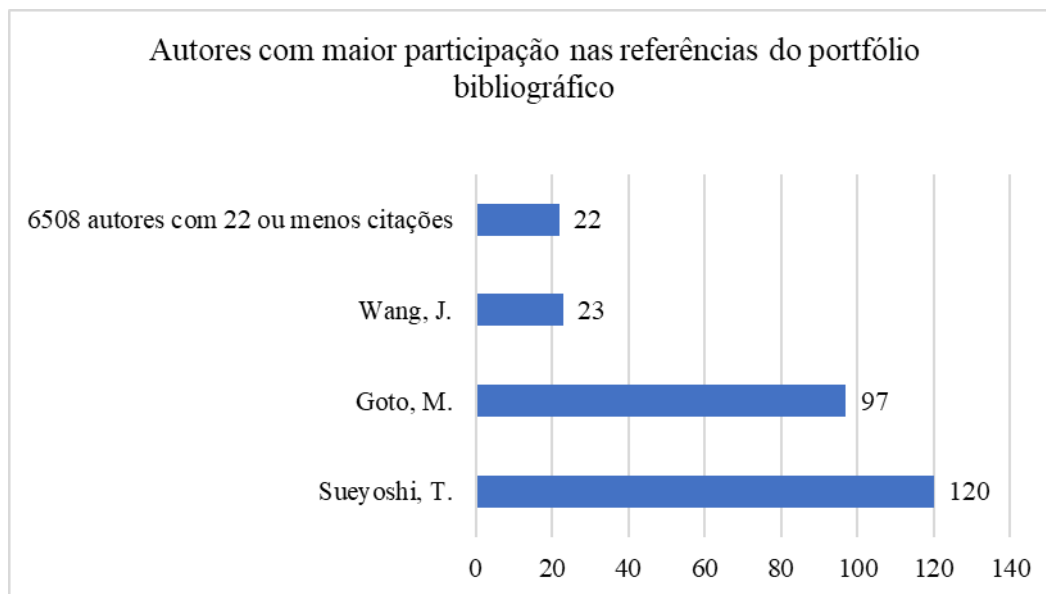


Fonte: Autor.

E, a relevância dos autores nas referências do PB5 – Figura 14.



Figura 14– Relevância dos autores nas referências do portfólio bibliográfico final (PB5), Eixo 3



Fonte: Autor.

Para concluir a análise bibliométrica, a frequência das palavras-chave pesquisadas – Quadro 15.

Quadro 15– Relação das palavras-chave pesquisadas no portfólio bibliográfico final (PB5), Eixo 3

	Palavras-chave do algoritmo de pesquisa																								
	appraisal	assessment	electric sector	electricity companies	electricity firms	electricity sector	energy companies	energy enterprises	energy firms	energy industry	energy sector	evaluation	example	framework	index	indicator	indice	measurement	metric	model	performance	power industry	sustainability	sustainable	template
<b>Total</b>	<b>0</b>	<b>35</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>8</b>	<b>13</b>	<b>12</b>	<b>6</b>	<b>37</b>	<b>43</b>	<b>25</b>	<b>2</b>	<b>20</b>	<b>10</b>	<b>34</b>	<b>5</b>	<b>11</b>	<b>5</b>	<b>91</b>	<b>179</b>	<b>14</b>	<b>95</b>	<b>87</b>	<b>0</b>
Título	0	9	0	0	0	1	2	3	1	7	7	4	0	0	0	2	0	1	0	6	16	1	12	12	0
Resumo	0	22	1	0	0	6	11	6	5	25	33	18	2	20	9	27	5	8	5	79	141	13	68	60	0
Palavras-chave	0	4	0	0	0	1	0	3	0	5	3	3	0	0	1	5	0	2	0	6	22	0	15	15	0

Fonte: Autor.

Assim, conclui-se em relação ao Eixo 3 de pesquisa:

- São 53 estudos sobre modelos de avaliação da sustentabilidade e desempenho para empresas de energia ou eletricidade, conforme portfólio bibliográfico final (PB5);
- Energies* foi a publicação mais relevante no PB5;
- Energy Police* e *Journal of Cleaner Production* foram as duas publicações mais citadas nas referências do PB5;
- Sueyoshi, T. e Goto, M. foram os autores mais citados nas referências em PB5;

- e) Kwakkel; Pruyt (2013) e Li; Tao (2017) foram os dois estudos mais citados no Google Acadêmico.

### 3.5. ATUALIZAÇÃO DA REVISÃO DA LITERATURA

Visando complementar a revisão da literatura, realizou-se uma busca nas mesmas bases da CAPES considerando-se os três Eixos de pesquisa, para o período de janeiro/2022 a agosto/2023. Houve o retorno de 27, 4 e 452 estudos para os Eixos 1, 2 e 3, respectivamente. Após a leitura dos títulos, resumos, verificação da aderência dos estudos à pesquisa e exclusão de estudos repetidos, foram selecionados estudos somente para os Eixos 2 e 3. Como resultado, houve a seleção de um e nove estudos para os Eixos 2 e 3, respectivamente, conforme Quadro 16.

Quadro 16- Complemento à revisão sistemática da literatura, período janeiro/2022-agosto/2023

Eixo	Estudos
2	(PAPAPOSTOLOU <i>et al.</i> , 2022) Papapostolou, A.; Mexis, F. D.; Karakosta, C.; Psarras, J. (2022), A Multicriteria Tool to Support Decision-Making in the Early Stages of Energy Efficiency Investments, 8th International Conference on Decision Support System Technology, ICDSST 2022, DOI: 10.1007/978-3-031-06530-9_15.
3	(ADAMKAITE <i>et al.</i> , 2022) Adamkaite, J.; Streimikiene, D.; Rudzioniene, K. (2022), The impact of social responsibility on corporate financial performance in the energy sector: Evidence from Lithuania, Corp Soc Responsib Environ Manag., 1-14, DOI: 10.1002/csr.2340.
	(ARORA <i>et al.</i> , 2023) Arora, H.; Anwer, N.; Narula, S. A. (2023), A Study of Corporate Sustainability Reporting Practices in the Indian Energy Sector, Strategic Planning for Energy and the Environment, 42_3, 529–562, doi: 10.13052/spee1048-5236.4235.
	(DONG <i>et al.</i> , 2023) Dong, R.; Shao, C.; Xin, S.; Lu, Z. (2023), A Sustainable Development Evaluation Framework for Chinese Electricity Enterprises Based on SDG and ESG Coupling, Sustainability, 15, 8960, <a href="https://doi.org/10.3390/su15118960">https://doi.org/10.3390/su15118960</a> .
	(ERSOY <i>et al.</i> , 2023) Ersoy, N.; Taslak, S. (2023), Comparative Analysis of MCDM Methods for the Assessment of Corporate Sustainability Performance in Energy Sector, Ege Academic Review, 23,03, 341-362, doi: 10.21121/eab.986122.
	(KUMAR <i>et al.</i> , 2022) Kumar, A.; Gupta, J.; Das, N. (2022), Revisiting the influence of corporate sustainability practices on corporate financial performance: An evidence from the global energy sector, Bus Strat Env., 31, 3231–3253, DOI: 10.1002/bse.3073.
	(MAKRIDOU <i>et al.</i> , 2023) Makridou, G.; Doumpos, M.; Lemonakis, C. (2023), Relationship between ESG and corporate financial performance in the energy sector: empirical evidence from European companies, International Journal of Energy Sector Management, DOI 10.1108/IJESM-01-2023-0012.
	(PATYAL <i>et al.</i> , 2023) Patyal, V. S.; Kumar, R.; Lamba, K.; Maheshwari, S. (2023), Performance evaluation of Indian electricity distribution companies: An integrated DEA-IRP-TOPSIS approach, Energy Economics, 124, 106796, <a href="https://doi.org/10.1016/j.eneco.2023.106796">https://doi.org/10.1016/j.eneco.2023.106796</a> .
(PONOMARENKO <i>et al.</i> , 2022) Ponomarenko, T.; Reshneva, E.; Urbano, A. P. M. (2022), Assessment of Energy Sustainability Issues in the Andean Community: Additional Indicators and Their Interpretation, Energies, 15, 3, DOI: 10.3390/en15031077.	
(RAMAIAH; JAYASANKAR, 2022) Ramaiah, V.; Jayasankar, V. (2022), Performance Assessment of Indian Electric Distribution Utilities Using Data Envelopment Analysis (DEA), International Journal of Electrical and Electronic Engineering & Telecommunications, 11, 3, DOI: 10.18178/ijeetc.11.3.192-202.	

Fonte: Autor.

Realizou-se uma busca das palavras-chave das *strings* de pesquisa dos Eixos 2 e 3 frente aos títulos, resumos e palavras-chave dos estudos, conforme Quadros 17 e 18, respectivamente.

Quadro 17- Resultado da pesquisa de palavras-chave, Eixo 2

	Palavras-chave do algoritmo de pesquisa																			
	benchmark	benchmarking	electric sector	electricity companies	electricity firms	electricity sector	energy companies	energy enterprises	energy firms	energy industry	energy sector	machine learning	multicriteria	multi-criteria	multiobjective	multi-objective	performance	power industry	sustainability	sustainable
<b>Total</b>	2	3	0	0	0	0	0	0	0	0	1	0	1	2	0	0	1	0	0	2
Título	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Resumo	2	2	0	0	0	0	0	0	0	0	1	0	0	2	0	0	1	0	0	1
Palavras-chave	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Fonte: Autor.

Quadro 18- Resultado da pesquisa de palavras-chave, Eixo 3

	Palavras-chave do algoritmo de pesquisa																								
	appraisal	assessment	electric sector	electricity companies	electricity firms	electricity sector	energy companies	energy enterprises	energy firms	energy industry	energy sector	evaluation	example	framework	index	indicator	indice	measurement	metric	model	performance	power industry	sustainability	sustainable	template
<b>Total</b>	0	9	0	0	0	3	12	0	1	0	23	10	1	7	11	14	1	2	4	5	37	0	32	27	0
Título	0	3	0	0	0	0	0	0	0	0	5	2	0	1	0	1	0	0	0	0	6	0	4	1	0
Resumo	0	6	0	0	0	3	12	0	1	0	13	8	1	5	9	12	1	1	4	2	26	0	22	23	0
Palavras-chave	0	0	0	0	0	0	0	0	0	0	5	0	0	1	2	1	0	1	0	3	5	0	6	3	0

Fonte: Autor.

### 3.6. ANÁLISE DOS RESULTADOS DA REVISÃO DA LITERATURA

Com base nos portfólios bibliográficos finais (PB5) dos três Eixos de pesquisa e na atualização da revisão sistemática da literatura, depreende-se que, com relação a *benchmarking* sobre sustentabilidade e desempenho para empresas de energia ou eletricidade (Eixo 1), há um estudo sobre a evolução de índices de eficiência técnica e de produtividade de *Malmquist* para 30 maiores empresas distribuidoras de energia elétrica no Brasil, que confronta a metodologia do órgão regulador ANEEL por meio da Análise Envoltória de Dados (GOULART, 2013); e, um segundo, que avalia o desempenho do setor elétrico brasileiro com base nas sustentabilidades econômica, social e ambiental, também por Análise Envoltória de Dados e com base em vários cenários (pesos flexíveis; *Triple Bottom Line*; questões sociais; questões econômicas; problemas ambientais) (SARTORI *et al.*, 2017). Consequentemente, foram identificadas as principais lacunas nos estudos do Eixo 1: i) conjunto mais amplo de indicadores das empresas, o qual permitiria considerar aspectos de governança corporativa, por exemplo; ii) avaliação das empresas por categoria ou classe, a qual permitiria identificar similaridades entre as

empresas; iii) oportunidade ou meios para identificar melhores práticas de gestão das empresas.

Com relação ao Eixo 2, que estende a pesquisa do primeiro Eixo com apoio de método multicritério à tomada de decisão ou *machine learning*, há um estudo que avalia o desempenho de empresas de energia renovável via internet, na Prefeitura de Thessaloniki, Grécia, em relação à sustentabilidade com o uso do método multicritério TOPSIS e que fornece um *ranking* dessas empresas (ANDREOPOULOU; KOLIOUSKA, 2018); um segundo, que propõe um *framework* como *benchmarking* para avaliações da sustentabilidade em sistemas de energia, com 65 atributos-chave e 13 critérios para três pilares de sustentabilidade (econômica, social e ambiental), com o uso de termos linguísticos para expressar o peso e taxa de desempenho para cada atributo, com base no julgamento de especialistas e lógica *fuzzy* (HENDIANI *et al.*, 2020). Um terceiro estudo utiliza o método multicritério ELECTRE Tri para estabelecer um *benchmarking* sobre investimentos em eficiência de energia (PAPAPOSTOULOU *et al.*, 2022). Por conseguinte, foram identificadas as principais lacunas nos estudos do Eixo 2: i) avaliação das empresas por categoria ou classe, a qual permitiria identificar similaridades entre as empresas; ii) análise preditiva de indicadores e classificação de empresas com o uso de *machine learning*; iii) predição de riscos das empresas.

E, em relação ao terceiro Eixo, há 62 (sessenta e dois) estudos que tratam sobre diversos assuntos sobre o setor de energia, inclusive empresas de energia ou eletricidade, a saber:

a) estudos que propõem modelos, *frameworks* ou avaliações de desempenho e sustentabilidade: RAFAJ *et al.* (2006); VAN GERVEN *et al.* (2007); PATARI *et al.* (2012); KWAKKEL; PRUYT (2013); HERBOHN *et al.* (2014); PATARI *et al.* (2014); SUEYOSHI; GOTO (2014); WANG *et al.* (2014); LI *et al.* (2016); SUEYOSHI; GOTO (2017); VIEIRA *et al.* (2017); ERVURAL *et al.* (2018a); PACHECO *et al.* (2018); TALBOT; BOIRAL (2018); DASHTPEYMA; GHODSI (2019); EBRAHIMI; RAHMANI (2019); HASSAN (2019); IRFAN *et al.* (2019); SALMI *et al.* (2019); SARANGI *et al.* (2019); TAJBAKSH; SHAMSI (2019); XU; LIU (2019); BARROS *et al.* (2020); CAIADO *et al.* (2020); CHEN (2020); COLLA *et al.* (2020); GARDAZI *et al.* (2020); KOLOSOK *et al.* (2020); LI; DONG (2020); MILOJEVIC *et al.* (2020); QI *et al.* (2020); ROSA *et al.* (2020); SHAHBAZ *et al.* (2020); SI *et al.* (2020); SUN *et al.* (2020); WANKE *et al.* (2020); XU; CHEN (2020); YANG; ZHANG (2020); YUSIFBAYLI; NASIBOV (2020); AHMAD *et al.* (2021a); KARNITIS *et al.* (2021);

SUEYOSHI *et al.* (2021); YORO *et al.* (2021); ADAMKAITE *et al.* (2022); PONOMARENKO *et al.* (2022); RAMAIAH; JAYASANKAR (2022); DONG *et al.* (2023); MAKRIDOU *et al.* (2023);

b) estudos que avaliam a sustentabilidade, desempenho ou projetos de investimentos em energia com o uso de método multicritério à tomada de decisão: ERVURAL *et al.* (2018b); ZHOU *et al.* (2019); CHAMANDOUST *et al.* (2020); PAPAGEORGIU *et al.* (2020); PHILLIS *et al.* (2020); RIGO *et al.* (2020); SAHABUDDIN; KHAN (2021); ERSOY *et al.* (2023); PATYAL *et al.* (2023);

c) estudos que tratam sobre práticas em sustentabilidade: JARVIS; SOVACOO (2011); LI; TAO (2017); KUMAR *et al.* (2022); ARORA *et al.* (2023);

d) Ahmad *et al.* (2021b) abordam o uso de Inteligência Artificial na geração de energias solar e de hidrogênio, no controle da gestão da oferta e da demanda, e avanços recentes na tecnologia de IA; e, Wang *et al.* (2021) tratam sobre previsão de desempenho com base em redes neurais de memória de longo e curto prazos para microrredes integradas de energia em Tianjin, China.

Em consequência, foram identificadas as principais lacunas nos estudos do Eixo 3: i) avaliação integrada da sustentabilidade e do desempenho, de forma uni e multidimensional; ii) identificação de padrões de dados sobre indicadores de sustentabilidade e de desempenho, por meio de Apoio Multicritério à Decisão e inferência de regras de decisão “if ... then ...”; iii) modelo preditivo com o uso de Aprendizado de Máquina, dados históricos de indicadores, regras de decisão e dados simulados; iv) previsão de riscos das empresas; v) identificação de melhores práticas de sustentabilidade e de desempenho das empresas como consequência do *benchmarking*.

Assim, pelos resultados da revisão sistemática da literatura e *gaps* identificados nos estudos, revela-se a oportunidade de proposição de um modelo de avaliação integrada da sustentabilidade e de desempenho para as empresas concessionárias de distribuição de energia elétrica no Brasil. Este modelo visa: a) classificar as empresas sob as dimensões econômica, social, ambiental, de governança corporativa e geral, e posicioná-las em uma matriz de sustentabilidade; b) realizar *benchmarking* das empresas; c) conhecer boas práticas em sustentabilidade e desempenho, inclusive para fins de políticas públicas; d) efetuar previsão das classificações por Aprendizado de Máquina com base em dados históricos de indicadores, regras de decisão inferidas e simulações; e) realizar efetiva gestão de riscos das empresas com base em simulações e previsões, bem como, oportunidade para se atingir metas. Utiliza-se de teoria em Apoio Multicritério à Decisão,

especificamente, *Rough Sets Theory* e sua extensão, *Dominance Rough Set Approach* (DRSA), as quais permitem o tratamento matemático da imprecisão e incerteza de dados (PAWLAK, 1991; SLOWINSKI *et al.*, 2012).

Ademais, Ferretti *et al.* (2020) demonstram o uso de DRSA na definição de políticas territoriais sustentáveis, em duas províncias do norte da Itália, na Cadeia Alpina, com base em indicadores selecionados e em dimensões sociodemográficas, ambientais, econômicas e de acessibilidade. E, Błaszczynski *et al.* (2021) mostram a aplicação de DRSA e *Rough Set Balanced Rule Ensemble* para previsão de fraude financeira, especificamente, empréstimo para transações comerciais de automóveis.

#### 4. DESENVOLVIMENTO DO MODELO CONCEITUAL DE AVALIAÇÃO INTEGRADA DA SUSTENTABILIDADE E DESEMPENHO

Como produto da pesquisa, foi desenvolvido um modelo conceitual para a avaliação integrada da sustentabilidade e do desempenho para empresas concessionárias de distribuição de energia elétrica no Brasil, conforme Figura 15.

Figura 15- Modelo conceitual para avaliação integrada da sustentabilidade e desempenho



Fonte: Autor.

A seguir, as etapas são descritas de forma sucinta e visa à implementação do modelo proposto.

a) Etapa 1: Definição dos Indicadores – Consiste na coleta e tratamento das informações quantitativas e qualitativas constantes no Relatório de Responsabilidade Socioambiental e Econômico-Financeiro (RSA) de cada empresa outorgada e distribuidora de energia elétrica, instituído de acordo com o Manual de Contabilidade do Setor Elétrico (MCSE), pela Agência Nacional de Energia Elétrica (ANEEL). A ANEEL é uma agência reguladora vinculada ao Ministério de Minas e Energia, criada pela Lei nº 9.427, de 26/12/1996. A ANEEL tem como finalidade regular e fiscalizar a produção, transmissão, distribuição e comercialização de energia elétrica, de acordo com a legislação e em conformidade com as diretrizes e as políticas do governo federal. O MCSE ao incorporar o Relatório de Responsabilidade Socioambiental e Econômico-Financeiro (RSA) estabeleceu que as organizações publiquem relatórios de



sustentabilidade, visando a melhora de seus processos de gestão, e a melhora de seus relacionamentos com as partes interessadas ou públicos de interesse. A obrigatoriedade do RSA foi estabelecida a partir do exercício de 2015 para as empresas outorgadas do setor de energia elétrica, com exceção dos autoprodutores e dos produtores independentes que utilizam totalmente ou parcialmente a energia elétrica como insumo para o seu processo produtivo. O RSA está estruturado em cinco partes, denominadas dimensões (geral, governança corporativa, econômico-financeira, social e setorial, ambiental), que contemplam descrições de atividades e indicadores de desempenho (quantitativos e qualitativos), inerentes à própria dimensão, de forma a fornecer uma visão ampla, consistente e consolidada de questões relevantes e peculiares ao setor elétrico, delineadas em seu marco regulatório, e de outras gerais de responsabilidade socioambiental. O RSA deve ser encaminhado à ANEEL até o dia 30 de abril do exercício subsequente e deve conter informações referentes aos últimos três exercícios (ano base, ano-1, ano-2). Assim sendo, os indicadores de sustentabilidade e desempenho visam informar, sob a forma quantitativa ou qualitativa, os resultados obtidos frente a uma meta determinada, tornando mais inteligíveis as tendências e os fenômenos que não sejam imediatamente percebidos. Podem ser considerados medidas que resumem informações relevantes, oferecendo parâmetros de avaliação para tomadas de decisão. As finalidades dos indicadores são: i) definir prioridades estratégicas que expressem valores das responsabilidades socioambiental e econômica; ii) viabilizar um equilibrado planejamento das vertentes econômica, social e ambiental; iii) monitorar o desempenho nas dimensões da sustentabilidade; iv) identificar ações corretivas de eventuais desvios; v) acumular dados e informações de desempenho comparáveis no tempo (ANEEL, 2015, 2022b).

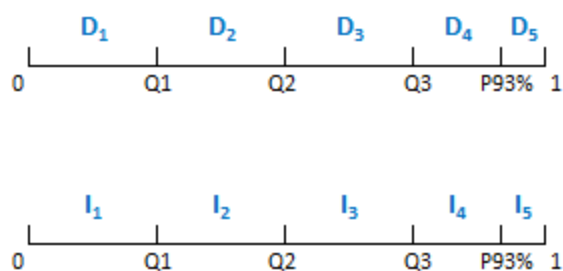
Neste contexto de pesquisa, distingue-se os indicadores diretamente relacionados à sustentabilidade, “dirSu” (sigla sugerida pelo autor da pesquisa), quando há referência ao padrão GRI (*Global Reporting Initiative*)/*Sustainability Reporting Guidelines & Electric Utility Sector Supplement* (GRI, 2000); e aqueles indiretamente relacionados à sustentabilidade, “indSu” (sigla sugerida pelo autor da pesquisa), quando não há referência ao padrão GRI.

Exemplos de indicadores direta e indiretamente relacionados à sustentabilidade: consumo total de água por fonte (m<sup>3</sup>) (sigla EN8, padrão GRI) e política de auxílio ao desenvolvimento de pequenas e microempresas fornecedoras, respectivamente. Além dos relatórios RSA, há outros relatórios com informações econômico-financeiras e de

desempenho das empresas disponibilizados diretamente no *site* da ANEEL (ANEEL, 2022a);

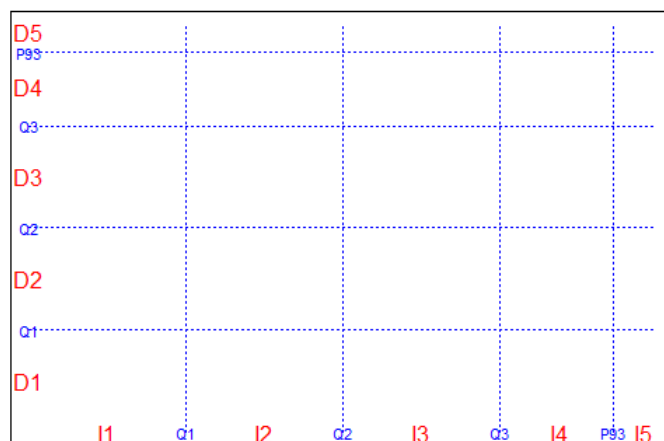
b) Etapa 2: Classificação das empresas – consiste em obter as classes por meio de índices de sustentabilidade, diretos (dirSu) ou indiretos (indSu). Para cada empresa, haverá indicadores normalizados (variação de 0 a 1) nas dimensões econômica, social, ambiental e de governança corporativa. Cada valor de indicador ( $v_i$ ) será então, normalizado ( $v_{in}$ ) e, visa obter um percentual em relação ao valor máximo encontrado ( $v_{max}$ ) (se indicador “ganho”; quanto maior, melhor) ou mínimo ( $v_{min}$ ) (se indicador “custo ou perda”; quanto menor, melhor):  $v_{in} = v_i / v_{max}$  ou  $v_{in} = 1 / (v_i / v_{min})$ , respectivamente. Assim, para cada empresa e para cada indicador, obtém-se sempre uma posição relativa (%) à empresa que pode ser considerada paradigma naquele indicador (que obteve os valores máximo ou mínimo, conforme o tipo do indicador, ganho ou custo, respectivamente). Para cada conjunto de indicadores de mesma dimensão, por exemplo, social, haverá um índice correspondente, calculado por média aritmética simples. Por fim, haverá um índice final a ser calculado pela média aritmética dos índices econômico, social, ambiental e de governança corporativa diretamente relacionados à sustentabilidade (dirSu). Procedimento análogo para a obtenção do índice indiretamente relacionado à sustentabilidade. A coleção de indicadores e classes diretamente relacionados à sustentabilidade das empresas constituir-se-á uma tabela de decisão; caso análogo para indicadores e classes indiretamente relacionados à sustentabilidade. Visando obter uma distribuição das classificações das empresas, propõe-se um modelo “quartis-percentil 93%”: os quartis Q1, Q2 (mediana) e Q3 e, o percentil de 93% (P93%), entre Q3 e 1, conforme ilustrado nas Figuras 16 e 17.

Figura 16 – Ordem crescente de preferência forte para as classes diretamente (Dn) e indiretamente (In) relacionadas à sustentabilidade



Fonte: Autor.

Figura 17 – Matriz de Sustentabilidade para o modelo quartis-percentil 93%



Fonte: Autor.

Como cada quartil identifica uma parte que representa  $\frac{1}{4}$  de uma amostra ou população, o percentil 93% representa aproximadamente um terceiro quartil da última parte:  $\frac{1}{4} \times \frac{1}{4} = 0,0625$  ou 0,7, considerando-se valor arredondado para cima; assim, o percentil 93% seria 100% - 7%. Entre as classes diretas, institui-se uma ordem crescente de preferência forte: D<sub>1</sub>, D<sub>2</sub>, D<sub>3</sub>, D<sub>4</sub>, D<sub>5</sub>. Análogo para as classes indiretas I<sub>1</sub>, I<sub>2</sub>, I<sub>3</sub>, I<sub>4</sub>, I<sub>5</sub>.

No contexto atual deste estudo, por que classificar as empresas? Inicialmente, desde a época de Aristóteles, já existia a preocupação de nomear, definir e categorizar. Com a ciência cognitiva, a categorização passou de um processo cognitivo individual para um processo cultural e social de construção da realidade, que organiza conceitos, parcialmente baseados na psicologia do pensamento. Na categorização, o reconhecimento das similaridades e diferenças leva à criação de um conhecimento novo, pelo agrupamento de entidades, de acordo com as similaridades e diferenças observadas. Conseqüentemente, categorizar coisas é inerente aos seres humanos desde os primeiros momentos de vida, porque o cérebro dá forma às estruturas que espelham o ambiente externo em uma forma categorial; trata-se de uma alternativa de estruturar a informação, pois ela procura refletir a organização da estrutura informacional de uma pessoa sobre determinado assunto (LIMA, 2010). Assim, identificar aquelas empresas com prováveis similaridades, sob uma mesma classe ou classes adjacentes, por exemplo, uso de boas práticas em sustentabilidade, permite que se obtenha um espectro maior de conhecimento sobre práticas em sustentabilidade frente à alternativa de um simples ranqueamento das empresas;

c) Etapa 3: Apoio Multicritério – uso de *Rough Sets* na verificação de (possíveis) “reduções” e “núcleo” de conhecimento nas tabelas de decisão. Por se tratar de

indicadores com base em medições quantitativas e qualitativas e, muito provavelmente, incorrer em imprecisões nessas medições, a aplicação da RST conjuntamente com o “princípio de Dominância” (*Dominance Rough Set Approach*, DRSA) permitem o tratamento matemático destas imprecisões para o universo considerado de empresas e atributos, por meio da identificação de conjuntos aproximativos inferior e superior. É possível verificar a existência de reduções e núcleo de conhecimento nas tabelas de decisão, isto é, subconjunto de indicadores com a mesma qualidade de classificação que o conjunto original (PAWLAK, 1991; COUTO; GOMES, 2010; SLOWINSKI *et al.*, 2012; COUTO; GOMES, 2020; KAROLAK *et al.*, 2020). Ademais, trata-se de um problema tipo “beta” ( $P\beta$ ) de classificação de alternativas (empresas);

d) Etapa 4: Padrões e predição – verificação de padrões sobre os indicadores por meio de regras de decisão “if ... then ...” (PAWLAK, 1991; COUTO; GOMES, 2010; SLOWINSKI *et al.*, 2012; COUTO; GOMES, 2020; KAROLAK *et al.*, 2020). Ademais, por meio de *Machine Learning*, é possível a predição das classes diretas e indiretas com base em dados históricos e em valores absolutos de indicadores informados ou simulados (HAN *et al.*, 2012; RIZA *et al.*, 2014, 2019). Os algoritmos de geração ou inferência de regras são aqueles propostos no package “RoughSets” em R (RIZA *et al.*, 2014, 2019), no *software* jMAF (*Dominance-Based Rough Set Data Analysis Framework*) (BLASZCZYNSKI *et al.*, 2013b) e no *software* Weka (WITTEN; FRANK, 2005), quais sejam: AQRules, CN2Rules, DecisionTable, HybridFS, indiscernibilityBasedRules, JRip, LEM2Rules, OneR, PART, VC-DomLEM e ZeroR (CLARK; NIBLETT, 1989; MICHALSKI *et al.*, 1991; COHEN, 1995; KOHAVI, 1995; GRZYMALA-BUSSE, 1997; WITTEN; FRANK, 2005; ALI; SMITH, 2006; JENSEN *et al.*, 2009; BLASZCZYNSKI *et al.*, 2009, 2011; RIZA *et al.*, 2014, 2019).

De acordo com Kahneman *et al.* (2021), os resultados das predições deverão ser comparados aos do cálculo das classes visando verificar possíveis divergências (erros, ruídos) e necessidade de ajustes nos algoritmos de predição. Kahneman *et al.* (2021) ressaltam que, as abordagens mecânicas tais como: Regras simples, Modelos lineares impróprios, Modelos de regressão linear e Modelos de Aprendizado de Máquina em apoio à tomada de decisão, apresentam desempenho superior frente ao julgamento humano, além de serem livres de ruído (dispersão aleatória). Ademais, torna-se desejável estabelecer uma “função de perda” com o objetivo de calcular provável custo de se classificar erroneamente uma empresa (RUSSELL, 2021).

Do Teorema de Bayes, utilizado na predição de classificação (HAN *et al.*, 2012; PROVOST; FAWCETT, 2013; ZAKI; MEIRA JR., 2020), Equação (13):

$$P(C_i | E) = P(E | C_i) P(C_i) / P(E) \quad (13),$$

$P(C_i | E)$  é a probabilidade condicional posterior de se classificar um exemplo na classe  $C_i$  dada a evidência ou vetor de atributos  $E$ ;  $P(E | C_i)$  representa a probabilidade condicional da evidência ou vetor de atributos  $E$  dado que um exemplo pertence à classe  $C_i$ ;  $P(C_i)$  é a probabilidade a priori calculada como porcentagem de todos os exemplos que são da classe  $C_i$ ;  $P(E)$  é a probabilidade calculada como a porcentagem de ocorrência da evidência  $E$  entre todos os exemplos. Para a probabilidade  $P(E | C_i)$  considera-se que, os atributos sejam condicionalmente independentes dada a classe, isto é, em  $P(e_1 \wedge e_2 \wedge \dots \wedge e_n | C_i)$ , cada  $e_j$  é independente de todos os outros  $e_k$  dada a classe  $C_i$ .

Em contraponto à Equação (13), o fator *certainty*, Equação (6), permite que se calcule diretamente a probabilidade de um exemplo pertencer a uma determinada classe dada a ocorrência de atributo(s) de condição associado(s) a uma regra de decisão específica [ $P(C_i | E)$ ], sem necessidade de se conhecer qualquer probabilidade a priori [ $P(C_i)$ ]. Como vantagem, precisão maior no cálculo da probabilidade em consequência de possíveis ajustes nas classificações, em função da aplicação do princípio de Dominância à tabela de decisão. Outra vantagem, seria o conhecimento dos atributos de condição e respectivos valores limites que determinam a classificação de um exemplo. Há ainda, o fator *coverage*, Equação (7), que permite calcular a probabilidade condicional da evidência  $E$  dado que um exemplo pertence à determinada classe [ $P(E | C_i)$ ].

Ademais, a interpretação *bayesiana* de que a probabilidade nada mais é do que uma quantificação da nossa expectativa de que um evento ocorra, sujeito à atualização cada vez que nova informação fique disponível que altere essa expectativa, torna-se uma alternativa à interpretação frequentista, de que a probabilidade de um evento incerto é em termos da frequência com que esse evento ocorre quando o experimento é repetido muitas vezes. A interpretação *bayesiana* permite quantificar a incerteza sobre algo, uma vez que está fundamentalmente relacionada à informação disponível e não à frequência de um evento (MURPHY, 2012; VIANA, 2022);

e) Etapa 5: Matriz de Sustentabilidade – elaboração de *benchmarking* em função das classificações diretas e indiretas das empresas do universo considerado. Para cada empresa, formar-se-á um par ordenado (Classe indSu, Classe dirSu) correspondente, a ser apontado na Matriz Sustentabilidade. Este posicionamento será sempre relativo, isto é, terá como referência a empresa melhor posicionada ou classificada na matriz. Este

*benchmarking* poderá ser efetuado com o uso de dados históricos de indicadores e/ou índices de sustentabilidade ou de desempenho para uma mesma empresa bem como, com relação a empresas concorrentes (não identificadas). O *benchmarking* é tradicionalmente pensado como uma ferramenta gerencial que melhora o desempenho, identificando e aplicando as melhores práticas, o qual permite que se compare o desempenho de empresas, produtos e processos externamente com concorrentes e as melhores empresas do setor, e internamente com outras operações de empresas que realizam atividades semelhantes (BOGETOFT, 2012). O *benchmarking* apresenta dois objetivos: saber se há empresa bem posicionada que possa servir de *benchmark* ou referência, bem como, conhecer o quanto uma determinada empresa está distante desse *benchmark*. Isto permite a uma empresa estabelecer metas e planos de ação, bem como, mitigar prováveis riscos associados ao seu desempenho (FERNANDEZ *et al.*, 2001; ANAND; KODALI, 2008; MORIARTY, 2011). E sob a ótica de operações, o desempenho possui os seguintes objetivos: qualidade, velocidade, confiabilidade, flexibilidade e custo (SLACK; BRANDON-JONES, 2021);

f) Etapa 6: Validação – tem como objetivo confrontar os resultados obtidos no *benchmarking* com aquelas empresas avaliadas. Assim, torna-se possível conhecer possíveis melhores práticas de sustentabilidade e desempenho adotadas por empresas situadas no setor verde da Matriz (Figura 28), de modo a viabilizar a criação de um ambiente para acesso e compartilhamento destas práticas pelas empresas (EPSTEIN; BUHOVAC, 2014; NAWAZ; KOÇ, 2019). Para o gestor público, por exemplo, ANEEL, o conhecimento destas práticas pode subsidiar a elaboração e/ou aperfeiçoamento de políticas públicas de desenvolvimento sustentável para o setor elétrico. Ao final, conforme Figura 15, reinicia-se o ciclo do modelo proposto de avaliação.

Consequentemente, as seis etapas descritas anteriormente visam de forma integrada, propor um modelo para avaliação da sustentabilidade e desempenho de empresas concessionárias de distribuição de energia elétrica, considerando-se as diversas dimensões de indicadores, de forma a se obter um índice final direto ou indiretamente relacionado à sustentabilidade e, uma correspondente classificação uni ou multidimensional.

A classificação de uma empresa apontada na Matriz de Sustentabilidade permitirá que se conheça de forma holística a sua situação frente a concorrentes, bem como, a sua evolução ao longo do tempo e, também, prever a classificação em função da simulação de indicadores e com base em dados históricos. Este é um resultado factível com o uso

de Aprendizado de Máquina e regras de decisão. Esta ferramenta de predição possibilitará às empresas uma gestão eficaz e efetiva dos seus riscos econômicos, ambientais, sociais e de governança corporativa.

O modelo conceitual desenvolvido permitiu ainda, conhecer boas práticas em sustentabilidade e desempenho daquelas empresas mais bem posicionadas e, possível externalização de conhecimento dessas práticas. Ademais, aos gestores públicos será possível obter subsídios à elaboração e aperfeiçoamento de políticas públicas direcionadas ao setor elétrico.

## **5. APLICAÇÃO DO MODELO CONCEITUAL DE AVALIAÇÃO INTEGRADA DA SUSTENTABILIDADE E DESEMPENHO**

Para a presente pesquisa, o modelo conceitual de avaliação integrada da sustentabilidade e desempenho foi aplicado em etapas, conforme estabelecido na Figura 15.

### **5.1. DEFINIÇÃO DOS INDICADORES**

Com base nos indicadores definidos pela ANEEL (ANEEL, 2014, 2015, 2021, 2022a, 2022b), procedeu-se à captura das informações constantes nos RSAs de 41 (quarenta e uma) empresas concessionárias distribuidoras de energia elétrica, no período (ano-base) de 2013 a 2021, de um total de 53 (cinquenta e três), conforme Quadro 19. A redução no número de empresas é devida a informações insuficientes para compor o quantitativo de indicadores pesquisados. Até a data de 10/07/2023, ainda não se encontravam disponíveis todos os RSAs dessas empresas relativos ao ano-base de 2022.

Ressalta-se que, o universo de empresas concessionárias de distribuição de energia elétrica considerado foi aquele existente em 2021, considerando-se as transformações, fusões, cisões ou incorporações de empresas porventura ocorridas no período 2013-2021.



Quadro 19– Empresas concessionárias de distribuição de energia elétrica consideradas na pesquisa, em 2021

Nº	Empresa	Razão Social	UF	Grupo Econômico
1	AmE	AmE - Amazonas Distribuidora de Energia S/A	AM	Oliveira Energia
2	Boa Vista	Boa Vista Energia S/A	RR	Oliveira Energia
3	CEA	Companhia de Eletricidade do Amapá	AP	Equatorial
4	CEAL	Companhia Energética de Alagoas	AL	Equatorial
5	Ceb-Dis	CEB Distribuição S/A	DF	Neoenergia
6	CEEE-D	Companhia Estadual de Distribuição de Energia Elétrica	RS	Equatorial
7	Celesc-Dis	Celesc Distribuição S/A	SC	Estadual
8	Celpa	Centrais Elétricas do Pará S/A	PA	Equatorial
9	Celpe	Companhia Energética de Pernambuco	PE	Neoenergia
10	Cemar	Companhia Energética do Maranhão	MA	Equatorial
11	Cemig-D	CEMIG Distribuição S/A	MG	Cemig
12	Cepisa	Companhia Energética do Piauí	PI	Equatorial
13	Ceron	Centrais Elétricas de Rondônia S/A	RO	Energisa
14	Chesp	Companhia Hidroelétrica São Patrício	GO	Não Aplicável
15	Cocel	Companhia Campolarguense de Energia	PR	Não Aplicável
16	Coelba	Companhia de Eletricidade do Estado da Bahia	BA	Neoenergia
17	Copel-Dis	Copel Distribuição S/A	PR	Estadual
18	Cosern	Companhia Energética do Rio Grande do Norte	RN	Neoenergia
19	EBO	Energisa Borborema – Distribuidora de Energia S/A	PB	Energisa
20	EDP ES	Espírito Santo Distribuição de Energia S/A	ES	EDP
21	EDP SP	São Paulo Distribuição de Energia S/A	SP	EDP
22	Elektro	Elektro Redes S/A	SP/MS	Neoenergia
23	Eletroacre	Companhia de Eletricidade do Acre	AC	Energisa
24	EMG	Energisa Minas Gerais - Distribuidora de Energia S/A	MG	Energisa
25	EMS	Energisa Mato Grosso do Sul - Distribuidora de Energia S/A	MS	Energisa
26	EMT	Energisa Mato Grosso - Distribuidora de Energia S/A	MT	Energisa
27	Enel CE	Enel Distribuição Ceará S/A	CE	Enel
28	Enel GO	Enel Distribuição Goiás S/A	GO	Enel
29	Enel RJ	Enel Distribuição Rio S/A	RJ	Enel
30	Enel SP	Enel Distribuição São Paulo S/A	SP	Enel
31	ENF	Energisa Nova Friburgo - Distribuidora de Energia S/A	RJ	Energisa
32	EPB	Energisa Paraíba - Distribuidora de Energia	PB	Energisa
33	ESE	Energisa Sergipe - Distribuidora de Energia S/A	SE	Energisa
34	ESS'	Energisa Sul Sudeste - Distribuidora de Energia S/A	SP	Energisa
35	ETO	Energisa Celtins - Distribuidora de Energia S/A	TO	Energisa
36	Iguaçu	Iguaçu Distribuidora de Energia Elétrica Ltda	SC	Não Aplicável
37	Light SESA	Light Serviços de Eletricidade S/A	RJ	Não Aplicável
38	Muxenergia	Muxfeldt Marin & Cia. Ltda	RS	Não Aplicável
39	Santa Maria	Empresa Luz e Força Santa Maria S/A	ES	Não Aplicável
40	Sulgipe	Companhia Sul Sergipana de Eletricidade	SE	Não Aplicável
41	Urussanga	Empresa Força e Luz Urussanga Ltda	SC	Não Aplicável

Fonte: Autor; adaptado de ANEEL (2021, 2022a).

Visando preservar a identificação das empresas, estabeleceu-se uma codificação para cada empresa, exceto para o grupo econômico ENEL, constituído de quatro

empresas, quando foi estabelecido um único código. Assim, as empresas foram designadas sequencialmente de E01 a E38.

Ainda que, o conjunto de indicadores estabelecidos seja expressivo, mais de duas centenas, nenhum indicador pesquisado se apresentava de forma completa para todas as empresas pesquisadas. Assim, foram considerados somente aqueles indicadores com o mínimo de 40% (quarenta por cento) de preenchimento ou resposta, para 38 (trinta e oito) empresas e no período 2013-2021, os quais resultaram em um total de 62 (sessenta e dois) indicadores: 33 (trinta e três) indicadores diretamente (dirSU) e 29 (vinte e nove) indicadores indiretamente (indSU) associados à sustentabilidade, incluindo os indicadores de governança corporativa, conforme Quadros 20 a 23. Os indicadores estão classificados segundo a dimensão: E, econômica; S, social; A, ambiental; GC, governança corporativa. Para cada indicador, encontra-se a sua natureza: G, ganho; C, custo; N, neutro.

Assim, como exemplo, o cálculo do índice econômico (IE) diretamente associado à sustentabilidade e desempenho (dirSU), compreende a média aritmética simples dos indicadores normalizados IE01 a IE10, conforme Quadro 20.

Quadro 20– Indicadores ESA diretamente associados à sustentabilidade e desempenho (dirSU)

<b>Código</b>	<b>Indicador</b>	<b>Dimensão</b>	<b>Unidade</b>	<b>G-C-N</b>	<b>Referência(s)</b>
IE01	Consumidores atendidos	E	un	G	ANEEL (2015); GRI (2000)
IE02	Energia comprada	E	GWh	C	ANEEL (2015); GRI (2000)
IE03	Perdas elétricas globais	E	GWh	C	ANEEL (2015); GRI (2000)
IE04	Subestações distribuidoras	E	un	G	ANEEL (2015); GRI (2000)
IE05	Capacidade instalada	E	MVA	G	ANEEL (2015); GRI (2000)
IE06	Folha de pagamento bruta	E	R\$ mil	C	ANEEL (2015); GRI (2000)
IE07	Energia vendida, setor Industrial	E	GWh	G	ANEEL (2015); GRI (2000)
IE08	Energia vendida, setor Rural	E	GWh	G	ANEEL (2015); GRI (2000)
IE09	Energia vendida, setor Iluminação pública	E	GWh	G	ANEEL (2015); GRI (2000)
IE10	Energia vendida, setor Poder público	E	GWh	G	ANEEL (2015); GRI (2000)
IS01	Número de empregados próprios	S	un	G	ANEEL (2015); GRI (2000)
IS02	Empregados com idade entre 31 e 40 anos	S	%	G	ANEEL (2015); GRI (2000)
IS03	Empregados com idade entre 41 e 50 anos	S	%	G	ANEEL (2015); GRI (2000)
IS04	Mulheres em relação ao total de empregados	S	%	G	ANEEL (2015); GRI (2000)
IS05	Empregados do programa de contratação de aprendizes	S	%	G	ANEEL (2015); GRI (2000)
IS06	Benefícios com Educação	S	R\$ mil	G	ANEEL (2015); GRI (2000)
IS07	Benefícios com Alimentação	S	R\$ mil	G	ANEEL (2015); GRI (2000)
IS08	Benefícios com Saúde	S	R\$ mil	G	ANEEL (2015); GRI (2000)
IS09	Benefícios com Creches ou auxílio-creches	S	R\$ mil	G	ANEEL (2015); GRI (2000)
IS10	Taxa de rotatividade (demissões)	S	%	C	ANEEL (2015); GRI (2000)
IA01	Consumo de energia direta, fonte Diesel	A	GJ	C	ANEEL (2015); GRI (2000)
IA02	Consumo de energia direta, fonte Gasolina	A	GJ	C	ANEEL (2015); GRI (2000)
IA03	Consumo total de água, fonte Rede pública	A	m <sup>3</sup>	C	ANEEL (2015); GRI (2000)

Fonte: Autor.

Quadro 21– Indicadores GC diretamente associados à sustentabilidade e desempenho (dirSU)

Código	Indicador	Dimensão	Unidade	G-C-N	Referência(s)
IG01	Estrutura de governança da organização, incluindo Comitês sob o mais alto órgão de governança responsável por tarefas específicas, como configuração da estratégia ou supervisão organizacional. Descreva o mandato e a composição (incluindo número de membros independentes e /ou membros não executivos) de tais comitês e indicar qualquer responsabilidade direta por questões econômicas, sociais e desempenho ambiental	GC	un	G	GRI (2000)
IG02	Indique se o presidente do mais alto órgão de governança também é um diretor executivo; em caso afirmativo, sua função dentro da organização e as razões para tal	GC	un	G	GRI (2000)
IG03	Para organizações que têm uma estrutura de conselho unitário, indique o número de membros do mais alto órgão de governança que são independentes e/ou membros não executivos. Declare como a organização define independente e não executivo. Este elemento se aplica apenas a organizações que têm estruturas de conselho unitárias	GC	un	G	GRI (2000)
IG04	Mecanismos para que acionistas e funcionários forneçam recomendações ou orientações para o mais alto órgão de governança. Incluir referência a processos relacionados a: <ul style="list-style-type: none"> <li>• O uso de resoluções de acionistas ou outros mecanismos para permitir à minoria de acionistas para expressar opiniões ao mais alto órgão de governança; e</li> <li>• Informar e consultar os funcionários sobre as relações de trabalho com órgãos de representação, como organização, nível de conselhos de trabalho e representação de funcionários do mais alto órgão de governança.</li> </ul> Identifique tópicos relacionados a aspectos econômicos, ambientais, e o desempenho social gerado por meio desses mecanismos durante o período coberto pelo relatório	GC	un	G	GRI (2000)

Fonte: Autor.

Quadro 21 (cont.) – Indicadores GC diretamente associados à sustentabilidade e desempenho (dirSU)

Código	Indicador	Dimensão	Unidade	G-C-N	Referência(s)
IG05	Relação entre a remuneração para membros do mais alto órgão de governança, gerentes seniores e executivos, e o desempenho da organização (incluindo atuações social e ambiental)	GC	un	G	GRI (2000)
IG06	Processos em vigor para a mais alta governança para garantir que conflitos de interesse sejam evitados	GC	un	G	GRI (2000)
IG07	Processo para determinar as qualificações e experiência dos membros do mais alto escalão de governança para orientar a organização quanto à estratégia econômica, ambiental e social	GC	un	G	GRI (2000)
IG08	Declarações de missão desenvolvidas internamente ou valores, códigos de conduta e princípios relevantes para desempenho econômico, ambiental e social e o status de sua implementação. Explique o grau em que estes: <ul style="list-style-type: none"> <li>• São aplicados em toda a organização em diferentes regiões e departamentos/unidades;</li> <li>• Estão relacionados com os padrões internacionalmente aceitos</li> </ul>	GC	un	G	GRI (2000)
IG09	Procedimentos do mais alto órgão de governança para supervisionar a organização: gestão econômica, ambiental e desempenho social, incluindo riscos relevantes e oportunidades e aderência ou conformidade com padrões internacionalmente acordados, códigos de conduta e princípios. Inclui a frequência com que o mais alto órgão de governança avalia o desempenho da sustentabilidade	GC	un	G	GRI (2000)
IG10	Processos para avaliar a mais alta governança, particularmente no que diz respeito a desempenho econômico, ambiental e social	GC	un	G	GRI (2000)

Fonte: Autor.

Quadro 22– Indicadores ESA indiretamente associados à sustentabilidade e desempenho (indSU)

<b>Código</b>	<b>Indicador</b>	<b>Dimensão</b>	<b>Unidade</b>	<b>G-C-N</b>	<b>Referência(s)</b>
IE01	Localidades atendidas (municípios)	E	un	G	ANEEL (2015)
IE02	Escritórios comerciais	E	un	G	ANEEL (2015)
IE03	Transformadores de distribuição	E	un	G	ANEEL (2015)
IE04	Energia vendida por empregado	E	MWh	G	ANEEL (2015)
IE05	Consumidores por empregado	E	un	G	ANEEL (2015)
IS01	Média de horas extras por empregado/ano	S	h	C	ANEEL (2015)
IS02	Chamadas Recebidas (atendimento ao cliente)	S	un	C	ANEEL (2015)
IS03	Número Médio de Atendentes	S	un	G	ANEEL (2015)
IS04	IAb - Índice de Abandono (atendimento ao cliente)	S	%	C	ANEEL (2015)
IS05	ICO - Índice de Chamadas Ocupadas (atendimento ao cliente)	S	%	C	ANEEL (2015)
IS06	TMA - Tempo Médio de Atendimento	S	s	C	ANEEL (2015)
IS07	DER (Duração Equivalente de Reclamação) (1)	S	h	C	ANEEL (2010; 2015)
IS08	FER (Frequência Equivalente de Reclamação a cada mil Unidades Consumidoras) (2)	S	un	C	ANEEL (2010; 2015)
IS09	Atendimentos realizados	S	un	G	ANEEL (2015)
IS10	Atendimentos realizados fora do prazo (violação de prazos de serviços comerciais)	S	un	C	ANEEL (2015)
IA01	Consumo de água por empregado	A	m <sup>3</sup>	C	ANEEL (2015)
IA02	Alunos atendidos	A	un	G	ANEEL (2015)

Fonte: Autor.

Nota:

(1)  $DER = (RP \times PMS) / RP$

(2)  $FER = (RP \times 1000) / NC$

RP = Quantidade de reclamações por tipo procedentes dos consumidores solucionadas pela distribuidora no período de apuração;

PMS = Prazo Médio de Solução das reclamações por tipo procedentes no período de apuração, expresso em horas e centésimos de horas;

NC = Número de consumidores da distribuidora, no final do período de apuração, coletado pelo Sistema de Acompanhamento de Informações de Mercado para Regulação Econômica - SAMP ou outro que vier a substituí-lo.

Quadro 23– Indicadores GC indiretamente associados à sustentabilidade e desempenho (indSU)

<b>Código</b>	<b>Indicador</b>	<b>Dimensão</b>	<b>Unidade</b>	<b>G-C-N</b>	<b>Referência(s)</b>
IG01	Emissão de ações preferenciais	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG02	Acionistas preferenciais têm direito a voto em matérias relevantes	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG03	Mecanismos de divulgação sobre os temas deliberados nas assembléias	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG04	Processos administrativos, arbitrais ou judiciais contra a companhia, os administradores ou o controlador, envolvendo tratamento não equitativo de acionistas minoritários nos últimos cinco anos	GC	un	C	PEREIRA; CÂNDIDO (2020)
IG05	Proibição de empréstimos e garantias em favor do controlador, dos administradores e de outras partes relacionadas	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG06	Existência de um canal para comunicações anônimas que se destina a receber denúncias, dúvidas e sugestões	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG07	Relatório anual e/ou de sustentabilidade com acessibilidade para portadores de necessidades especiais	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG08	Programas para educação sobre sustentabilidade e os públicos atingidos	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG09	Aderência de compromissos voluntários relacionados ao desenvolvimento sustentável	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG10	Existência de comitê de sustentabilidade	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG11	Parecer de auditoria independente	GC	un	G	PEREIRA; CÂNDIDO (2020)
IG12	Compromisso com o combate à corrupção que abranja o público interno	GC	un	G	PEREIRA; CÂNDIDO (2020)

Fonte: Autor.

Para a captura de dados, foi utilizado o VBA (*Visual Basic Application*, Microsoft Excel) na programação de procedimentos para a extração de dados a partir de dicionários de termos indexados (níveis 1 e 2; colunas C e D) com base nas descrições dos indicadores, conforme Figuras 18 e 19. Inicialmente, foram extraídos os textos dos RSAs em formato PDF das empresas e transportados para planilhas, individualmente, com o uso do *software* PDF Candy Desktop PRO v 2.87. Como geralmente, cada RSA refere-se a um triênio, foram necessários três ou mais RSAs para cada empresa para abranger o período 2013-2021. Em média, foram importadas para processamento 6.000 (seis mil) linhas de texto para cada triênio, no total de 6.000 linhas x 3 triênios x 38 empresas ou 684.000 linhas de texto.



Figura 18– Indicadores de sustentabilidade e desempenho, e dicionário de termos de indexação

	A	B	C	D
1	eficiência	Indicadores operacionais e de produtividade	nível 1	nível 2
40				
41		Indicadores sociais externos		
42	3.5.2	Chamadas Recebidas (unid.)	chamadas recebidas	
43	3.5.2	Número Médio de Atendentes (unid.)	atendentes	
44	3.5.2	INS Índice de Nivel de Serviço (%)	nível de serviço	
45	3.5.2	IAb - Índice de Abandono (%)	índice de abandono	
46	3.5.2	ICO - Índice de Chamadas Ocupadas (%)	chamadas ocupadas	
47	3.5.2	TMA - Tempo Médio de Atendimento (s)	tempo médio de atendimento	
48	3.5.2	Indenização por Danos Elétricos: Volume de Solicitações (unid.)	danos elétricos	
49	3.5.2	Indenização por Danos Elétricos: Procedentes (unid.)	danos elétricos	
50	3.5.2	Reclamações Procedentes (unid.)	reclamações	
51	3.5.2	DER (horas)	DER (horas)	
52	3.5.2	FER (unid.)	FER (unid.)	
53	3.5.2	Atendimentos realizados (unid.)	atendimentos realizados	
54	3.5.2	Atendimentos realizados fora do prazo (unid.)	atendimentos realizados	
55	3.5.2	Eficiência do Atendimento (%)	eficiência do atendimento	
56	3.5.2	Número de reclamações de consumidores encaminhadas	reclamações de consumidores	
57	3.5.2	Número de reclamações de consumidores encaminhadas: À Empresa	reclamações de consumidores	
58	3.5.2	Número de reclamações de consumidores encaminhadas: À ANEEL – agências estaduais / regionais	reclamações de consumidores	
59	3.5.2	Número de reclamações de consumidores encaminhadas: Ao PROCON	reclamações de consumidores	
60	3.5.2	Número de reclamações de consumidores encaminhadas: À Justiça	reclamações de consumidores	

Fonte: Autor; adaptado de Baeza-Yates; Ribeiro-Neto (2013).

Figura 19– Indicadores de governança corporativa e dicionário de termos de indexação

	A	B	C	D	E	F	G	H
1	#	<b>Indicadores de governança corporativa</b>	<b>nível 1</b>	<b>nível 2</b>	ad	ad	<b>2021-2019</b>	<b>2018-2016</b>
2	1	1 Emissão de ações preferenciais.	ações preferenciais					
3	2	2 Acionistas preferenciais têm direito a voto em matérias relevantes.	acionistas preferenciais	direito a voto				
4	3	3 Mecanismos de divulgação sobre os temas deliberados nas assembleias.	mecanismo de divulgação					
5	4	4 Processos administrativos, arbitrais ou judiciais contra a companhia, os administradores ou o controlador, envolvendo tratamento não equitativo de acionistas minoritários, nos últimos cinco anos.	processo administrativo	contra				
6	5	4 Processos administrativos, arbitrais ou judiciais contra a companhia, os administradores ou o controlador, envolvendo tratamento não equitativo de acionistas minoritários, nos últimos cinco anos.	processo arbitral	contra				
7	6	4 Processos administrativos, arbitrais ou judiciais contra a companhia, os administradores ou o controlador, envolvendo tratamento não equitativo de acionistas minoritários, nos últimos cinco anos.	processo judicial	contra				
8	7	5 Proibição de empréstimos e garantias em favor do controlador, dos administradores e de outras partes relacionadas.	proibição de empréstimo					
9	8	5 Proibição de empréstimos e garantias em favor do controlador, dos administradores e de outras partes relacionadas.	proibição de garantia					
10	9	6 Existência de um canal para comunicações anônimas que se destina a receber denúncias, dúvidas e sugestões.	canal de denúncias					869
11	10	6 Existência de um canal para comunicações anônimas que se destina a receber denúncias, dúvidas e sugestões.	canal de dúvidas					

Fonte: Autor; adaptado de Baeza-Yates; Ribeiro-Neto (2013).

Após a importação do texto de cada RSA, procedeu-se à conferência dos conteúdos dos indicadores transportados. Para os indicadores de governança corporativa, os conteúdos transportados (trechos de textos) foram confrontados visualmente com as descrições dos indicadores, objetivando verificar a existência (“1”, independentemente de repetições) ou não (“0”) de cada indicador (BAEZA-YATES; RIBEIRO-NETO, 2013). Para os demais indicadores de conteúdo numérico, prosseguiu-se ao tratamento de dados conforme as etapas descritas a seguir (HAIR *et al.*, 2009; HAN *et al.*, 2012; ILYAS; CHU, 2019):

a) Para cada indicador e para cada empresa, calculou-se a relação entre os valores máximo e mínimo visando verificar a variação máxima no período 2013-2021 e, a ocorrência de possíveis valores inconsistentes. Essa inconsistência pode ser devida a valores capturados erroneamente ou valores com unidades de medida diferentes da unidade de medida do indicador;

b) Apontamento dos valores máximo e mínimo para cada indicador, para todas as empresas e no período 2013-2021, visando identificar possíveis valores *outliers*, com base no intervalo interquartil,  $IQ = Q3 - Q1$ : acima de  $Q3 + 1,5 \times IQ$  ou abaixo de  $Q1 - 1,5 \times IQ$ . Os valores *outliers* encontrados foram então verificados quanto à exatidão.

Com relação a valores ausentes dos indicadores, para cada empresa e considerando-se o mesmo período de tempo, esses foram preenchidos pela mediana dos valores presentes visando a não influência de valores extremos, o que ocorreria se fosse utilizada a média aritmética simples;

c) No caso de valores ausentes para todo o período de tempo, para determinada empresa e indicador, esses foram preenchidos pela mediana dos valores informados pelas demais empresas, para cada ano. Neste caso, para evitar a duplicação de valores para duas empresas, por exemplo, utilizou-se da mediana com intervalo aleatório de  $\pm 5\%$ . Evitou-se também em caso de indicador “custo”, a geração do valor “zero”, o qual foi substituído por “0,01”.

A Figura 20 ilustra os resultados do preenchimento dos valores ausentes pelas medianas horizontal (células em azul) e vertical (células em verde).

Figura 20– Preenchimento de valores ausentes pelas medianas horizontal (em azul) e vertical (em verde)

2021	2020	2019	2018	2017	2016	2015	2014	2013
84	84	86	84	84	84	84	84	84
1	1	1	15	15	1	1	1	1
259	259	258	260	261	261	252	205	259
777	777	777	777	777	777	777	777	777
9	9	9	9	8	8	8	9	9
43	61	60	51	52	63	52	61	57
43	61	62	54	53	65	53	59	60
69	69	69	30	30	69	69	69	69
30	30	30	30	30	30	30	30	30
24	24	26	24	24	24	24	25	24
6	6	6	6	6	6	6	6	6
79	78	78	78	79	79	79	79	79
144	143	142	142	144	144	144	145	144
66	66	66	66	66	66	66	66	66
1	1	1	1	1	1	1	1	1
217	217	217	217	217	217	218	218	217
56	56	55	59	56	56	56	56	56
63	63	63	63	63	63	63	63	63
83	83	83	83	83	83	82	83	83
139	140	140	140	139	64	11	63	139
4	3.5	3	3.5	3.5	3.5	3.5	3.5	3.5

Fonte: Autor.

## 5.2. CLASSIFICAÇÃO DAS EMPRESAS

Uma vez que os indicadores tenham sido extraídos e tratados, procedeu-se à sua agregação para todas as empresas, anualmente, no período 2013-2021. Este procedimento foi realizado tanto para os indicadores diretamente (dirSU) como indiretamente (indSU) relacionados à sustentabilidade. Estabeleceu-se então, duas agregações de indicadores: uma com valores absolutos e outra, com valores normalizados. A partir dos valores normalizados, foram calculados os índices por dimensão econômica, social, ambiental, governança corporativa e geral, bem como, os quartis e o percentil 93% para cada dimensão. Como consequência, foram obtidas as classificações nas dimensões anteriormente citadas para cada empresa, anualmente, para os eixos dirSU (classes D1, D2, D3, D4, D5) e indSU (classes I1, I2, I3, I4, I5).

As Figuras 21 a 23 ilustram os passos anteriormente citados para as dez primeiras empresas, no período 2013-2021. Na Figura 22, parte 2, os valores em azul representam os índices finais por dimensão (IE, econômica; IS, social; IA, ambiental; IG, governança corporativa) e geral (Dir).

Figura 21– Indicadores dirSU, valores absolutos, ano 2021, dez primeiras empresas – parte 1

	IE01	IE02	IE03	IE04	IE05	IE06	IE07	IE08	IE09	IE10
	G	C	C	G	G	C	G	G	G	G
<b>E01</b>	953647	14879	4882	28	2729.4	147003	448	330	188	435
<b>E02</b>	136949.5	961.415	143.84	17	212.8	45115	22714	60.76	34389	85140
<b>E03</b>	3228811	19070.82	2272.29	166	3829	654010	2563.495	1357.725	620.3	434.305
<b>E04</b>	8882293	29482.6	4791	416	10980	846747	1695	3944	1226	729
<b>E05</b>	38835	140.07	12.2	9	96.5	6185	9.98	22.93	5	14.27
<b>E06</b>	22104	225	26	55469	1650	14998	368	1760	1	319
<b>E07</b>	4519630	25337	3071	378	10742.3	511280.5	71651	2562	1001	556
<b>E08</b>	1511594	8960.64	1599.15	14	3761	352825.2	14.51	13.02	5.82	4.39
<b>E09</b>	1804425	11541	1421.46	14	4010	352825.2	12402	7962	3420	9938
<b>E10</b>	263115	1328.4	276.7	13	312	41302	38.25	48.67	48.64	125.23

	IS01	IS02	IS03	IS04	IS05	IS06	IS07	IS08	IS09	IS10	IA01	IA02	IA03
	G	G	G	G	G	G	G	G	G	C	C	C	C
<b>E01</b>	1121	24.35	25	30.24	2.14	118	25.06	20165664	443	9	93059	10959	15200
<b>E02</b>	350	38.735	26.8	23.1	4.28	120	2877	6486	433	10	2953	1583.61	8063
<b>E03</b>	3747	25.45	24.71	18.65	4.875	591.5	34750.5	47891	1433.5	10.3	6129	22143	59045
<b>E04</b>	3761	41	23	12.98	0.01	609	57672	43893	1377	7.4	2247475	420356.2	118465
<b>E05</b>	117	27.35	34.19	20.51	4	36	1060	385	445	9	0.165	0.106	1078
<b>E06</b>	116	40	22	26	4	60	1673	945	441	8	99197	10779	14599
<b>E07</b>	4430	37.56	28.65	21.51	3.83	1344	83970	134231	0.01	2.945	135047	980.79	77927.5
<b>E08</b>	956	55.86	23	23	4.08	91.95	13336120	9505.68	352180	9	113278.7	10910.85	9099118
<b>E09</b>	1214	66.23	25	23	1.4	182.49	42980.91	37458.34	1859.84	9	113278.7	10910.85	9099118
<b>E10</b>	298	42	20.9	19	6	307	4324	1494.5	226	0.505	1161.072	336.4712	5446.5

Fonte: Autor.

Figura 21 – Indicadores dirSU, valores absolutos, ano 2021, dez primeiras empresas – parte 2

	<b>IG01</b>	<b>IG02</b>	<b>IG03</b>	<b>IG04</b>	<b>IG05</b>	<b>IG06</b>	<b>IG07</b>	<b>IG08</b>	<b>IG09</b>	<b>IG10</b>
	G	G	G	G	G	G	G	G	G	G
<b>E01</b>	1	0	0	0	0	1	0	1	1	0
<b>E02</b>	1	0	0	0	0	1	0	1	0	0
<b>E03</b>	0	0	0	0	0	0	0	1	0	0
<b>E04</b>	0	0	0	0	0	1	0	1	0	0
<b>E05</b>	0	0	0	0	1	0	0	0	0	0
<b>E06</b>	0	0	0	0	0	0	0	0	0	0
<b>E07</b>	1	0	0	0	0	0	0	1	1	0
<b>E08</b>	1	0	0	0	0	0	0	0	0	0
<b>E09</b>	1	0	0	0	0	0	0	0	0	0
<b>E10</b>	1	0	0	0	0	1	0	0	0	0

Fonte: Autor.

Figura 22– Indicadores dirSU, valores normalizados, ano 2021, dez primeiras empresas – parte 1

	<b>IE01</b>	<b>IE02</b>	<b>IE03</b>	<b>IE04</b>	<b>IE05</b>	<b>IE06</b>	<b>IE07</b>	<b>IE08</b>	<b>IE09</b>	<b>IE10</b>
<b>E01</b>	0.0542	0.0022	0.0009	0.0005	0.0871	0.0382	0.0016	0.0414	0.0052	0.0044
<b>E02</b>	0.0078	0.0342	0.0312	0.0003	0.0068	0.1246	0.0805	0.0076	0.9515	0.8565
<b>E03</b>	0.1837	0.0017	0.0020	0.0030	0.1221	0.0086	0.0091	0.1705	0.0172	0.0044
<b>E04</b>	0.5052	0.0011	0.0009	0.0075	0.3503	0.0066	0.0060	0.4954	0.0339	0.0073
<b>E05</b>	0.0022	0.2349	0.3680	0.0002	0.0031	0.9086	0.0000	0.0029	0.0001	0.0001
<b>E06</b>	0.0013	0.1462	0.1727	1.0000	0.0526	0.3747	0.0013	0.2210	0.0000	0.0032
<b>E07</b>	0.2571	0.0013	0.0015	0.0068	0.3427	0.0110	0.2540	0.3218	0.0277	0.0056
<b>E08</b>	0.0860	0.0037	0.0028	0.0003	0.1200	0.0159	0.0001	0.0016	0.0002	0.0000
<b>E09</b>	0.1026	0.0029	0.0032	0.0003	0.1279	0.0159	0.0440	1.0000	0.0946	0.1000
<b>E10</b>	0.0150	0.0248	0.0162	0.0002	0.0100	0.1361	0.0001	0.0061	0.0013	0.0013

	<b>IS01</b>	<b>IS02</b>	<b>IS03</b>	<b>IS04</b>	<b>IS05</b>	<b>IS06</b>	<b>IS07</b>	<b>IS08</b>	<b>IS09</b>	<b>IS10</b>	<b>IA01</b>	<b>IA02</b>	<b>IA03</b>
<b>E01</b>	0.1264	0.3677	0.4545	0.9147	0.2779	0.0878	0.0000	1.0000	0.0013	0.0331	0.0000	0.0000	0.0709
<b>E02</b>	0.0395	0.5849	0.4873	0.6987	0.5558	0.0893	0.0002	0.0003	0.0012	0.0298	0.0001	0.0001	0.1337
<b>E03</b>	0.4224	0.3843	0.4493	0.5641	0.6331	0.4401	0.0026	0.0024	0.0041	0.0289	0.0000	0.0000	0.0183
<b>E04</b>	0.4240	0.6191	0.4182	0.3926	0.0013	0.4531	0.0043	0.0022	0.0039	0.0402	0.0000	0.0000	0.0091
<b>E05</b>	0.0132	0.4130	0.6216	0.6204	0.5195	0.0268	0.0001	0.0000	0.0013	0.0331	1.0000	1.0000	1.0000
<b>E06</b>	0.0131	0.6040	0.4000	0.7864	0.5195	0.0446	0.0001	0.0000	0.0013	0.0372	0.0000	0.0000	0.0738
<b>E07</b>	0.4994	0.5671	0.5209	0.6506	0.4974	1.0000	0.0063	0.0067	0.0000	0.1010	0.0000	0.0001	0.0138
<b>E08</b>	0.1078	0.8434	0.4182	0.6957	0.5299	0.0684	1.0000	0.0005	1.0000	0.0331	0.0000	0.0000	0.0001
<b>E09</b>	0.1369	1.0000	0.4545	0.6957	0.1818	0.1358	0.0032	0.0019	0.0053	0.0331	0.0000	0.0000	0.0001
<b>E10</b>	0.0336	0.6342	0.3800	0.5747	0.7792	0.2284	0.0003	0.0001	0.0006	0.5891	0.0001	0.0003	0.1979

Fonte: Autor.



Figura 22 – Indicadores dirSU, valores normalizados e índices finais (em azul), ano 2021, dez primeiras empresas – parte 2

	IG01	IG02	IG03	IG04	IG05	IG06	IG07	IG08	IG09	IG10	IE	IS	IA	IG	Dir	CLASSE
<b>E01</b>	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	1.0000	1.0000	0.0000	0.0236	0.3263	0.0236	0.4000	0.1934	D4
<b>E02</b>	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	1.0000	0.0000	0.0000	0.2101	0.2487	0.0446	0.3000	0.2009	D4
<b>E03</b>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0522	0.2931	0.0061	0.1000	0.1129	D1
<b>E04</b>	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	1.0000	0.0000	0.0000	0.1414	0.2359	0.0030	0.2000	0.1451	D2
<b>E05</b>	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1520	0.2249	1.0000	0.1000	0.3692	D5
<b>E06</b>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1973	0.2406	0.0246	0.0000	0.1156	D1
<b>E07</b>	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	1.0000	0.0000	0.1229	0.3849	0.0046	0.3000	0.2031	D5
<b>E08</b>	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0231	0.4697	0.0000	0.1000	0.1482	D3
<b>E09</b>	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1491	0.2648	0.0000	0.1000	0.1285	D2
<b>E10</b>	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0211	0.3220	0.0661	0.2000	0.1523	D3

Fonte: Autor.

Figura 23– Resultado da classificação para as dez primeiras empresas, período 2021-2013

<i>Código</i>	2021		2020		2019		2018		2017		2016		2015		2014		2013	
	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>	<i>Dir</i>	<i>Ind</i>
E01	D4	I1	D4	I1	D4	I1	D2	I1	D2	I1	D1	I1	D2	I1	D2	I1	D1	I1
E02	D4	I1	D4	I1	D4	I1	D3	I1	D3	I1	D3	I1	D3	I1	D3	I1	D3	I1
E03	D1	I2	D2	I3	D1	I2	D4	I4	D4	I2	D4	I2	D3	I3	D3	I2	D2	I3
E04	D2	I4	D3	I4	D2	I4	D4	I5	D4	I5	D4	I5	D4	I4	D4	I4	D4	I4
E05	D5	I1	D5	I1	D5	I1	D5	I1	D5	I1	D5	I1	D5	I1	D5	I1	D5	I1
E06	D1	I1	D1	I1	D1	I1	D2	I1	D2	I2	D2	I2	D2	I1	D2	I1	D2	I1
E07	D5	I5	D5	I5	D4	I5	D5	I5	D5	I4	D4	I5	D4	I5	D4	I5	D4	I5
E08	D3	I2	D3	I2	D3	I2	D3	I3	D3	I3	D3	I3	D3	I2	D3	I2	D3	I2
E09	D2	I2	D2	I2	D2	I3	D2	I4	D2	I3	D2	I4	D2	I3	D2	I3	D2	I3
E10	D3	I2	D1	I2	D1	I2	D4	I2	D4	I2	D4	I3	D4	I3	D4	I3	D5	I2

Fonte: Autor.

A Tabela 3 ilustra então, a distribuição das classes dirSU e indSU para o conjunto de empresas. As classes encontram-se perfeitamente equilibradas no que se referem aos conjuntos de registros dirSU e indSU, porém, entre as classes, há um desbalanceamento quanto à classe D5 (7,9%).

Tabela 3– Distribuição das classes dirSU e indSU, período 2013-2021

CLASSE dirSU	QUANT.	%	CLASSE indSU	QUANT.
<b>D1</b>	90	26.3	<b>I1</b>	90
<b>D2</b>	81	23.7	<b>I2</b>	81
<b>D3</b>	81	23.7	<b>I3</b>	81
<b>D4</b>	63	18.4	<b>I4</b>	63
<b>D5</b>	27	7.9	<b>I5</b>	27
<b>Totais</b>	<b>342</b>	<b>100.0</b>		<b>342</b>

Fonte: Autor.

### 5.3. APOIO MULTICRITÉRIO À DECISÃO

Visando o tratamento matemático de prováveis imprecisões de dados, aplicou-se a Teoria dos Conjuntos Aproximativos (*Rough Sets Theory*, RST)/princípio de Dominância (*Dominance Rough Sets Approach*, DRSA) como Apoio Multicritério à Decisão, a todo o conjunto de indicadores.

No contexto deste estudo, os indicadores foram considerados critérios de condição e a classe, critério de decisão. Conseqüentemente, foram geradas duas tabelas de decisão: uma relacionada aos indicadores e classe dirSU e, outra, aos indicadores e classe indSU. Em cada tabela de decisão, foram reunidos 342 (trezentos e quarenta e dois) registros sobre os indicadores e classes das empresas, no período 2013-2021.

Essas tabelas de decisão foram então processadas com o uso dos *softwares* jMAF (*Dominance-Based Rough Set Data Analysis Framework*; v. 08-03-2019; BLASZCZYNSKI *et al.*, 2013b) e *package RoughSets* em R (v. 1.3-7, de 15/12/2019; RIZA *et al.*, 2014, 2019).

#### 5.3.1. ANÁLISE COM APOIO DO SOFTWARE jMAF

Cada tabela de decisão foi então exportada de acordo com o padrão estabelecido pelo *software* jMAF, conforme *script* a seguir e a Figura 24 para indicadores dirSU. A seção “Examples” ilustra os três primeiros registros de um total de 342. Procedimento

análogo foi realizado para os indicadores indSU. O processamento das tabelas de decisão foi realizado individualmente.

**\*\*ATTRIBUTES**

+ codigo\_empresa: (nominal), none, description  
 + ano: (nominal), none, description  
 + IE01: (continuous), gain  
 + IE02: (continuous), cost  
 + IE03: (continuous), cost  
 + IE04: (continuous), gain  
 + IE05: (continuous), gain  
 + IE06: (continuous), cost  
 + IE07: (continuous), gain  
 + IE08: (continuous), gain  
 + IE09: (continuous), gain  
 + IE10: (continuous), gain  
 + IS01: (continuous), gain  
 + IS02: (continuous), gain  
 + IS03: (continuous), gain  
 + IS04: (continuous), gain  
 + IS05: (continuous), gain  
 + IS06: (continuous), gain  
 + IS07: (continuous), gain  
 + IS08: (continuous), gain  
 + IS09: (continuous), gain  
 + IS10: (continuous), cost  
 + IA01: (continuous), cost  
 + IA02: (continuous), cost  
 + IA03: (continuous), cost  
 + IG01: (continuous), gain  
 + IG02: (continuous), gain  
 + IG03: (continuous), gain  
 + IG04: (continuous), gain  
 + IG05: (continuous), gain  
 + IG06: (continuous), gain  
 + IG07: (continuous), gain  
 + IG08: (continuous), gain  
 + IG09: (continuous), gain  
 + IG10: (continuous), gain  
 + CLASSE: [D1, D2, D3, D4, D5], gain, decision  
 decision: CLASSE

**\*\*PREFERENCES**

codigo\_empresa: none  
 ano: none  
 IE01: gain  
 IE02: cost  
 IE03: cost  
 IE04: gain  
 IE05: gain  
 IE06: cost  
 IE07: gain  
 IE08: gain  
 IE09: gain  
 IE10: gain  
 IS01: gain  
 IS02: gain  
 IS03: gain  
 IS04: gain  
 IS05: gain  
 IS06: gain  
 IS07: gain  
 IS08: gain  
 IS09: gain

IS10: cost  
 IA01: cost  
 IA02: cost  
 IA03: cost  
 IG01: gain  
 IG02: gain  
 IG03: gain  
 IG04: gain  
 IG05: gain  
 IG06: gain  
 IG07: gain  
 IG08: gain  
 IG09: gain  
 IG10: gain  
 CLASSE: gain

\*\*EXAMPLES

Example\_1: E01, 2021, 953647, 14879, 4882, 28, 2729.4, 147003, 448, 330, 188, 435, 1121, 24.35, 25, 30.24, 2.14, 118, 25.06, 20165663.81, 443, 9, 93059, 10959, 15200, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, D4

Example\_2: E02, 2021, 136949.5, 961.415, 143.84, 17, 212.8, 45115, 22714, 60.76, 34389, 85140, 350, 38.735, 26.8, 23.1, 4.28, 120, 2877, 6486, 433, 10, 2953, 1583.61, 8063, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, D4

Example\_3: E03, 2021, 3228811, 19070.82, 2272.29, 166, 3829, 654010, 2563.495, 1357.725, 620.3, 434.305, 3747, 25.45, 24.71, 18.65, 4.875, 591.5, 34750.5, 47891, 1433.5, 10.3, 6129, 22143, 59045, 0, 0, 0, 0, 0, 0, 1, 0, 0, D1

Figura 24– Processamento de tabela de decisão, indicadores e classe dirSU com apoio do software jMAF

No	Object	co.	ano	IE0	IE1	IE2	IE3	IE4	IE5	IE6	IE7	IE8	IE9	IE10	IE11	IE12	IE13	IE14	IE15	IE16	IE17	IE18	IE19	IE20	IE21	IE22	IE23	IE24	IE25	IE26	IE27	IE28	IE29	IE30	IE31	IE32	IE33	IE34	IE35	IE36
1	Example_1	E01	2021	953647	14879.0	4882.000	28.000	2729.400	147003.0	448.000	330.000	188.000	435.000	1121.000	24.350	25.000	30.240	2.140	118.000	25.060	20165663.810	443.000	9.000	93059.000	10959.000	15200.000	1.000	0.000	0.000	0.000	1.000	0.000	1.000	1.000	0.000	D4				
2	Example_2	E02	2021	136949.5	961.415	143.840	17.000	212.800	45115.0	22714.0	60.760	34389.0	85140.0	350.000	38.735	26.800	23.100	4.280	120.000	2877.000	6486.000	433.000	10.000	2953.000	1583.610	8063.000	1.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.000	D4				
3	Example_3	E03	2021	3228811	19070.82	2272.290	166.000	3829.000	654010.0	2563.495	1357.725	620.300	434.305	3747.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
4	Example_4	E04	2021	888229	29482.6	4791.000	416.000	10980.0	846747.0	1695.000	3944.000	1226.000	728.000	3760.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
5	Example_5	E05	2021	38833.0	140.070	12.200	9.000	96.500	6185.000	9.980	22.930	5.000	14.270	11.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
6	Example_6	E06	2021	22104.0	225.000	26.000	55469.0	1650.000	14998.0	368.000	1760.000	1.000	319.000	13.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
7	Example_7	E07	2021	451963	25337.0	3071.000	378.000	10742.3	511280.0	71651.0	2562.000	1001.000	556.000	443.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
8	Example_8	E08	2021	151139	8960.640	1599.150	14.000	3761.800	352825.0	14.510	13.020	5.820	4.390	996.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
9	Example_9	E09	2021	180442	11541.0	1421.460	14.000	4030.000	352825.0	12402.0	7962.000	3420.000	8938.000	121.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
10	Example_10	E10	2021	263115	1328.400	276.700	13.000	312.000	41302.0	38.250	48.670	48.640	125.230	296.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
11	Example_11	E11	2021	209981	696.150	48.000	8.000	185.000	14646.0	69.290	23.920	34.390	32.700	23.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
12	Example_12	E12	2021	101552	5610.270	808.000	101.000	2466.000	354905.0	324.000	530.000	226.000	245.000	1394.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
13	Example_13	E13	2021	130365	8971.200	1446.200	159.000	3841.000	206084.0	660.200	1131.000	363.000	365.000	2366.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					
14	Example_14	E14	2021	445597	1513.870	175.300	47.000	3012.000	47981.0	130.910	181.630	85.470	33.070	631.000	25.450	24.710	18.650	4.875	591.500	34750.500	47891.000	1433.500	10.300	6129.000	22143.000	59045.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	D1					

Fonte: Autor.

A seguir, os seguintes passos foram executados: *Calculate: P-Dominance sets, DP+ Calculate dominating set, DP- Calculate dominated set; Unions of classes, standard, consistency level, 1.0; Reducts, all reducts; Rules, VC-DOMLEM algorithm, consistency level, 1.0, type of rules, certain, type of unions, standard.* Consequentemente, foram gerados o núcleo e 664 (seiscentos e sessenta e quatro) redutos para a tabela de

decisão dirSU. O resultado a seguir mostra o núcleo e os três primeiros redutos para indicadores dirSU.

1 core () of dados\_dirSU.isf (all attributes 36, active condition attributes 33, decision CLASSE, objects 342); Thu Jul 20 16:48:41 GMT-03:00 2023

{IE01,IE05,IA03,IG08,IG09}

664 reducts (all) of dados\_dirSU.isf (all attributes 36, active condition attributes 33, decision CLASSE, objects 342); Thu Jul 20 16:48:41 GMT-03:00 2023

{IE01,IE05,IE06,IS01,IS05,IS06,IS08,IS09,IA02,IA03,IG08,IG09}

{IE01,IE05,IE06,IS01,IS05,IS06,IS08,IS10,IA02,IA03,IG08,IG09}

{IE01,IE05,IE06,IS01,IS03,IS05,IS06,IS08,IA02,IA03,IG08,IG09}

O núcleo formado pelos indicadores {IE01, IE05, IA03, IG08, IG09} indica que esse subconjunto está presente (interseção) em todos os redutos, e que o torna indispensável ao conjunto de todos os indicadores.

### 5.3.2. ANÁLISE COM APOIO DO PACKAGE ROUGHSETS EM R

A tabela de decisão dirSU com 342 (trezentos e quarenta e dois) registros foi exportada como um *data frame*, de acordo com o padrão estabelecido pelo *package RoughSets* em R, conforme Figura 25. Procedimento análogo foi realizado para os indicadores indSU. O processamento das tabelas de decisão foi realizado de forma individual.

Figura 25– Exportação da tabela de decisão dirSU como *data frame* para o *package RoughSets* em R

```
dt.ex1 <- data.frame (
  c( 953647, 136949.5, 3228811, 8882293, 38835, 22104, 4519630, 1511594, 1804425, 263115, 209981, 1015526, 1365659, 445557, 105555, 1404298, 645131,
  637804, 1564347, 1886693, 263729, 212744, 1018108, 1403355, 455359, 108287, 1424082, 641995, 776347, 784064, 586458, 198051, 1158036, 1719675, 264356
  470, 102533, 1355210, 645131, 731480, 741943, 567469, 191254, 1137246, 1604207, 2311003, 2261602, 1172997, 6730430, 6205000, 3820000, 1494000, 277000
  c( 14879, 961.415, 19070.82, 29482.6, 140.07, 225, 25337, 8960.64, 11541, 1328.4, 696.15, 5610.27, 8971.2, 1513.87, 319.84, 4707.95, 4620.8, 3726.6
  4933, 11273, 5959, 5250, 4933, 21128, 14482, 5581, 13344, 9085, 224.2, 28747, 4921, 564.18, 394.5, 34.75, 13850, 1017.06, 18290.4, 33173, 135.31, 211
  696.15, 5610.27, 8971.2, 1513.87, 319.84, 4707.95, 4620.8, 3726.96, 4189.01, 2633.385, 1996, 4789, 4710, 9644, 4619, 5250, 4533, 21383, 14482, 5595,
  c( 4882, 143.84, 2272.29, 4791, 12.2, 26, 3071, 1599.15, 1421.46, 276.7, 48, 808, 1446.2, 175.3, 16, 651.8, 1225.73, 345.8, 295.25, 355.2, 948, 11
  1467.75, 1421.46, 289.6, 41.7, 805, 1448.37, 175.3, 16, 640.68, 1225.73, 334.26, 283.4, 334.5, 948, 1193.5, 1730, 3325, 1512, 1054, 840, 3276, 2736,
  651.8, 1225.73, 345.8, 295.25, 355.2, 948, 1193.5, 1736.5, 3590.5, 1512, 1054, 892, 3276, 2736, 632, 1534, 1059, 17.899, 935, 4.49, 66.35, 59, 4.54 )
  c( 28, 17, 166, 416, 9, 55469, 378, 14, 14, 13, 8, 101, 159, 47, 5, 63, 56, 33, 93, 101, 143, 44, 92, 145, 143, 98, 844, 341, 147, 65, 144, 41, 81,
  , 8, 55469, 378, 14, 14, 11, 7, 98, 156, 46, 5, 62, 56, 31, 90, 99, 140, 40, 65.5, 112, 127, 84, 977, 341, 147, 65, 144, 41, 64, 108, 0.01, 10, 6, 66
  c( 2729.4, 212.8, 3829, 10980, 96.5, 1650, 10742.3, 3761, 4010, 312, 185, 2466, 3841, 1012, 119, 1230, 1294, 741, 2334, 1454, 1703, 1038.75, 2186.6
  4106, 309.2, 182.5, 2466, 3841, 1012, 119, 1181, 1294, 741, 2335, 1454, 1611, 976, 2332.35, 3331, 3005, 1334, 14178, 6396, 3969, 1664, 6396, 3000, 17
  , 2186.625, 3331, 3005, 1334, 30706, 6396, 3969, 1664, 6118, 3000, 1738, 1153.5, 33.3, 661, 65, 56.63 )
  c( 147003, 45115, 654010, 846747, 6185, 14998, 511280.5, 352825.21, 352825.21, 41302, 14646, 154965, 206004, 47981, 8246, 97814, 157506, 63189, 877
  89, 88070, 94594, 33476, 133118, 376721, 140088, 128365.77, 103308, 125733, 416231, 256568, 108907, 375958, 209195, 547713, 364930, 124389, 11199, 21
  364930, 104160, 10148, 22076.5, 5254.715, 170029.5, 19594, 531624, 665288, 5039, 15573.5, 459923, 352825.21, 352825.21, 41302, 13938, 154965, 82355,
  c( 448, 22714, 2563.495, 1695, 9.98, 368, 71651, 14.51, 12402, 38.25, 69.29, 324, 660.2, 130.91, 39.65, 335.6, 305.39, 213.8, 383, 173, 280532, 386, 575, 1321, 392, 380, 400, 1688, 1089, 377,
  20, 76328, 11.45, 12402, 38.4, 64.8, 324, 656, 130.91, 39.65, 335.6, 305.39, 213.8, 383, 173, 280532, 386, 575, 1321, 392, 380, 400, 1688, 1089, 377,
  39, 213.8, 365.83, 173, 280532, 366, 1281, 722.5, 379, 389, 371, 1688, 1089, 377, 1778, 82.7, 30.96, 5052, 26.27, 383, 371, 7.17 )
  c( 330, 60.76, 1357.725, 3944, 22.93, 1760, 2562, 13.02, 7962, 48.67, 23.92, 530, 1131, 103.63, 5.44, 275, 337.06, 117, 305, 221, 4743, 327, 559, 1
  131, 183.63, 5.56, 267, 337.06, 108.81, 295, 221, 4743, 276, 599, 271, 278, 281, 279, 1997, 678, 447, 1005, 135.34, 35.71, 287, 0.22, 286, 286, 5.05,
  , 0.21, 262, 272, 3.39 )
  c( 188, 34389, 620.3, 1226, 5, 1, 1001, 5.82, 3420, 48.64, 34.39, 226, 363, 83.47, 20.68, 258.6, 154.15, 184.59, 188.5, 119.69, 36142, 200, 239, 46
  8.6, 154.15, 184.59, 190, 121, 36142, 196, 238, 476, 192, 199, 184, 1079, 487, 194, 527, 436.71, 9.925, 194, 3.24, 186, 199, 2.47, 188, 41674, 605.6:
  2.47 )
)
```

Fonte: Autor.

### 5.3. PADRÕES E PREDIÇÃO

Consequentemente, após o processamento descrito na etapa anterior, foi possível a indução de regras de decisão “if ... then ...”, bem como, a identificação de padrões sobre os indicadores e classes.

#### 5.3.1. INDUÇÃO DE REGRAS COM APOIO DO *SOFTWARE* jMAF

Com o uso do algoritmo VC-DomLEM (Blaszczynski *et al.*, 2009, 2011), Anexo V, foram geradas 172 (cento e setenta e duas) regras de decisão para indicadores e classes dirSU, cujas três primeiras estão descritas a seguir, Anexo VI. Procedimento análogo foi obtido para indicadores e classes indSU, com a indução de 138 (cento e trinta e oito) regras de decisão, Anexo VII.

[RULES]

#Certain at least rules

1: (IA01 <= 0.204) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|

2: (IE07 >= 76328.0) & (IS04 >= 22.52) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|

3: (IE03 <= 276.7) & (IS10 <= 0.17) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|

Foram selecionadas aquelas regras mais representativas, ou seja, com maiores valores de *Coveragefactor*, Equação (7), visando identificar padrões sobre os indicadores e classes dirSU, conforme as regras “1” e “172” descritas a seguir.

1: (IA01 <= 0.204) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|

LearningPositiveExamples: 5, 7, 21, 43, 45, 59, 81, 97, 101, 119, 121, 135, 157, 159, 173, 195, 211, 214, 233, 249, 252, 271, 287, 290, 309, 314, 325

Support: 12

SupportingExamples: 5, 43, 81, 119, 157, 195, 214, 233, 252, 271, 290, 309

Strength: 0.03508771929824561

Confidence: 1.0

CoverageFactor: 0.4444444444444444

Coverage: 12

CoveredExamples: 5, 43, 81, 119, 157, 195, 214, 233, 252, 271, 290, 309

172: (IS03 <= 26.8) & (IA02 >= 1583.61) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|

LearningPositiveExamples: 1, 2, 3, 4, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98, 99, 100, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 120, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 158, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 212, 213, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 250, 251, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 288, 289, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 310, 311, 312, 313, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342

Support: 245

SupportingExamples: 1, 2, 3, 4, 6, 8, 9, 14, 15, 16, 17, 18, 20, 25, 27, 28, 29, 30, 31, 33, 34, 35, 37, 38, 39, 40, 41, 42, 44, 46, 47, 48, 49, 50, 51, 52, 54, 55, 56, 58, 60, 63, 64, 65, 66, 67, 68, 69, 71, 72, 73, 75, 76, 77, 78, 79, 80, 82, 84, 85, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98, 102, 103, 104, 105, 106, 107, 109, 110, 111, 113, 114, 115, 116, 117, 118, 120, 122, 125, 126, 127, 128, 130, 131, 132, 133, 134, 136, 140, 141, 142, 143, 144, 145, 147, 148, 149, 151, 152, 153, 154, 155, 156, 158, 160, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 174, 178, 179, 180, 181, 182, 183, 185, 186, 187, 189, 190, 191, 192, 194, 196, 198, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 212, 217, 218, 219, 220, 221, 223, 224, 225, 227, 228, 229, 231, 232, 234, 236, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 250, 251, 255, 256, 257, 258, 259, 261, 262, 263, 265, 266, 267, 269, 270, 272, 274, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 288, 289, 293, 294, 295, 296, 297, 299, 300, 301, 303, 304, 305, 306, 307, 308, 310, 312, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 326, 330, 331, 332, 333, 334, 335, 337, 338, 339, 341, 342

Strength: 0.716374269005848

Confidence: 1.0

CoverageFactor: 0.7777777777777778

Coverage: 245

CoveredExamples: 1, 2, 3, 4, 6, 8, 9, 14, 15, 16, 17, 18, 20, 25, 27, 28, 29, 30, 31, 33, 34, 35, 37, 38, 39, 40, 41, 42, 44, 46, 47, 48, 49, 50, 51, 52, 54, 55, 56, 58, 60, 63, 64, 65, 66, 67, 68, 69, 71, 72, 73, 75, 76, 77, 78, 79, 80, 82, 84, 85, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98, 102, 103, 104, 105, 106, 107, 109, 110, 111, 113, 114, 115, 116, 117, 118, 120, 122, 125, 126, 127, 128, 130, 131, 132, 133, 134, 136, 140, 141, 142, 143, 144, 145, 147, 148, 149, 151, 152, 153, 154, 155, 156, 158, 160, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 174, 178, 179, 180, 181, 182, 183, 185, 186, 187, 189, 190, 191, 192, 194, 196, 198, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 212, 217, 218, 219, 220, 221, 223, 224, 225, 227, 228, 229, 231, 232, 234, 236, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 250, 251, 255, 256, 257, 258, 259, 261, 262, 263, 265, 266, 267, 269, 270, 272, 274, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 288, 289, 293, 294, 295, 296, 297, 299, 300, 301, 303, 304, 305, 306, 307, 308, 310, 312, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 326, 330, 331, 332, 333, 334, 335, 337, 338, 339, 341, 342

Porém, somente 39,4% dos indicadores dirSU foram referenciados em 70,1% do total de referências pelas 172 regras de decisão induzidas e, 34,5% dos indicadores indSU foram referenciados em 70,4% do total de referências pelas 138 regras de decisão induzidas. As Tabelas 4 e 5 ilustram estes indicadores.

Tabela 4– Indicadores dirSU referenciados em 70,1% do total de referências pelas regras de decisão, algoritmo VC-DomLEM

<b>Indicador</b>	<b>Descrição</b>	<b>% Referência</b>
IS03	Empregados com idade entre 41 e 50 anos (%)	9,9
IA03	Consumo total de água, fonte Rede pública	8,5
IS04	% mulheres/total de empregados	7,0
IS06	Benefícios com Educação	5,4
IS10	Taxa de rotatividade	5,4
IE07	Energia vendida, setor Industrial	5,1
IS05	% empregados programa contratação aprendizes	5,1
IS09	Benefícios com Auxílio-creche	4,8
IE01	Número de consumidores atendidos	4,5
IA01	Consumo de energia direta, fonte Diesel	3,9
IE03	Perdas elétricas globais	3,7
IA02	Consumo de energia direta, fonte Gasolina	3,7
IE06	Folha de pagamento bruta	3,4

Fonte: Autor.



Tabela 5– Indicadores indSU referenciados em 70,4% do total de referências pelas regras de decisão, algoritmo VC-DomLEM

Indicador	Descrição	% Referência
IA01	Consumo de água por empregado (em m <sup>3</sup> )	12,6
IE05	Número de consumidores por empregado	10,5
IE03	Número de transformadores de distribuição	7,9
IS03	Número médio de atendentes	6,9
IS09	Número de atendimentos realizados	6,9
IA02	Número de alunos atendidos	6,5
IE01	Número de localidades atendidas (municípios)	5,1
IS08	FER (unid.)	5,1
IE04	Energia vendida por empregado	4,7
IS01	Média de horas extras por empregado/ano	4,3

Fonte: Autor.

Com o uso de indicadores mais referenciados nas regras de decisão, conforme Tabelas 4 e 5, propõe-se uma alternativa de forma heurística à identificação de atributos ou subconjuntos de atributos de maior importância por meio do cálculo da região positiva de uma classificação (PAWLAK, 1991).

Visando obter os riscos de ascensão e descensão das empresas concessionárias de distribuição de energia elétrica nas classes dirSU e indSU e, considerando-se todo o conjunto de indicadores e classes no período 2013-2021, propõe-se a elaboração de uma Matriz de Risco Condicional: quando há risco associado a algum conhecimento prévio específico e consequências, neste caso, probabilidade condicional de indicadores e classes de decisão, respectivamente (AVEN, 2015; AVEN; YLONEN, 2018). A partir dos maiores valores do fator *coverage* e respectivas regras de decisão inferidas, mas não de forma exaustiva, e com apoio do algoritmo VC-DomLEM, as matrizes de risco condicional propostas estão ilustradas conforme Quadros 24 e 25.

Quadro 24– Matriz de Risco Condicional para indicadores e classes dirSU

#	Probabilidade condicional	Regra	Classes dirSU				
			D1	D2	D3	D4	D5
1	44,4%: (IA01 ≤ 0.204)	1					✓
2	44,4%: (IE02 ≤ 25751.0) & (IE07 ≥ 71651.0)	4					✓
3	22,8%: (IE07 ≥ 21565.0) & (IG08 ≥ 1.0)	9				✓	✓
4	77,7%: (IS03 ≤ 26.8) & (IA02 ≥ 1583.61)	172	✓	✓	✓	✓	
5	18,0%: (IS03 ≤ 16.5) & (IS09 ≤ 535.0)	153	✓	✓	✓		
6	29,1%: (IS03 ≤ 18.37) & (IS09 ≤ 359.0)	125	✓	✓			
7	22,3%: (IE01 ≤ 1365659.0) & (IS03 ≤ 15.22) & (IS04 ≤ 13.61)	99	✓				

Fonte: Autor; adaptado de Aven (2015); Aven; Ylonen (2018).

Quadro 25– Matriz de Risco Condicional para indicadores e classes indSU

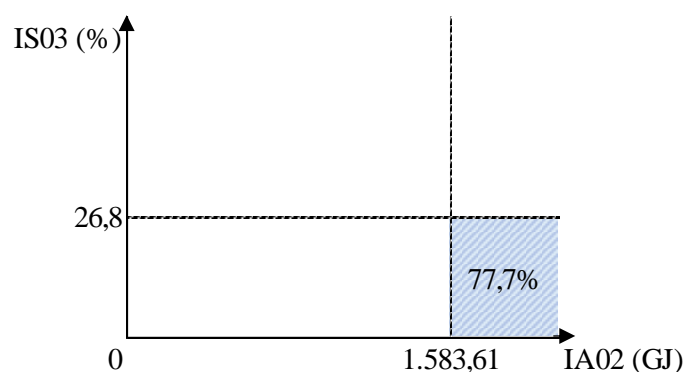
#	Probabilidade condicional	Regra	Classes indSU				
			I1	I2	I3	I4	I5
1	29,6%: (IS03 >= 404.0) & (IS08 <= 8.21)	2					✓
2	45,5%: (IE03 >= 182079.0) & (IS07 <= 199.32)	13				✓	✓
3	31,3%: (IE05 <= 562.0) & (IA01 >= 4.5)	111	✓	✓	✓		
4	47,9%: (IE01 <= 72.0) & (IG10 <= 0.0)	92	✓	✓			
5	48,8%: (IE05 <= 390.45) & (IS07 >= 128.0)	71	✓				

Fonte: Autor; adaptado de Aven (2015); Aven; Ylonen (2018).

Do Quadro 24, constata-se pela probabilidade condicional #4: há uma probabilidade de 77,7% das empresas concessionárias de distribuição de energia elétrica não pertencentes à classe D5, apresentarem “(IS03 <= 26.8) & (IA02 >= 1583.61)”, ou seja, possuírem 26,8% ou menos de empregados com idade entre 41 e 50 anos, e consumirem gasolina como fonte de energia direta em quantidade igual ou superior a 1.583,61 GJ (gigajoule). E, do Quadro 25, probabilidade condicional #1: há probabilidade de 29,6% daquelas empresas que possuem classificação mais elevada (I5), apresentarem “(IS03 >= 404.0) & (IS08 <= 8.21)”, isto é, ter a média mínima de 404 atendentes, e Frequência Equivalente de Reclamação a cada mil unidades consumidoras (FER) de no máximo 8,21.

E, para a probabilidade condicional #4 do Quadro 24, a Figura 26 ilustra a probabilidade de uma empresa pertencer à região hachurada dada que a classe dirSU não seja D5. Essa probabilidade condicional pode perfeitamente ser verificada e comprovada mediante análise da tabela de decisão dirSU: há 245 registros que satisfazem à regra “(IS03 <= 26.8) & (IA02 >= 1583.61)” frente ao total de 315 registros com classificação D1, D2, D3 ou D4; ou,  $245 / 315 = 0,777$ .

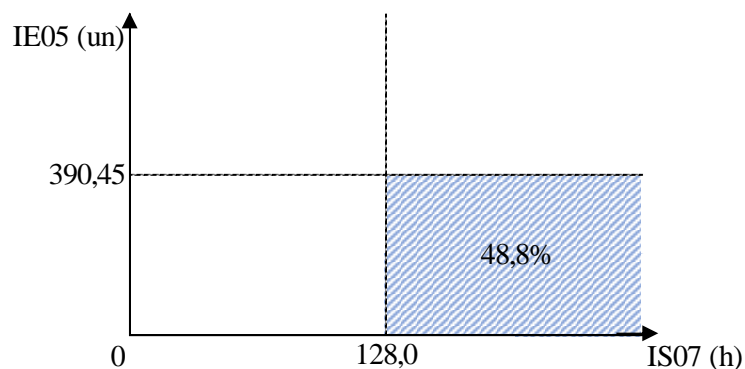
Figura 26– Probabilidade condicional #4, Quadro 24: probabilidade de uma empresa pertencer à região hachurada dada que a classe dirSU não seja D5



Fonte: Autor.

Do Quadro 25, pela probabilidade condicional #5, há uma probabilidade de 48,8% das empresas pertencentes à classe I1, de apresentarem “(IE05  $\leq$  390.45) & (IS07  $\geq$  128.0)”, ou seja, um número máximo de 390 consumidores por empregado, e mínimo de 128 horas para a Duração Equivalente de Reclamação (DER) por unidade consumidora, relacionado ao prazo médio de solução. A Figura 27 ilustra essa probabilidade condicional.

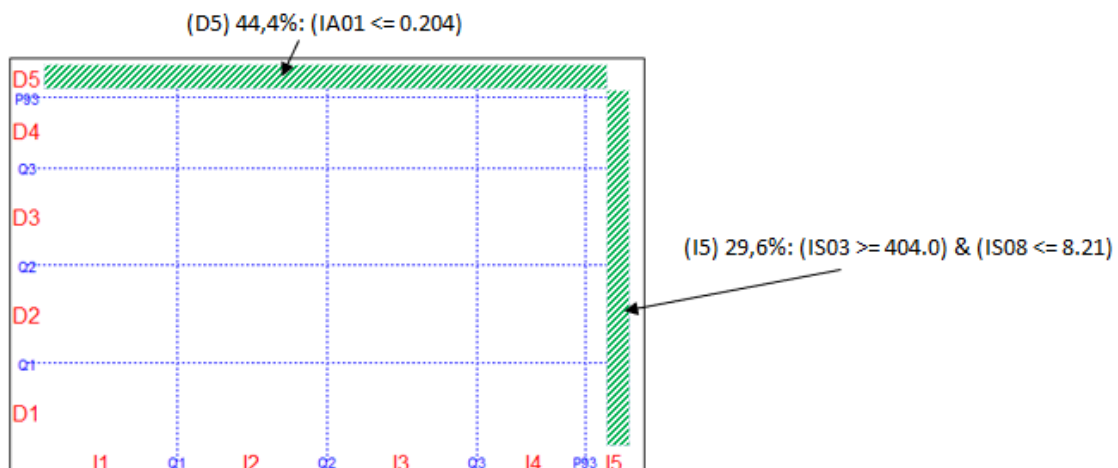
Figura 27– Probabilidade condicional #5, Quadro 25: probabilidade de uma empresa pertencer à região hachurada dada que a classe indSU seja I1



Fonte: Autor.

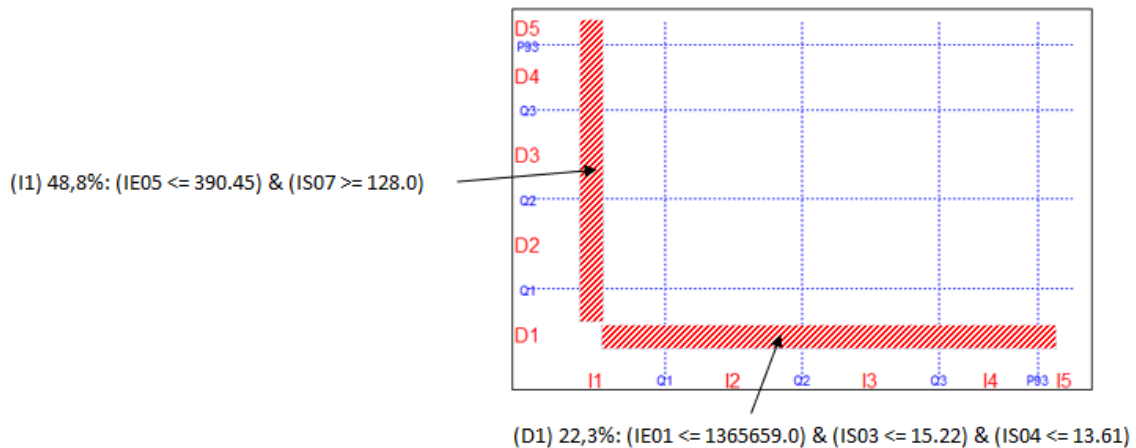
Com base nos Quadros 24 e 25, as Figuras 28 e 29 exemplificam de forma simultânea, as regiões mais (hachuradas em verde) e menos (hachuradas em vermelho) vantajosas na Matriz de Sustentabilidade, respectivamente, vis-à-vis às probabilidades condicionais.

Figura 28– Regiões mais vantajosas (hachuradas em verde) na Matriz de Sustentabilidade vis-à-vis às probabilidades condicionais



Fonte: Autor.

Figura 29– Regiões menos vantajosas (hachuradas em vermelho) na Matriz de Sustentabilidade vis-à-vis às probabilidades condicionais



Fonte: Autor.

### 5.3.2. INDUÇÃO DE REGRAS COM APOIO DO *PACKAGE ROUGHSETS* EM R

Com o uso dos algoritmos *RI.AQRules.RST*, *RI.CN2Rules.RST*, *RI.indiscernibilityBasedRules.RST* e *RI.LEM2Rules.RST* (CLARK; NIBLETT, 1989; MICHALSKI *et al.*, 1991; GRZYMALA-BUSSE, 1997; RIZA *et al.*, 2014, 2019) foram inferidas 827 (oitocentos e vinte e sete) regras de decisão com base na tabela de decisão *dirSU*, conforme resultados a seguir. Procedimento análogo para a tabela de decisão *indSU*.

```
[1] "RI.AQRules.RST"
> rules1
A set consisting of 538 rules:
1. IF IS04 is [19.9,24.3) and IE03 is [915,1.83e+03) and IS10 is [-Inf,6.68) THEN CLASSE is D3;
   (supportSize=5; laplace=0.5)
2. IF IE06 is [1.31e+05,2.57e+05) and IE02 is [-Inf,1.02e+04) and IS03 is [21.9,28.5) THEN CLASSE is D3;
   (supportSize=1; laplace=0.3333)
3. IF IS04 is [15.4,19.9) and IS02 is [35.2,42.2) and IG01 is [-Inf,0.143) and IE06 is [-Inf,1.31e+05) THEN CLASSE
   is D3;
   (supportSize=3; laplace=0.5)
4. IF IS01 is [3.03e+03,4.53e+03) and IG06 is [0.857, Inf] THEN CLASSE is D3;
   (supportSize=11; laplace=0.6875)
5. IF IS10 is [6.68,13.3) and IG06 is [-Inf,0.143) and IS05 is [5.59,8.38) THEN CLASSE is D3;
   (supportSize=1; laplace=0.3333)
6. IF IG06 is [-Inf,0.143) and IE10 is [7.41e+04,8.89e+04) THEN CLASSE is D3;
   (supportSize=2; laplace=0.4286)
7. IF IG01 is [-Inf,0.143) and IS10 is [-Inf,6.68) THEN CLASSE is D3;
   (supportSize=4; laplace=0.5556)
8. IF IS07 is [1.14e+07, Inf] THEN CLASSE is D3;
   (supportSize=5; laplace=0.6)
9. IF IS06 is [-Inf,576) and IE03 is [2.74e+03,3.65e+03) and IG06 is [-Inf,0.143) THEN CLASSE is D3;
```

(supportSize=1; laplace=0.3333)  
 10. IF IS05 is [-Inf,2.8) and IG06 is [-Inf,0.143) and IS02 is [49.3,56.4) THEN CLASSE is D3;  
 (supportSize=1; laplace=0.3333)  
 ... and 528 other rules.

[1] "RI.CN2Rules.RST"

> rules2

A set consisting of 36 rules:

1. IF IA01 is [1.12e+08,1.5e+08) THEN CLASSE is D3;  
 (supportSize=8; laplace=0.6923)
  2. IF IE07 is [2.42e+05, Inf] THEN CLASSE is D5;  
 (supportSize=7; laplace=0.6667)
  3. IF IE08 is [6.82e+03, Inf] THEN CLASSE is D2;  
 (supportSize=5; laplace=0.6)
  4. IF IS07 is [1.14e+07, Inf] THEN CLASSE is D3;  
 (supportSize=5; laplace=0.6)
  5. IF IE05 is [4.51e+03,8.98e+03) and IS06 is [576,1.15e+03) and IG01 is [0.857, Inf] THEN CLASSE is D4;  
 (supportSize=10; laplace=0.7333)
  6. IF IE03 is [1.83e+03,2.74e+03) and IS06 is [-Inf,576) THEN CLASSE is D3;  
 (supportSize=9; laplace=0.6429)
  7. IF IS03 is [-Inf,15.3) and IG08 is [-Inf,0.143) and IA02 is [-Inf,2.97e+06) and IS06 is [-Inf,576) and IS01 is [-Inf,1.53e+03) THEN CLASSE is D1;  
 (supportSize=14; laplace=0.7895)
  8. IF IS01 is [1.53e+03,3.03e+03) and IS03 is [15.3,21.9) and IE08 is [-Inf,1.14e+03) THEN CLASSE is D2;  
 (supportSize=13; laplace=0.7778)
  9. IF IG05 is [0.857, Inf] THEN CLASSE is D5;  
 (supportSize=4; laplace=0.5556)
  10. IF IS06 is [576,1.15e+03) and IS02 is [35.2,42.2) and IG06 is [0.857, Inf] and IG09 is [-Inf,0.143) THEN CLASSE is D2;  
 (supportSize=6; laplace=0.6364)
- ... and 26 other rules.

[1] "RI.indiscernibilityBasedRules.RST"

> rules3

A set consisting of 173 rules:

1. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [-Inf,28.1) and IS03 is [28.5,35.2) and IS04 is [19.9,24.3) and IS05 is [2.8,5.59) and IS10 is [6.68,13.3) and IA01 is [-Inf,3.74e+07) and IG01 is [-Inf,0.143) and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D5;  
 (supportSize=1; laplace=0.3333)
2. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [15.3,21.9) and IS04 is [15.4,19.9) and IS05 is [2.8,5.59) and IS10 is [6.68,13.3) and IA01 is [-Inf,3.74e+07) and IG01 is [-Inf,0.143) and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D1;  
 (supportSize=4; laplace=0.5556)
3. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [15.3,21.9) and IS04 is [6.64,11) and IS05 is [-Inf,2.8) and IS10 is [6.68,13.3) and IA01 is [-Inf,3.74e+07) and IG01 is [0.857, Inf] and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D4;  
 (supportSize=1; laplace=0.3333)
4. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [15.3,21.9) and IS04 is [6.64,11) and IS05 is [2.8,5.59) and IS10 is [6.68,13.3) and IA01 is [-Inf,3.74e+07) and IG01 is [-Inf,0.143) and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D1;  
 (supportSize=2; laplace=0.4286)
5. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [21.9,28.5) and IS04 is [11,15.4) and IS05 is [5.59,8.38) and IS10 is [6.68,13.3) and IA01 is [-Inf,3.74e+07) and IG01 is [0.857, Inf] and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D3;  
 (supportSize=1; laplace=0.3333)
6. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [21.9,28.5) and IS04 is [15.4,19.9) and IS05 is [2.8,5.59) and IS10 is [6.68,13.3) and IA01 is [-Inf,3.74e+07) and IG01 is [-Inf,0.143) and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D1;  
 (supportSize=1; laplace=0.3333)
7. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [21.9,28.5) and IS04 is [15.4,19.9) and IS05 is [2.8,5.59) and IS10 is [6.68,13.3) and IA01 is [-Inf,3.74e+07) and IG01 is [0.857, Inf] and IG06 is [0.857, Inf] and IG09 is [-Inf,0.143) THEN CLASSE is D4;  
 (supportSize=1; laplace=0.3333)

8. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [21.9,28.5) and IS04 is [6.64,11) and IS05 is [-Inf,2.8) and IS10 is [-Inf,6.68) and IA01 is [-Inf,3.74e+07) and IG01 is [-Inf,0.143) and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D3;  
 (supportSize=2; laplace=0.4286)

9. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [21.9,28.5) and IS04 is [6.64,11) and IS05 is [-Inf,2.8) and IS10 is [-Inf,6.68) and IA01 is [-Inf,3.74e+07) and IG01 is [0.857, Inf] and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D4;  
 (supportSize=1; laplace=0.3333)

10. IF IE01 is [-Inf,2.52e+06) and IE03 is [-Inf,915) and IE06 is [-Inf,1.31e+05) and IE08 is [-Inf,1.14e+03) and IS02 is [28.1,35.2) and IS03 is [21.9,28.5) and IS04 is [6.64,11) and IS05 is [-Inf,2.8) and IS10 is [13.3,19.9) and IA01 is [-Inf,3.74e+07) and IG01 is [0.857, Inf] and IG06 is [-Inf,0.143) and IG09 is [-Inf,0.143) THEN CLASSE is D4;  
 (supportSize=1; laplace=0.3333)

... and 163 other rules.

[1] "RI.LEM2Rules.RST"

> rules4

A set consisting of 80 rules:

1. IF IE06 is [-Inf,1.31e+05) and IG01 is [-Inf,0.143) and IS04 is [15.4,19.9) and IS02 is [35.2,42.2) THEN CLASSE is D3;  
 (supportSize=3; laplace=0.5)

2. IF IE01 is [-Inf,2.52e+06) and IE03 is [915,1.83e+03) and IE06 is [-Inf,1.31e+05) and IS05 is [-Inf,2.8) and IS02 is [42.2,49.3) THEN CLASSE is D3;  
 (supportSize=3; laplace=0.5)

3. IF IG08 is [0.857, Inf] and IE02 is [1.02e+04,2.04e+04) and IS05 is [-Inf,2.8) THEN CLASSE is D3;  
 (supportSize=10; laplace=0.7333)

4. IF IE06 is [2.57e+05,3.82e+05) and IS02 is [42.2,49.3) and IG06 is [-Inf,0.143) THEN CLASSE is D3;  
 (supportSize=3; laplace=0.5)

5. IF IS10 is [-Inf,6.68) and IG01 is [-Inf,0.143) THEN CLASSE is D3;  
 (supportSize=4; laplace=0.5556)

6. IF IS01 is [-Inf,1.53e+03) and IG06 is [-Inf,0.143) and IS02 is [49.3,56.4) THEN CLASSE is D3;  
 (supportSize=6; laplace=0.6364)

7. IF IE01 is [2.52e+06,5.03e+06) and IS03 is [21.9,28.5) and IE06 is [1.31e+05,2.57e+05) THEN CLASSE is D3;  
 (supportSize=1; laplace=0.3333)

8. IF IG01 is [0.857, Inf] and IS03 is [21.9,28.5) and IE03 is [-Inf,915) and IG06 is [-Inf,0.143) and IS10 is [6.68,13.3) THEN CLASSE is D3;  
 (supportSize=5; laplace=0.6)

9. IF IA02 is [-Inf,2.97e+06) and IS01 is [-Inf,1.53e+03) and IG08 is [0.857, Inf] and IE02 is [-Inf,1.02e+04) and IS05 is [2.8,5.59) and IS10 is [-Inf,6.68) THEN CLASSE is D3;  
 (supportSize=4; laplace=0.5556)

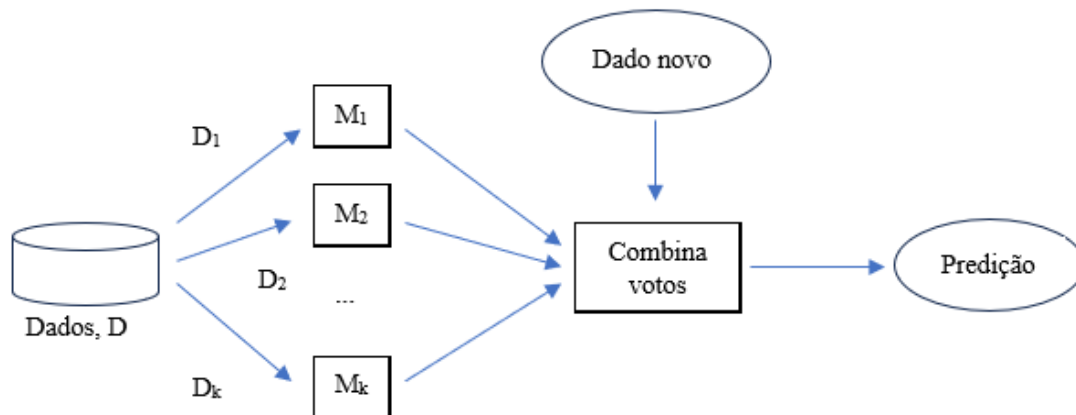
10. IF IE06 is [2.57e+05,3.82e+05) and IS05 is [5.59,8.38) THEN CLASSE is D3;  
 (supportSize=1; laplace=0.3333)

... and 70 other rules.

### 5.5.3. PREDIÇÃO DE CLASSES

Visando aumentar a acurácia na predição de classes com base em dados históricos dos indicadores e correspondentes classificações das empresas, algoritmos de indução de regras de decisão e valores simulados de indicadores, utilizou-se de método conjunto (*ensemble*) de modelos classificadores (HAN *et al.*, 2012), conforme Figura 30. Ademais, o método *ensemble* aplica-se ao contexto atual tendo em vista a distribuição desbalanceada de classes.

Figura 30– Método *ensemble* com “k” modelos classificadores, votação de classes e classe predita



Fonte: Han *et al.* (2012).

De acordo com Han *et al.* (2012), neste método, gera-se um conjunto de modelos de classificação  $M_1, M_2, \dots, M_k$ , com base nas amostras  $D_1, D_2, \dots, D_k$ , e, havendo um novo dado a classificar, cada classificador “vota” (induz) uma classe. O método então, combina os votos para retornar a classe mais votada ou predita. Os métodos *ensemble* são os seguintes: a) *Bagging*: a classe predita é a mais votada, sem qualquer atribuição de peso ou valor a cada voto; b) *Boosting*: quando se assinala um peso ou valor a cada voto; c) *Random Forests*: cada classificador é uma árvore de decisão com uma seleção aleatória de atributos em cada nó, a coleção de classificadores forma uma “floresta” de árvores de decisão e, vota-se na classe mais popular.

Para o contexto deste estudo e, inicialmente, propõe-se o método *ensemble Bagging*, constituído dos algoritmos classificadores indutores de regras de decisão “if ... then ...”, descritos no Quadro 26.

Quadro 26– Algoritmos classificadores indutores de regras de decisão

Classificador	Descrição resumida	Nome externo; software de apoio	Referência(s)
AQ	Trata-se de uma família de algoritmos de indução de regras, tais como AQ14-NT, AQ15-GA, AQ15-FCLS, AQ17-DCI, AQ17-HCI e AQ17-1. Com base em exemplos encontrados, gera-se regras e escolhe a melhor regra por meio de uma função de qualidade (lexicográfica); remove-se exemplos cobertos por regra inferida. Repete-se o processo.	RI.AQRules.RST; <i>package</i> RoughSets em R	Michalski <i>et al.</i> (1991); Riza <i>et al.</i> (2014, 2019)
CN2	Em cada iteração procura-se por exemplos de uma única classe “C” e poucos de outras classes. Ao encontrar exemplos que cobrem o conjunto de treinamento adiciona a regra ao final da lista de regras. Este processo se repete até que não sejam encontrados mais exemplos satisfatórios.	RI.CN2Rules.RS T; <i>package</i> RoughSets em R	Clark; Niblett (1989); Riza <i>et al.</i> (2014, 2019)
DecisionTable	Utiliza-se de tabela de decisão para construir um classificador. Pode utilizar a validação cruzada ( <i>cross-validation</i> ) para avaliação, bem como, o método do vizinho mais próximo ( <i>nearest-neighbor</i> ) para determinar a classe para cada instância, que não seja coberta por uma entrada na tabela de decisão.	DecisionTable; Weka	Kohavi (1995); Witten; Frank (2005)
HybridFS	<i>Hybrid fuzzy-rough rule and induction and feature selection</i> . Algoritmo para indução de regras <i>fuzzy-rough</i> híbridas e <i>feature selection</i> (seleciona recursos significativos e elimina os dispensáveis), utilizadas em tarefas de classificação e regressão.	RI.hybridFS.FRS T; <i>package</i> RoughSets em R	Jensen <i>et al.</i> (2009); Riza <i>et al.</i> (2014, 2019)
Indiscernibility Based Rules	Gera regras de decisão "if-then" a partir de classes de indiscernibilidade definidas por um determinado subconjunto de atributos condicionais.	RI.indiscernibility BasedRules.RST; <i>package</i> RoughSets em R	Pawlak (1982); Riza <i>et al.</i> (2014, 2019)
JRip	Utiliza o algoritmo RIPPER ( <i>Repeated Incremental Pruning to Produce Error Reduction</i> ) com otimização heurística e métricas para poda ou não de regras induzidas.	JRip; Weka	Cohen (1995); Witten; Frank (2005)
LEM2	Origem no algoritmo LERS ( <i>Learning from Examples using Rough Sets</i> ), que trabalha com “conceito”, isto é, critério de decisão com mesmo valor. Cada regra de decisão gerada é avaliada com base nas qualidades das aproximações inferior e superior, e probabilidade condicional entre critérios de condição e de decisão.	RI.LEM2Rules.R ST; <i>package</i> RoughSets em R	Grzymala-Busse (1997); Riza <i>et al.</i> (2014, 2019)

Fonte: Autor; adaptado dos autores relacionados.



Quadro 26 (cont.) – Algoritmos classificadores indutores de regras de decisão

Classificador	Descrição resumida	Nome externo; <i>software de apoio</i>	Referência(s)
OneR	Gera regra para cada atributo dos dados de treinamento, determina a classe mais frequente para cada atributo e seleciona a regra com menor percentagem de erro.	OneR; Weka	Ali; Smith (2006); Witten; Frank (2005)
PART	Algoritmo de árvore de decisão parcial; versão desenvolvida dos algoritmos C4.5 e RIPPER.	PART; Weka	Ali; Smith (2006); Witten; Frank (2005)
VC-DomLEM	Algoritmo de indução de conjunto mínimo de regras de decisão com base em cobertura sequencial e consistência variável (VC) por meio de <i>Rough Sets Theory</i> . As regras de decisão são utilizadas tanto para descrever os dados como para realizar novas classificações (predição). Utiliza tanto <i>Indiscernibility-based Rough Set Approach (IRSA)</i> como <i>Dominance-based Rough Set Approach (DRSA)</i> .	VC-DomLEM; jMAF	Blaszczynski <i>et al.</i> (2009, 2011, 2013b); Greco <i>et al.</i> (2001)
ZeroR	Realiza a predição pela média, quando valor numérico e, pela moda, para classes nominais.	ZeroR; Weka	Witten; Frank (2005)

Fonte: Autor; adaptado dos autores relacionados.

Segundo Swets (1988), Han *et al.* (2012), Provost; Fawcett (2013) e Vluymans (2019), a acurácia (*accuracy*) de um classificador em um determinado conjunto de teste, é a porcentagem de tuplas ou registros do conjunto de teste que são classificados corretamente pelo classificador. Porém, nem sempre a acurácia pode refletir perfeitamente o desempenho de um classificador. Quando a distribuição de classes ocorre de forma desbalanceada (Tabela 3), há a necessidade de se usar outras medidas, tal como, a sensibilidade (*sensitivity*, ou *recall*) ou *TP Rate (true positive / positive)*, que reflete a proporção de tuplas positivas que são identificadas corretamente, considerando-se o cálculo por cada classe. Há ainda, a especificidade (*specificity*) ou taxa negativa: proporção de tuplas negativas que são identificadas corretamente. Tuplas positivas (TP) são aquelas consideradas da classe principal de interesse e que são identificadas de forma correta pelo classificador; tuplas negativas (TN) são as demais que são identificadas também de forma correta pelo classificador; e, falsas positivas (FP) ou falsas negativas (FN) são aquelas tuplas identificadas sob forma incorreta pelo classificador. A matriz que resume essas situações é representada pela Matriz de Confusão (Han *et al.*, 2012), Quadro 27.

Quadro 27– Matriz de Confusão

		Classes preditas		Total
		Sim	Não	
Classes reais	Sim	TP	FN	$P$
	Não	FP	TN	$N$
	Total	$P'$	$N'$	$P + N$

Fonte: Han *et al.* (2012).

Assim, procedeu-se à apuração de medidas que viessem a demonstrar o quanto cada algoritmo conseguiria efetuar de forma correta a predição de classes para novos dados, conforme Tabela 6.

Tabela 6– Algoritmos de classificação indutores de regras de decisão e respectivas medidas de aferição

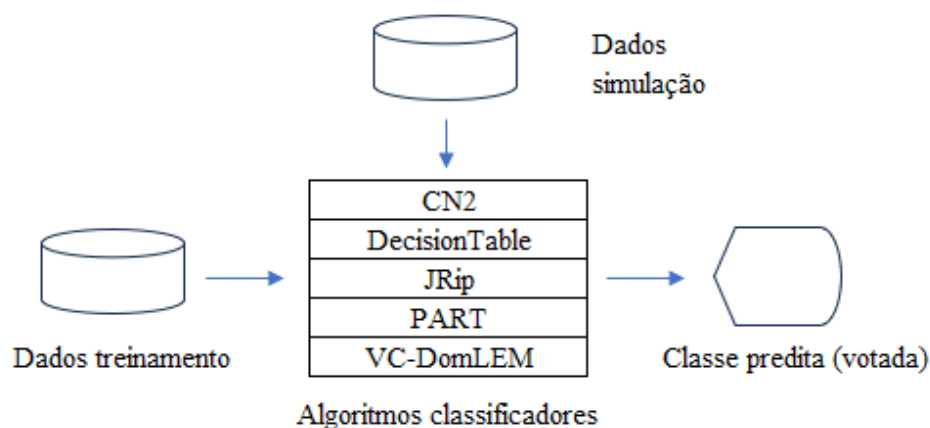
Classificador	Nome externo; <i>software</i> de apoio	Acurácia geral (%)		Sensibilidade média (%)	
		dirSu	indSU	dirSU	indSU
AQ	RI.AQRules.RST; <i>package RoughSets</i> em R	49,6	59,8	46,2	60,1
CN2	RI.CN2Rules.RST; <i>package RoughSets</i> em R	61,5	66,7	64,1	68,3
DecisionTable	DecisionTable; Weka	66,3	64,0	66,4	64,0
HybridFS	HybridFS; Weka	46,4	40,6	33,6	51,2
Indiscernibility Based Rules	RI.indiscernibilityBased Rules.RST; <i>package RoughSets</i> em R	47,0	40,2	44,0	38,8
JRip	JRip; Weka	62,2	65,2	62,3	65,2
LEM2	RI.LEM2Rules.RST; <i>package RoughSets</i> em R	56,4	61,5	57,7	63,5
OneR	OneR; Weka	54,3	58,1	54,4	58,2
PART	PART; Weka	71,6	71,9	71,6	71,9
VC-DomLEM	VC-DomLEM; jMAF	54,3	58,2	78,7	80,5
ZeroR	ZeroR; Weka	26,9	26,3	26,9	26,3

Fonte: Autor; adaptado de Han *et al.* (2012).

Para os algoritmos classificadores AQ, CN2, *HybridFS*, *Indiscernibility Based Rules* e LEM2, foi utilizado o método *Holdout* (Han *et al.*, 2012) para aferição, considerando-se os conjuntos de treinamento e teste na proporção de 2/3 e 1/3 do total de registros, respectivamente. O conjunto de teste foi constituído de forma aleatória: 117 registros de um total de 342, para cada conjunto dirSU e indSU. Para os demais algoritmos classificadores, *DecisionTable*, JRip, *OneR*, PART, VC-DomLEM e *ZeroR*, foi utilizado o método de aferição *Stratified k-fold cross-validation*,  $k = 10$  (Han *et al.*, 2012): os registros são particionados aleatoriamente em  $k$  subconjuntos mutuamente exclusivos,  $D_1, D_2, \dots, D_k$ , cada um de tamanho aproximadamente igual e mantida aproximadamente a mesma distribuição de classes dos dados iniciais. Os registros de treinamento e de teste são executados  $k$  vezes. Na iteração  $i$ , a partição  $D_i$  é reservada como conjunto de teste e as partições restantes são usadas coletivamente para treinar o modelo, considerando-se os conjuntos dirSU e indSU.

Para a escolha dos algoritmos classificadores para compor o método *ensemble bagging* para a predição de classes, adotou-se o critério de sensibilidade média superior a 60% (sessenta por cento), para ambos os conjuntos dirSU e indSU. Ao final, tem-se a(s) classe(s) predita(s) mais votada(s) ou de maior frequência, conforme ilustrado na Figura 31.

Figura 31– Método *ensemble bagging* e algoritmos classificadores para predição de classe (votada)



Fonte: Autor; adaptado de Han *et al.* (2012).

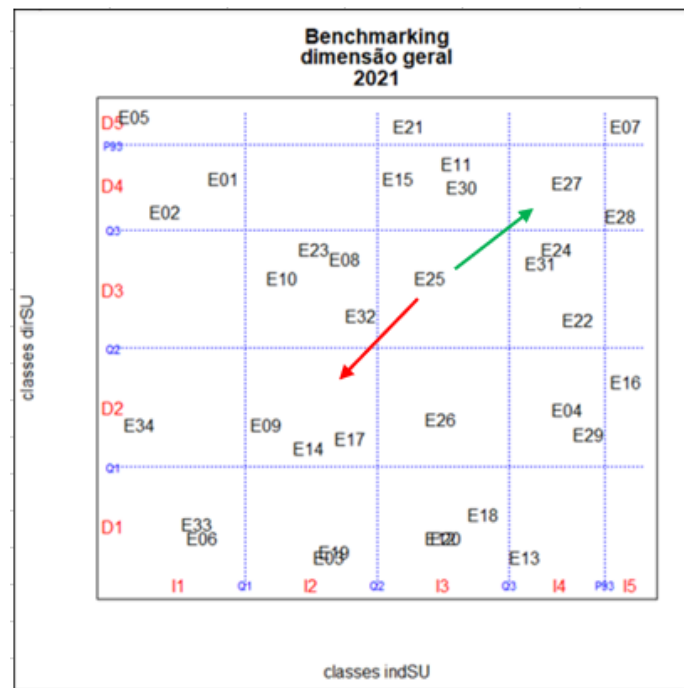
Ademais, a utilização de mais de um algoritmo de classificação contribui para aumentar a probabilidade de se obter uma predição correta. Se for adotada uma sensibilidade de 60% como probabilidade de sucesso (P), se cada algoritmo ao predizer

uma classe for considerado um evento independente e se dois ou mais eventos independentes podem ocorrer simultaneamente, a probabilidade de sucesso (P) de ou um ou outro, ou ambos, pode ser obtida da seguinte forma: a) P (predição correta de 1 ou 2 algoritmos) =  $0,6 + 0,6 - (0,6 \times 0,6) = 84,0\%$ ; b) P (predição correta de 1, 2 ou 3 algoritmos) =  $0,6 + 0,6 + 0,6 - 3 \times (0,6 \times 0,6) + (0,6 \times 0,6 \times 0,6) = 93,6\%$ .

Sob uma abordagem heurística e, tendo empresas como alvo, E27 e E17, a ascensão ou descensão de classes para a empresa E25, em 2021, situada no setor (D3, I3) na Matriz de Sustentabilidade conforme Figura 32, pode ser simulada por meio de método *ensemble bagging* e algoritmos classificadores, comparando-se os seus indicadores com aqueles das empresas E27 e E17. No caso da ascensão equiparada à empresa E27, partiu-se dos maiores ou menores valores para indicadores de ganho (G) ou custo (C), respectivamente, entre as empresas E25 e E27, conforme Quadros 28 e 29, setas verdes. No caso da descensão equiparada à empresa E17, partiu-se dos menores ou maiores valores para indicadores ganho ou custo, respectivamente, entre as empresas E25 e E17, conforme Quadros 28 e 29, setas vermelhas. Os valores hachurados em verde utilizados na simulação visam a ascensão de classe; os valores hachurados em vermelho visam a descensão de classe. Com base em dados históricos dos indicadores, período 2013-2021, e valores simulados, obteve-se as classes preditas pelos algoritmos CN2, DecisionTable, JRip, PART e VC-DomLEM, com a escolha da(s) classes(s) de maior frequência ou votada(s). O Anexo VIII reporta o procedimento em R para a predição de classe com o uso do algoritmo classificador CN2.

Desta forma, permite-se uma efetiva gestão de riscos quanto à descensão de classes, bem como, à ascensão como melhor aproveitamento de oportunidades. Isso se constitui em ferramenta para se investigar indicadores mais sensíveis a mudanças de classes, além da possibilidade de se conhecer valores de referência de indicadores a se atingir ou se distanciar, conforme o caso.

Figura 32– Matriz de Sustentabilidade visando simular a ascensão ou descensão de classes, empresa E25, ano 2021, método *ensemble bagging* e algoritmos classificadores



Fonte: Autor.

Quadro 28– Simulação da ascensão (setas verdes) e descensão (setas vermelhas) de classe dirSU, empresa E25, classe D3, método *ensemble bagging* e algoritmos classificadores

	IE01	IE02	IE03	IE04	IE05	IE06	IE07	IE08	IE09	IE10	IS01	IS02	IS03	IS04	IS05	IS06	IS07	IS08	IS09	IS10
	G	C	C	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	C
E27 ↑	17580637	4773	842	844	31349	103814	378	328	204	255	8870	39	24	2.93	4	115	12092	8457	440	25.61
E25 ↑	2629206	4832	1599	143	3105.85	105778	378	339	197	245	1340	45.3	25.9	33.06	1.7	213.43	15068	13942	1473	15.74
E25 ↓	2629206	4832	1599	143	3105.85	105778	378	339	197	245	1340	45.3	25.9	33.06	1.7	213.43	15068	13942	1473	15.74
E17 ↓	645131	4620.8	1225.73	56	1294	157506	305.39	337.06	154.15	221.65	937	41.71	13.11	17.77	3.7	723	14837	16472	1014	39.09

	IA01	IA02	IA03	IG01	IG02	IG03	IG04	IG05	IG06	IG07	IG08	IG09	IG10						
	C	C	C	G	G	G	G	G	G	G	G	G	G	CN2	DecisionTable	JRip	PART	VC-DomLEM	CLASSE(S) VOTADA(S)
E27 ↑	21852993	15022766	129627.9	1	0	0	1	0	1	0	0	0	0						
E25 ↑	328797.2	319904.2	34449	1	0	0	0	0	1	0	1	0	0	D4	D1	D2	D3	D4	D4
E25 ↓	328797.2	319904.2	34449	1	0	0	0	0	1	0	1	0	0	D1	D3	D1	D1	D2	D1
E17 ↓	96732.36	56342.79	13998	1	0	0	0	0	1	0	0	0	0						

Fonte: Autor.

Quadro 29– Simulação da ascensão (setas verdes) e descensão (setas vermelhas) de classe indSU, empresa E25, classe I3, método *ensemble bagging* e algoritmos classificadores

	IE01	IE02	IE03	IE04	IE05	IS01	IS02	IS03	IS04	IS05	IS06	IS07	IS08	IS09	IS10	IA01	IA02
	G	G	G	G	G	C	C	G	C	C	C	C	C	G	C	C	G
E27 ↑	511	532	633374	7.88	1982.03	80	1565051	78	2	0.01	192	150	7	440432	14867	16	1206
E25 ↑	217	70	140734	5223	1962.27	55	1970592	74	1.69	0.01	152	155.59	5.56	777089	37853	28.34	3120
E25 ↓	217	70	140734	5223	1962.27	55	1970592	74	1.69	0.01	152	155.59	5.56	777089	37853	28.34	3120
E17 ↓	52	56	112116	3157	685	47.23	1044281	97	1.7	0.01	178	274.92	14.19	259237	11818	26.83	1124

	IG01	IG02	IG03	IG04	IG05	IG06	IG07	IG08	IG09	IG10	IG11	IG12						
	G	G	G	C	G	G	G	G	G	G	G	G	CN2	DecisionTable	JRip	PART	VC-DomLEM	CLASSE(S) VOTADA(S)
E27 ↑	0	0	0	0	0	1	0	0	1	0	0	1						
E25 ↑	0	0	0	0	0	1	0	0	1	0	1	1	I4	I1	I4	I4	I4	I4
E25 ↓	0	0	0	0	0	1	0	0	0	1	0	1	I2	I2	I2	I2	I2	I2
E17 ↓	1	0	0	0	0	1	0	0	1	1	0	0						

Fonte: Autor.

Os Quadros 28 e 29 ilustram para cada indicador, o quanto a empresa E25 precisa aumentar, diminuir ou mesmo manter, em termos absolutos, para ascender ou descender de classe, sob a perspectiva de variações concomitantes. Isso se constitui em ferramenta de efetiva gestão de riscos.

Com referência ao Quadro 28, a Figura 33 ilustra como exemplo, a classe predita D4 pelo algoritmo VC-DomLEM, *software* jMAF, para a empresa E25, equiparada à empresa E27, utilizando-se dos dados históricos dos indicadores dirSU no período 2013-2021, e dos valores simulados.

Figura 33– Predição da classe D4 pelo algoritmo VC-DomLEM para a empresa E25, com base em dados históricos de indicadores dirSU e valores simulados

Example name	Original decision	Classification result	Certainty	No. of matching rules	codigo_empresa	ano	IE01	IE02	IE03
Example_1	D3	D4	1.000	13	E25	2021	17580637.000	4773.000	842.000

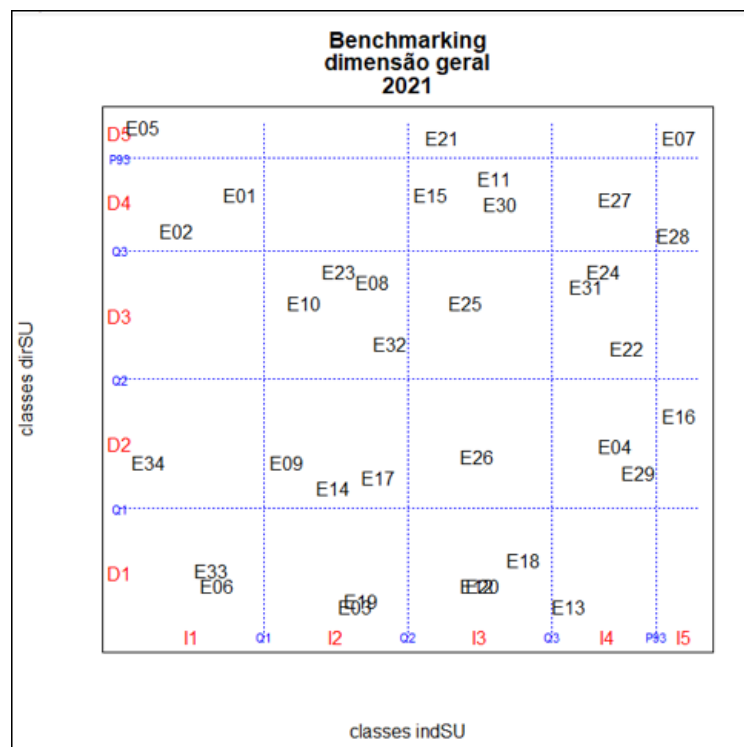
Tested Examples: 1  
 Correct: 0 (0.000%)  
 Incorrect: 0 (0.000%)  
 Unclassified: 1 (100.000%)

Fonte: Autor.

#### 5.4. MATRIZ DE SUSTENTABILIDADE

Com base nas classificações das empresas obtidas na etapa 5.2, foram gerados os gráficos de *benchmarking* e evolução de empresa, com o apoio do *software* RStudio, v. 2023.06.1. Os resultados obtidos estão ilustrados nas Figuras 34 a 47. Neste caso, considerando a dimensão geral, ano 2021, a empresa E07 desponta-se como um *benchmark* ou referência no setor (D5, I5), conforme ilustrado na Figura 34.

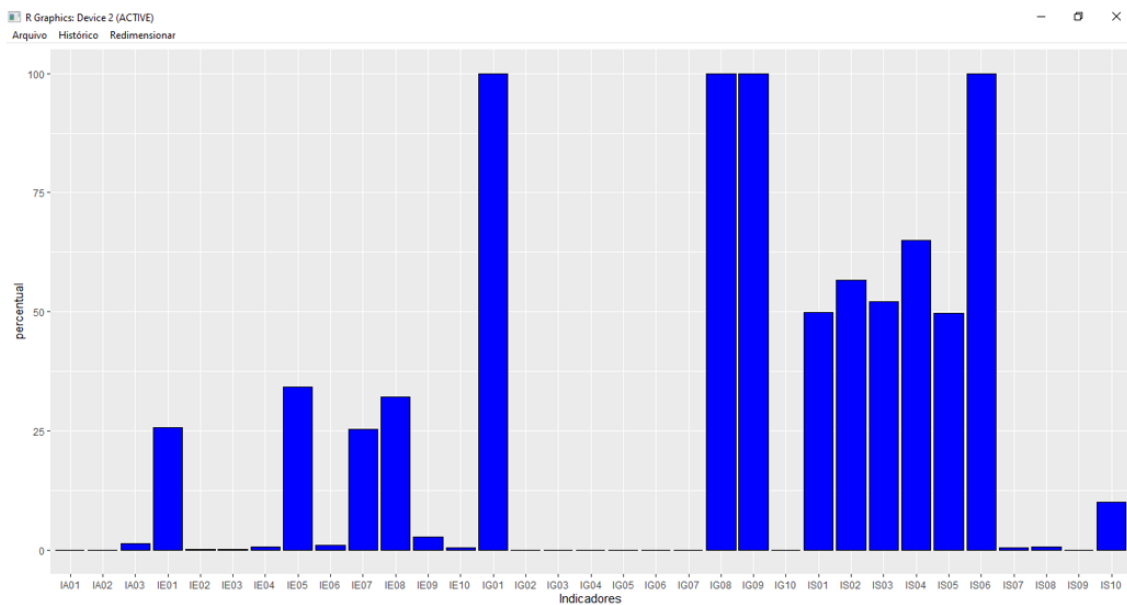
Figura 34– Resultado do *benchmarking*, dimensão geral, ano 2021; empresa E07 como benchmark



Fonte: Autor.

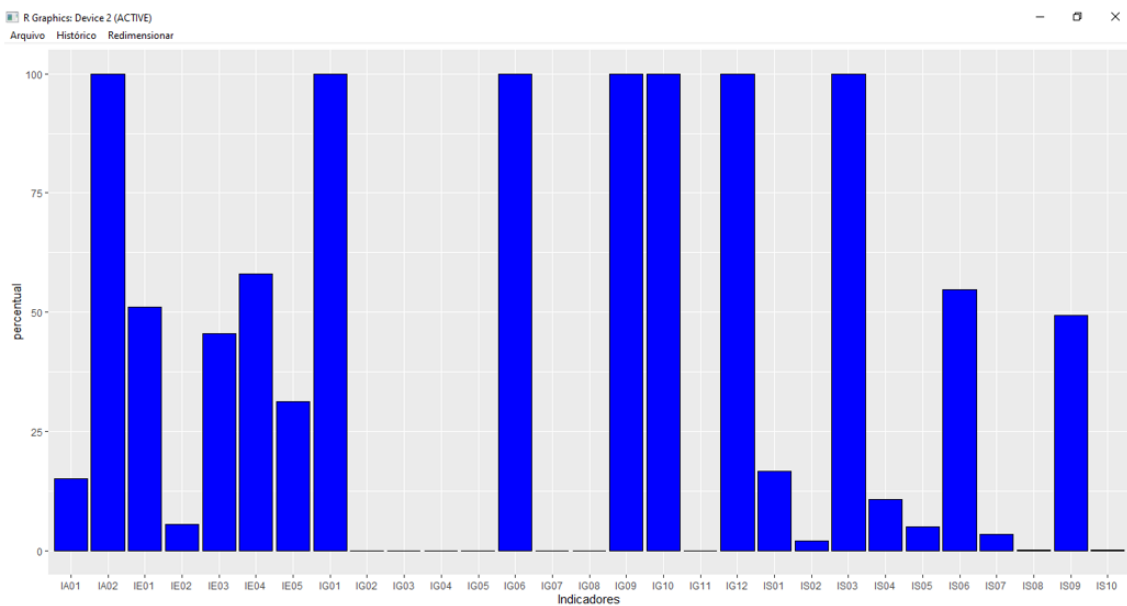


Figura 35– Posição relativa dos indicadores dirSU, empresa E07, ano 2021



Fonte: Autor.

Figura 36– Posição relativa dos indicadores indSU, empresa E07, ano 2021

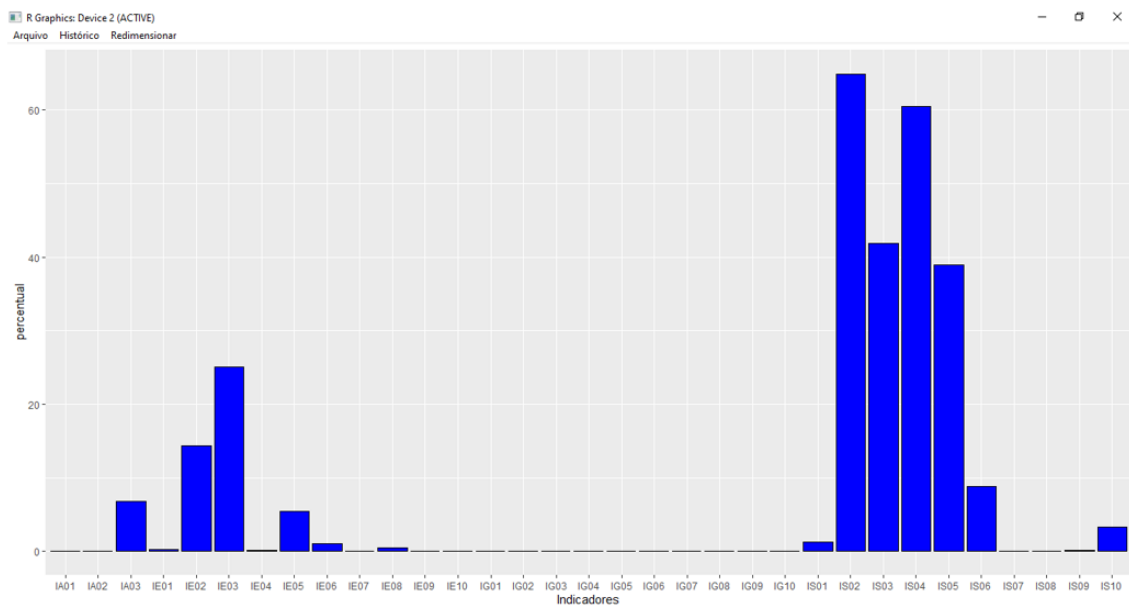


Fonte: Autor.

Neste caso, como ilustram as Figuras 35 e 36, a empresa E07 atingiu posições relativas de 100% em vários indicadores.

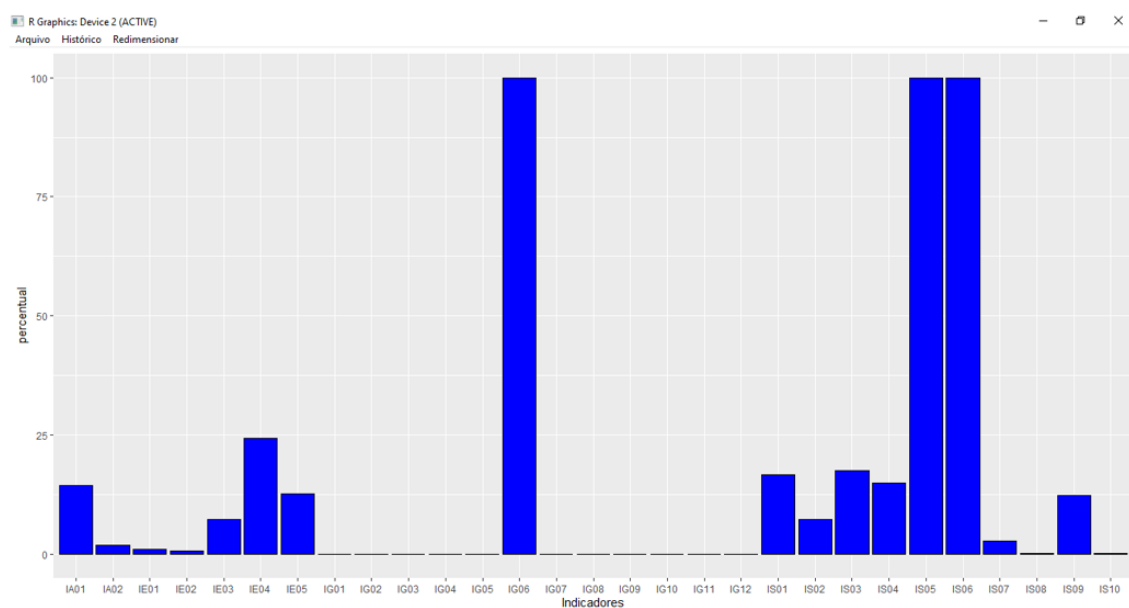
Para a empresa E33, situada em (D1, I1), essa obteve 100% somente em três indicadores indSU, conforme Figura 38.

Figura 37– Posição relativa dos indicadores dirSU, empresa E33, ano 2021



Fonte: Autor.

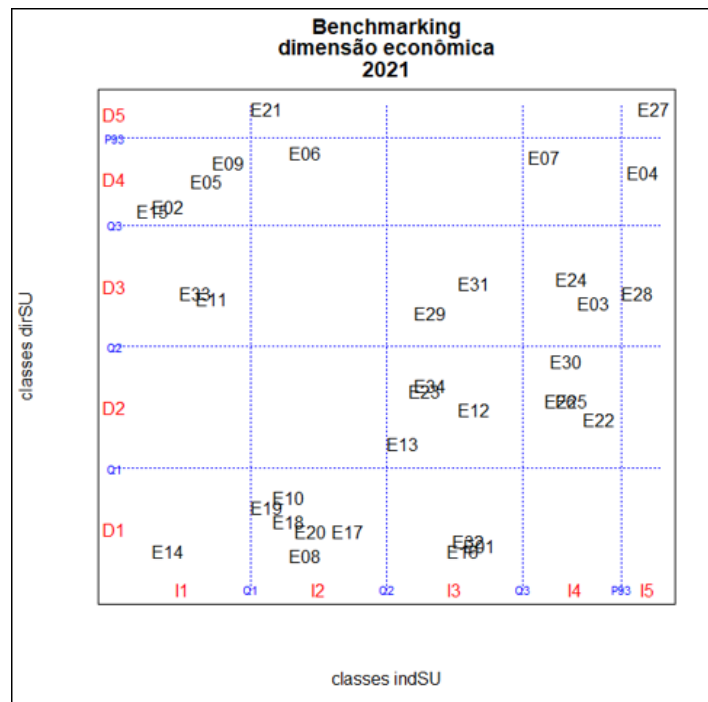
Figura 38– Posição relativa dos indicadores indSU, empresa E33, ano 2021



Fonte: Autor.

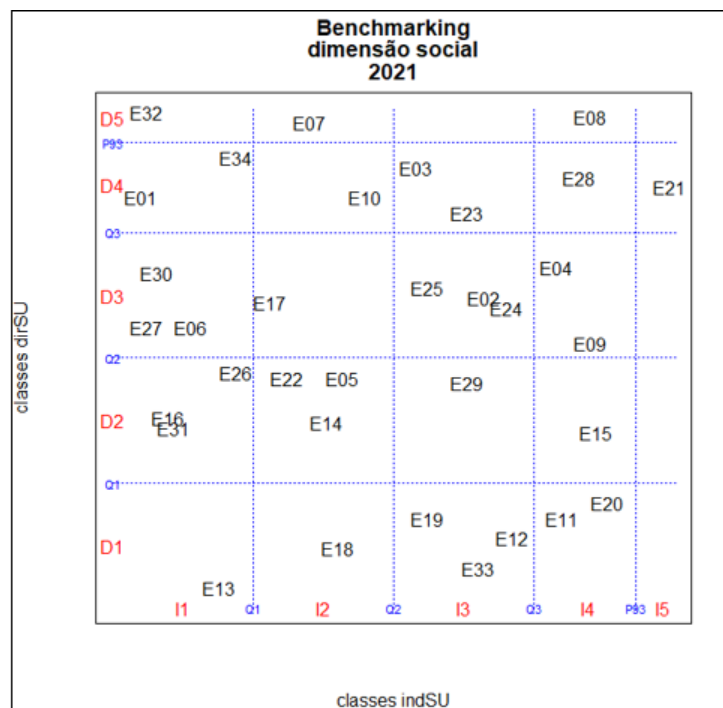
As Figuras 39 a 42 reportam *benchmarks* nas dimensões econômica, social, ambiental e de governança corporativa, respectivamente, para o período de 2021.

Figura 39– Resultado do *benchmarking*, dimensão econômica, ano 2021



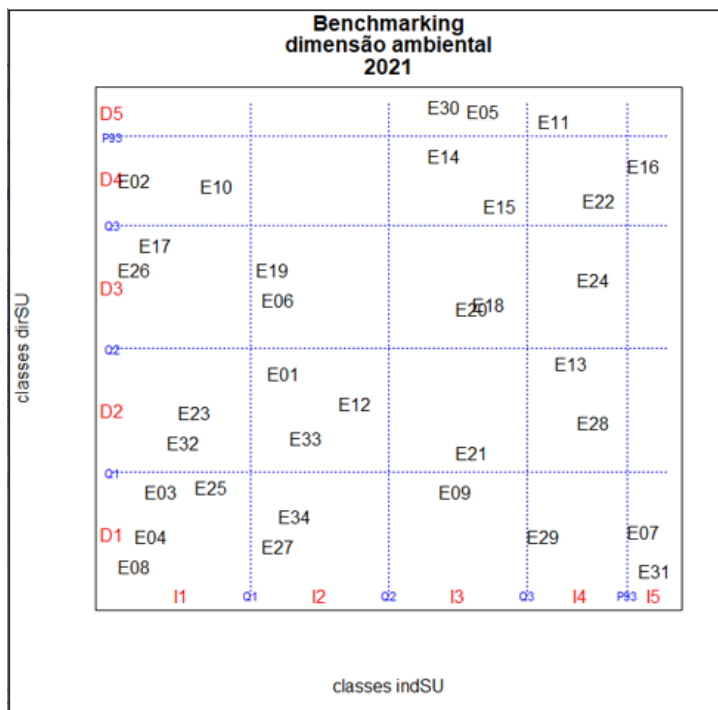
Fonte: Autor.

Figura 40– Resultado do *benchmarking*, dimensão social, ano 2021



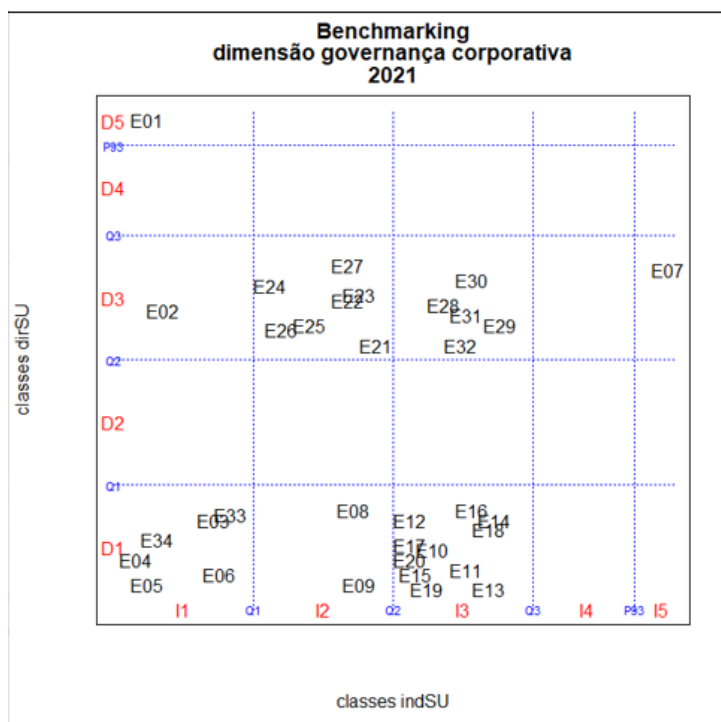
Fonte: Autor.

Figura 41– Resultado do *benchmarking*, dimensão ambiental, ano 2021



Fonte: Autor.

Figura 42– Resultado do *benchmarking*, dimensão governança corporativa, ano 2021



Fonte: Autor.

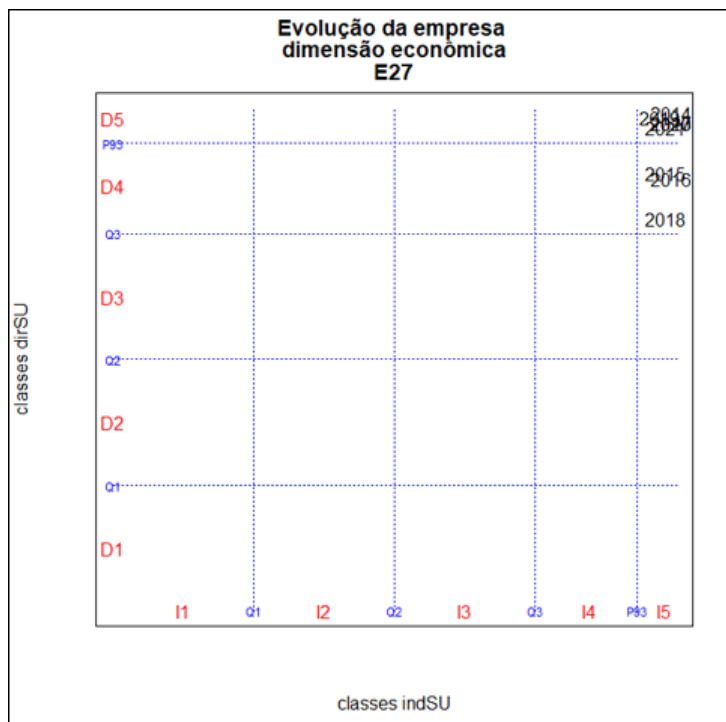
As Figuras 43 a 47 ilustram a evolução da empresa E27 ao longo do período 2013-2021, nas dimensões geral, econômica, social, ambiental e de governança corporativa, respectivamente.

Figura 43– Resultado da evolução da empresa E27, dimensão geral, período 2013-2021



Fonte: Autor.

Figura 44– Resultado da evolução da empresa E27, dimensão econômica, período 2013-2021



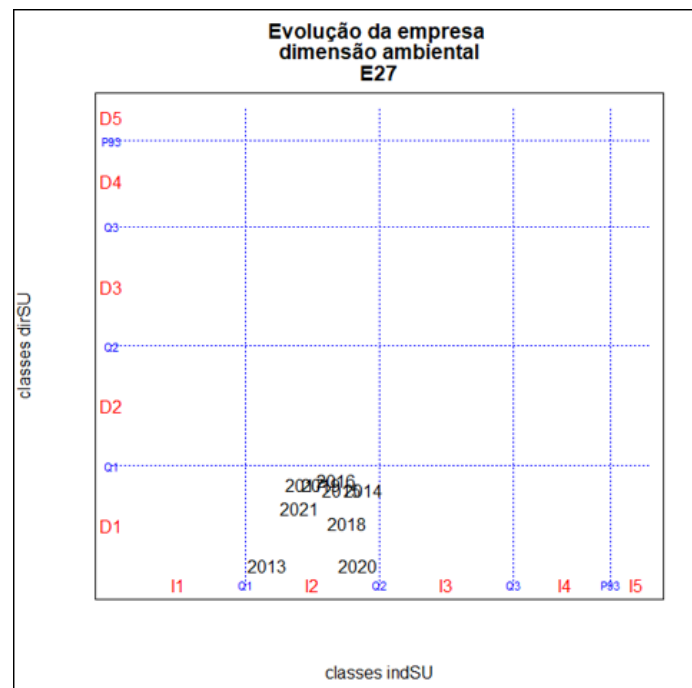
Fonte: Autor.

Figura 45– Resultado da evolução da empresa E27, dimensão social, período 2013-2021



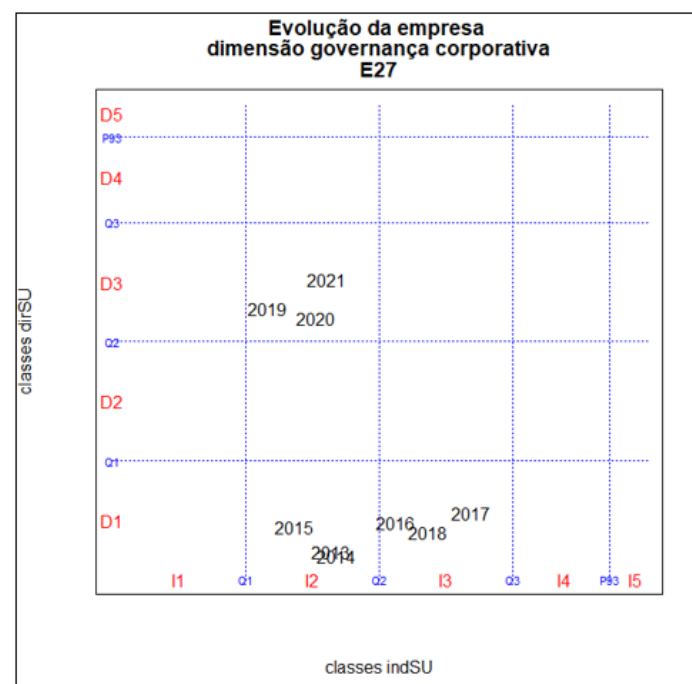
Fonte: Autor.

Figura 46– Resultado da evolução da empresa E27, dimensão ambiental, período 2013-2021



Fonte: Autor.

Figura 47– Resultado da evolução da empresa E27, dimensão governança corporativa, período 2013-2021



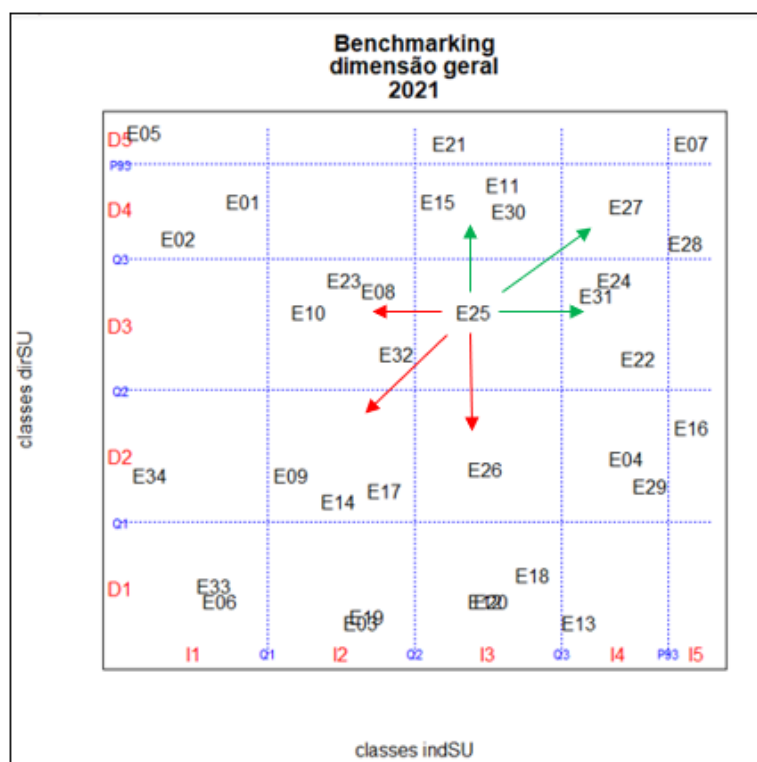
Fonte: Autor.

#### 5.4.1. ANÁLISE DE SENSIBILIDADE

Visando obter uma análise de sensibilidade para verificar de que forma as variações introduzidas nos indicadores influenciam os resultados obtidos (GOMES; GOMES, 2019), procedeu-se à simulação da variação de +/- 1,0% para cada indicador, exclusivamente. Assim, a depender da classe a prospectar ou de destino, elaborou-se um algoritmo em VBA/Microsoft Excel para simular um acréscimo ou decréscimo do valor do indicador sob análise, com o objetivo de identificar o risco, como percentual, na ascensão ou descensão de classe.

A Figura 48 ilustra movimentos de migração de classe para a empresa E25, ano de 2021, considerando-se as possibilidades de migrar para classes superiores (I4, D4; setas em verde) ou inferiores (I2, D2; setas em vermelho). A variação total para cada indicador foi limitada em 100% (cem por cento).

Figura 48– Movimentos de migração de classes para a empresa E25, ano de 2021



Fonte: Autor.

Assim, pelos Quadros 30 e 31, identifica-se que para ascender para a classe D4 ou I4, no ano de 2021, a empresa E25 pode simular a variação necessária em cada indicador, considerando-se se o indicador é de ganho (quanto maior, melhor) ou de custo (quanto menor, melhor). O valor de referência diz respeito ao maior valor entre as empresas,



quando indicador de ganho; e, ao menor valor, quando indicador de custo. O % do valor de referência diz da relação entre os valores final e de referência.

Os Quadros 32 e 33 ilustram simulações com os indicadores para descender para a classe D2 ou I2, empresa E25, ano 2021.

No caso de ascensão de classe, por exemplo, de D3 para D4, identifica-se que a menor variação seria de 88%, para menor, no caso do indicador IA03, de custo: Consumo total de água (m<sup>3</sup>), fonte rede pública. Ou seja, haveria um esforço grande para se atingir o valor final. Neste caso, a adoção de práticas de governança corporativa (7) pode melhor contribuir para a migração de classe. Para a ascensão da classe I3 para I4, a menor variação é de 69%, para menor, para o indicador IA01, de custo: Consumo de água (m<sup>3</sup>) por empregado.

Para o caso inverso, de descensão de classe, de D3 para D2, é necessária uma variação de 33%, para menor, para o indicador IS04, de ganho: % mulheres sobre o total de empregados; ou, descartar alguma prática de governança corporativa apontada. Para descender da classe I3 para I2, é necessária uma variação mínima de 77%, para menor, para o indicador IE05, de ganho: Número de consumidores por empregado.

Conseqüentemente, pelas avaliações anteriores, a possibilidade de descensão de classes torna-se mais exequível frente ao esforço necessário para a ascensão.

Quadro 30– Simulação de ascensão de classe de D3 para D4, empresa E25, ano 2021

ANÁLISE DE SENSIBILIDADE - INDICADORES dirSU										
EMPRESA: E25			ANO: 2021		CLASSE ATUAL: D3		CLASSE A PROSPECTAR: D4			
IND.	DESCRIÇÃO RESUMIDA	GRI/ ANEEL (1)	UN.	GANHO/ CUSTO	CLASSE FINAL	VALOR INICIAL	VALOR FINAL	VARIAÇÃO (%)	VALOR REFERÊNCIA	% VALOR REFERÊNCIA
IE01	Número de consumidores atendidos	EU3	un	G						
IE02	Energia comprada	EU10	GWh	C						
IE03	Perdas elétricas globais	EU12	GWh	C						
IE04	Número de subestações	EU1	un	G						
IE05	Capacidade instalada	EU1	MVA	G						
IE06	Folha de pagamento bruta	EC1	R\$ mil	C	D4	105778	6346.68	94	5619.61	112.94
IE07	Energia vendida, setor Industrial	EU3	GWh	G						
IE08	Energia vendida, setor Rural	EU3	GWh	G						
IE09	Energia vendida, Iluminação pública	EU3	GWh	G						
IE10	Energia vendida, Poder público	EU3	GWh	G						
IS01	Número de empregados próprios	LA1	un	G						
IS02	Empregados com idade entre 31 e 40 anos (%)	LA13	%	G						
IS03	Empregados com idade entre 41 e 50 anos (%)	LA13	%	G						
IS04	% mulheres/total de empregados	LA13	%	G						
IS05	% empregados programa contratação aprendizes	LA1	%	G						
IS06	Benefícios com Educação	LA3	R\$ mil	G						
IS07	Benefícios com Alimentação	LA3	R\$ mil	G						
IS08	Benefícios com Saúde	LA3	R\$ mil	G						
IS09	Benefícios com Auxílio-creche	LA3	R\$ mil	G						
IS10	Taxa de rotatividade	LA2	%	C	D4	15.74	0.31	98	0.30	105.82
IA01	Consumo de energia direta, fonte Diesel	EN3	GJ	C	D4	328797.19	0.00	100	0.17	0.00
IA02	Consumo de energia direta, fonte Gasolina	EN3	GJ	C	D4	319904.17	0.00	100	0.11	0.00
IA03	Consumo total de água, fonte Rede pública	EN3	m3	C	D4	34449	4133.88	88	1078.00	383.48
IG01	Estrutura de governança da organização	4.1	un	G						
IG02	Presidente órgão de governança é ou não diretor executivo	4.2	un	G	D4	0	1			
IG03	Membros da governança independentes	4.3	un	G	D4	0	1			
IG04	Mecanismos para recomendações para governança	4.4	un	G	D4	0	1			
IG05	Relação remuneração governança e desempenho	4.5	un	G	D4	0	1			
IG06	Garantia para evitar conflitos interesse	4.6	un	G						
IG07	Qualificações para membros da governança	4.7	un	G	D4	0	1			
IG08	Declarações de missão e princípios conduta	4.8	un	G						
IG09	Procedimentos para supervisionar a organização	4.9	un	G	D4	0	1			
IG10	Processos para avaliar a governança	4.10	un	G	D4	0	1			

1- ANEEL (2015); GRI (2000)

Fonte: Autor.

Quadro 31– Simulação de ascensão de classe de I3 para I4, empresa E25, ano 2021

ANÁLISE DE SENSIBILIDADE - INDICADORES indSU									
EMPRESA: E25			ANO: 2021		CLASSE ATUAL: I3		CLASSE A PROSPECTAR: I4		
IND.	DESCRIÇÃO RESUMIDA	UN.	GANHO/ CUSTO	CLASSE FINAL	VALOR INICIAL	VALOR FINAL	VARIÇÃO (%)	VALOR REFERÊNCIA	% VALOR REFERÊNCIA
IE01	Número de localidades atendidas (municípios)	un	G						
IE02	Número de escritórios comerciais	un	G						
IE03	Número de transformadores de distribuição	un	G						
IE04	Energia vendida por empregado	MWh	G	I4	5223	8879.10	70	6634.00	133.84
IE05	Número de consumidores por empregado	un	G	I4	1962.27	3630.20	85	2342.49	154.97
IS01	Média de horas extras por empregado/ano	h	C						
IS02	Número de chamadas recebidas	un	C	I4	1970592	98529.60	95	115576.00	85.25
IS03	Número médio de atendentes	un	G						
IS04	% IAb - Índice de Abandono	%	C	I4	1.69	0.12	93	0.15	78.87
IS05	% ICO - Índice de Chamadas Ocupadas	%	C	I4	0.01	0.00	93	0.01	7.00
IS06	TMA - Tempo Médio de Atendimento (s)	s	C	I4	152	6.08	96	117.80	5.16
IS07	DER (horas)	h	C	I4	155.59	3.11	98	3.81	81.67
IS08	FER (unid.)	un	C						
IS09	Número de atendimentos realizados	un	G						
IS10	Número de atendimentos realizados fora do prazo	un	C						
IA01	Consumo de água por empregado (em m3)	m3	C	I4	28.34	8.79	69	2.46	357.13
IA02	Número de alunos atendidos	un	G						
IG01	Emissão de ações preferenciais	un	G						
IG02	Acionistas preferenciais têm direito a voto	un	G						
IG03	Mecanismos de divulgação sobre temas deliberados	un	G						
IG04	Processos administrativos, arbitrais ou judiciais contra a companhia	un	C						
IG05	Proibição de empréstimos e garantias em favor do controlador	un	G						
IG06	Existência de um canal para comunicações anônimas	un	G						
IG07	Relatório anual e/ou de sustentabilidade	un	G						
IG08	Programas para educação sobre sustentabilidade	un	G						
IG09	Aderência de compromissos voluntários	un	G						
IG10	Existência de comitê de sustentabilidade	un	G						
IG11	Parecer de auditoria independente	un	G						
IG12	Compromisso com o combate à corrupção	un	G						

Fonte: Autor.

Quadro 32– Simulação de descensão de classe de D3 para D2, empresa E25, ano 2021

ANÁLISE DE SENSIBILIDADE - INDICADORES dirSU										
EMPRESA: E25			ANO: 2021		CLASSE ATUAL: D3		CLASSE A PROSPECTAR: D2			
IND.	DESCRIÇÃO RESUMIDA	GRI/ ANEEL (1)	UN.	GANHO/ CUSTO	CLASSE FINAL	VALOR INICIAL	VALOR FINAL	VARIAÇÃO (%)	VALOR REFERÊNCIA	% VALOR REFERÊNCIA
IE01	Número de consumidores atendidos	EU3	un	G						
IE02	Energia comprada	EU10	GWh	C						
IE03	Perdas elétricas globais	EU12	GWh	C						
IE04	Número de subestações	EU1	un	G						
IE05	Capacidade instalada	EU1	MVA	G						
IE06	Folha de pagamento bruta	EC1	R\$ mil	C						
IE07	Energia vendida, setor Industrial	EU3	GWh	G						
IE08	Energia vendida, setor Rural	EU3	GWh	G						
IE09	Energia vendida, Iluminação pública	EU3	GWh	G						
IE10	Energia vendida, Poder público	EU3	GWh	G						
IS01	Número de empregados próprios	LA1	un	G						
IS02	Empregados com idade entre 31 e 40 anos (%)	LA13	%	G	D2	45.3	24.01	47	66.23	36.25
IS03	Empregados com idade entre 41 e 50 anos (%)	LA13	%	G	D2	25.9	8.29	68	55.00	15.07
IS04	% mulheres/total de empregados	LA13	%	G	D2	33.06	22.15	33	33.06	67.00
IS05	% empregados programa contratação aprendizes	LA1	%	G						
IS06	Benefícios com Educação	LA3	R\$ mil	G						
IS07	Benefícios com Alimentação	LA3	R\$ mil	G						
IS08	Benefícios com Saúde	LA3	R\$ mil	G						
IS09	Benefícios com Auxílio-creche	LA3	R\$ mil	G						
IS10	Taxa de rotatividade	LA2	%	C						
IA01	Consumo de energia direta, fonte Diesel	EN3	GJ	C						
IA02	Consumo de energia direta, fonte Gasolina	EN3	GJ	C						
IA03	Consumo total de água, fonte Rede pública	EN3	m3	C						
IG01	Estrutura de governança da organização	4.1	un	G	D2	1	0			
IG02	Presidente órgão de governança é ou não diretor executivo	4.2	un	G						
IG03	Membros da governança independentes	4.3	un	G						
IG04	Mecanismos para recomendações para governança	4.4	un	G						
IG05	Relação remuneração governança e desempenho	4.5	un	G						
IG06	Garantia para evitar conflitos interesse	4.6	un	G	D2	1	0			
IG07	Qualificações para membros da governança	4.7	un	G						
IG08	Declarações de missão e princípios conduta	4.8	un	G	D2	1	0			
IG09	Procedimentos para supervisionar a organização	4.9	un	G						
IG10	Processos para avaliar a governança	4.10	un	G						

1- ANEEL (2015); GRI (2000)

Fonte: Autor.

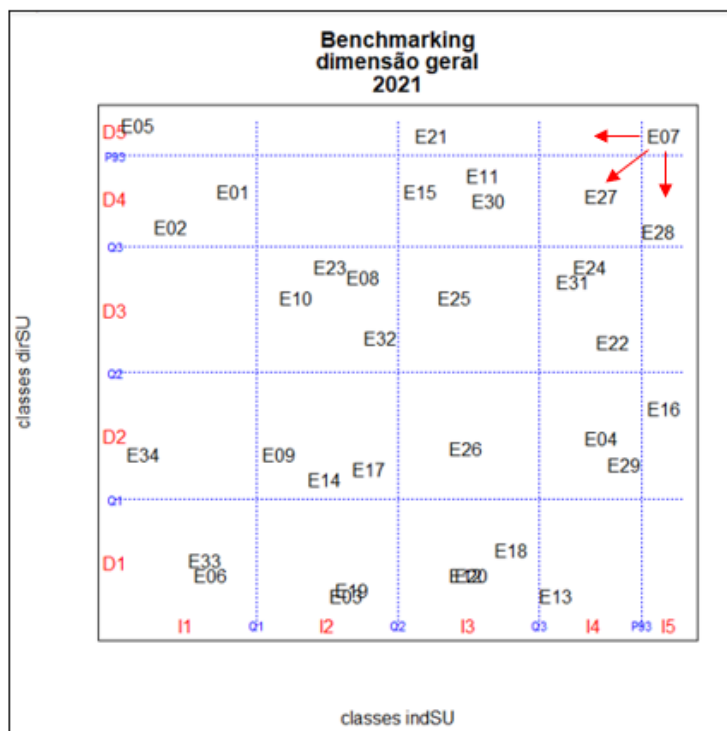
Quadro 33– Simulação de descensão de classe de I3 para I2, empresa E25, ano 2021

ANÁLISE DE SENSIBILIDADE - INDICADORES indSU									
EMPRESA: E25			ANO: 2021		CLASSE ATUAL: I3		CLASSE A PROSPECTAR: I2		
IND.	DESCRIÇÃO RESUMIDA	UN.	GANHO/ CUSTO	CLASSE FINAL	VALOR INICIAL	VALOR FINAL	VARIAÇÃO (%)	VALOR REFERÊNCIA	% VALOR REFERÊNCIA
IE01	Número de localidades atendidas (municípios)	un	G						
IE02	Número de escritórios comerciais	un	G						
IE03	Número de transformadores de distribuição	un	G						
IE04	Energia vendida por empregado	MWh	G	I2	5223	940.14	82	6634.00	14.17
IE05	Número de consumidores por empregado	un	G	I2	1962.27	451.32	77	2342.49	19.27
IS01	Média de horas extras por empregado/ano	h	C						
IS02	Número de chamadas recebidas	un	C						
IS03	Número médio de atendentes	un	G						
IS04	% IAb - Índice de Abandono	%	C						
IS05	% ICO - Índice de Chamadas Ocupadas	%	C						
IS06	TMA - Tempo Médio de Atendimento (s)	s	C						
IS07	DER (horas)	h	C						
IS08	FER (unid.)	un	C						
IS09	Número de atendimentos realizados	un	G						
IS10	Número de atendimentos realizados fora do prazo	un	C						
IA01	Consumo de água por empregado (em m <sup>3</sup> )	m <sup>3</sup>	C						
IA02	Número de alunos atendidos	un	G						
IG01	Emissão de ações preferenciais	un	G						
IG02	Acionistas preferenciais têm direito a voto	un	G						
IG03	Mecanismos de divulgação sobre temas deliberados	un	G						
IG04	Processos administrativos, arbitrais ou judiciais contra a companhia	un	C						
IG05	Proibição de empréstimos e garantias em favor do controlador	un	G						
IG06	Existência de um canal para comunicações anônimas	un	G						
IG07	Relatório anual e/ou de sustentabilidade	un	G						
IG08	Programas para educação sobre sustentabilidade	un	G						
IG09	Aderência de compromissos voluntários	un	G						
IG10	Existência de comitê de sustentabilidade	un	G						
IG11	Parecer de auditoria independente	un	G						
IG12	Compromisso com o combate à corrupção	un	G						

Fonte: Autor.

Para a empresa E07, simulou-se a condição de descensão de classes, para I4 ou D4 (setas em vermelho), ano 2021, conforme Figura 49.

Figura 49– Simulação de descensão de classes, empresa E07, ano 2021



Fonte: Autor.

Os Quadros 34 e 35 ilustram a simulação por indicador, da descensão de classe para I4 ou D4, ano de 2021.

Neste caso, é necessária uma variação mínima de 15%, para menor, do indicador IS04, de ganho: % mulheres / total empregados, para se migrar da classe D5 para D4; ou, deixar de adotar alguma prática de governança corporativa apontada. No caso de migração de I5 para I4, há somente uma variação de 59%, para menor, para o indicador IA02, de ganho: Número de alunos atendidos.

Consequentemente, a análise de sensibilidade aponta que a migração de classe de D5 para D4 torna-se mais factível frente à migração de I5 para I4.

O Anexo IX reporta o procedimento em VBA para análise da sensibilidade.

Quadro 34– Simulação de descensão de classe de D5 para D4, empresa E07, ano 2021

ANÁLISE DE SENSIBILIDADE - INDICADORES dirSU										
EMPRESA: E07			ANO: 2021		CLASSE ATUAL: D5		CLASSE A PROSPECTAR: D4			
IND.	DESCRIÇÃO RESUMIDA	GRI/ ANEEL <sup>(1)</sup>	UN.	GANHO/ CUSTO	CLASSE FINAL	VALOR INICIAL	VALOR FINAL	VARIAÇÃO (%)	VALOR REFERÊNCIA	% VALOR REFERÊNCIA
IE01	Número de consumidores atendidos	EU3	un	G	D4	4519630	2892563.20	36	17580637.00	16.45
IE02	Energia comprada	EU10	GWh	C						
IE03	Perdas elétricas globais	EU12	GWh	C						
IE04	Número de subestações	EU1	un	G						
IE05	Capacidade instalada	EU1	MVA	G	D4	10742.3	7841.88	27	31349.00	25.01
IE06	Folha de pagamento bruta	EC1	R\$ mil	C						
IE07	Energia vendida, setor Industrial	EU3	GWh	G	D4	71651	45856.64	36	282062.00	16.26
IE08	Energia vendida, setor Rural	EU3	GWh	G	D4	2562	1819.02	29	7962.00	22.85
IE09	Energia vendida, Iluminação pública	EU3	GWh	G						
IE10	Energia vendida, Poder público	EU3	GWh	G						
IS01	Número de empregados próprios	LA1	un	G	D4	4430	3588.30	19	8870.00	40.45
IS02	Empregados com idade entre 31 e 40 anos (%)	LA13	%	G	D4	37.56	31.17	17	66.23	47.07
IS03	Empregados com idade entre 41 e 50 anos (%)	LA13	%	G	D4	28.65	23.49	18	55.00	42.71
IS04	% mulheres/total de empregados	LA13	%	G	D4	21.51	18.28	15	33.06	55.30
IS05	% empregados programa contratação aprendizes	LA1	%	G	D4	3.83	3.10	19	7.70	40.29
IS06	Benefícios com Educação	LA3	R\$ mil	G	D4	1344	913.92	32	1344.00	68.00
IS07	Benefícios com Alimentação	LA3	R\$ mil	G						
IS08	Benefícios com Saúde	LA3	R\$ mil	G						
IS09	Benefícios com Auxílio-creche	LA3	R\$ mil	G						
IS10	Taxa de rotatividade	LA2	%	C						
IA01	Consumo de energia direta, fonte Diesel	EN3	GJ	C						
IA02	Consumo de energia direta, fonte Gasolina	EN3	GJ	C						
IA03	Consumo total de água, fonte Rede pública	EN3	m3	C						
IG01	Estrutura de governança da organização	4.1	un	G	D4	1	0			
IG02	Presidente órgão de governança é ou não diretor executivo	4.2	un	G						
IG03	Membros da governança independentes	4.3	un	G						
IG04	Mecanismos para recomendações para governança	4.4	un	G						
IG05	Relação remuneração governança e desempenho	4.5	un	G						
IG06	Garantia para evitar conflitos interesse	4.6	un	G						
IG07	Qualificações para membros da governança	4.7	un	G						
IG08	Declarações de missão e princípios conduta	4.8	un	G	D4	1	0			
IG09	Procedimentos para supervisionar a organização	4.9	un	G	D4	1	0			
IG10	Processos para avaliar a governança	4.10	un	G						

1- ANEEL (2015); GRI (2000)

Fonte: Autor.

Quadro 35– Simulação de descensão de classe de I5 para I4, empresa E07, ano 2021

ANÁLISE DE SENSIBILIDADE - INDICADORES indSU									
EMPRESA: E07			ANO: 2021		CLASSE ATUAL: I5		CLASSE A PROSPECTAR: I4		
IND.	DESCRIÇÃO RESUMIDA	UN.	GANHO/ CUSTO	CLASSE FINAL	VALOR INICIAL	VALOR FINAL	VARIAÇÃO (%)	VALOR REFERÊNCIA	% VALOR REFERÊNCIA
IE01	Número de localidades atendidas (municípios)	un	G						
IE02	Número de escritórios comerciais	un	G						
IE03	Número de transformadores de distribuição	un	G						
IE04	Energia vendida por empregado	MWh	G						
IE05	Número de consumidores por empregado	un	G						
IS01	Média de horas extras por empregado/ano	h	C						
IS02	Número de chamadas recebidas	un	C						
IS03	Número médio de atendentes	un	G						
IS04	% IAb - Índice de Abandono	%	C						
IS05	% ICO - Índice de Chamadas Ocupadas	%	C						
IS06	TMA - Tempo Médio de Atendimento (s)	s	C						
IS07	DER (horas)	h	C						
IS08	FER (unid.)	un	C						
IS09	Número de atendimentos realizados	un	G						
IS10	Número de atendimentos realizados fora do prazo	un	C						
IA01	Consumo de água por empregado (em m3)	m3	C						
IA02	Número de alunos atendidos	un	G	I4	65000	26650.00	59	65000	41
IG01	Emissão de ações preferenciais	un	G						
IG02	Acionistas preferenciais têm direito a voto	un	G						
IG03	Mecanismos de divulgação sobre temas deliberados	un	G						
IG04	Processos administrativos, arbitrais ou judiciais contra a companhia	un	C						
IG05	Proibição de empréstimos e garantias em favor do controlador	un	G						
IG06	Existência de um canal para comunicações anônimas	un	G						
IG07	Relatório anual e/ou de sustentabilidade	un	G						
IG08	Programas para educação sobre sustentabilidade	un	G						
IG09	Aderência de compromissos voluntários	un	G						
IG10	Existência de comitê de sustentabilidade	un	G						
IG11	Parecer de auditoria independente	un	G						
IG12	Compromisso com o combate à corrupção	un	G						

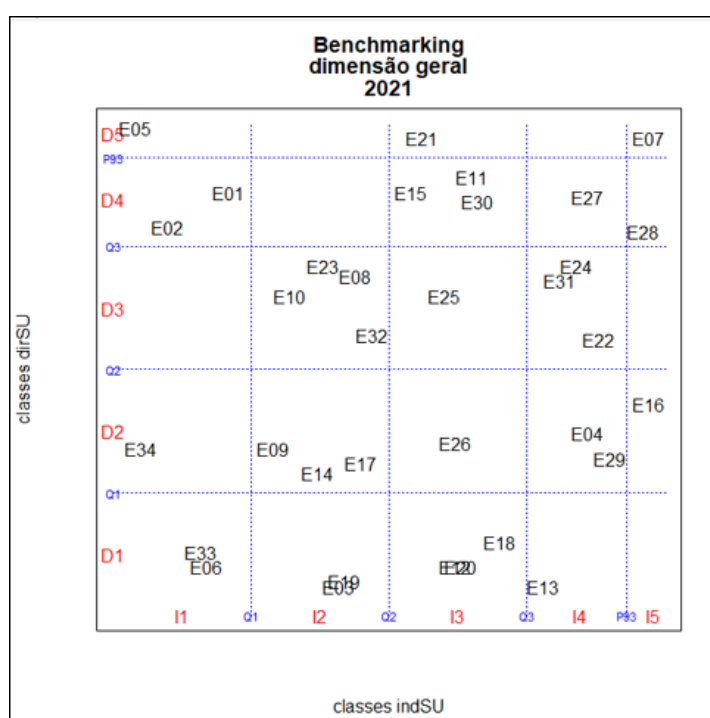
Fonte: Autor.



## 5.5. VALIDAÇÃO

A validação foi realizada com base nos resultados do *benchmarking*, visando verificar as prováveis relações entre os indicadores de sustentabilidade ou desempenho com as práticas das empresas concessionárias de distribuição de energia elétrica. Em destaque e, como exemplos, estão as empresas E07, E28 e E27 vis-à-vis à posição simultânea acima do terceiro quartil (Q3), em ambos os eixos, conforme mostrado na Figura 50.

Figura 50– Resultado do *benchmarking* e posições das empresas E07, E28 e E27



Fonte: Autor.

A empresa E07, posição na matriz (I5, D5), em seu Relatório de Responsabilidade Socioambiental e Econômico-Financeiro (RSA), ano 2021, demonstra várias práticas relacionadas à sustentabilidade ou desempenho, tais como: práticas de compra, de ESG, de gestão, sustentabilidade empresarial, ambiental, de governança, etc. Assim, visando identificar possíveis boas práticas da empresa que possam confirmar a sua posição de liderança no *benchmarking* em 2021 e, conseqüentemente, validar o modelo proposto de avaliação da sustentabilidade e desempenho, inicialmente, foram analisados os indicadores normalizados. Foram adotados os seguintes critérios para análise: indicadores

com posição relativa maior que 0,5 ou 50%, e os indicadores de governança corporativa com valor “1”:

dirSU: IS02, IS03, IS04, IS06, IG01, IG08, IG09;

indSU: IE01, IE04, IS03, IS06, IA02, IG01, IG06, IG09, IG10, IG12.

Para cada indicador analisado, evidenciou-se o trecho do RSA relacionado à alguma prática de sustentabilidade ou desempenho, conforme Quadros 36 e 37.

Quadro 36– Práticas de sustentabilidade ou desempenho da empresa E07 relacionadas aos indicadores dirSU em destaque, ano 2021

Origem	Indicador	Descrição	Principal(is) prática(s) evidenciada(s)
dirSU	IS02	Empregados com idade entre 31 e 40 anos (%)	O tempo médio de vínculo trabalhista dos empregados desligados foi de 13,8 anos em 2021. A empresa monitora o nível de satisfação dos empregados por meio de pesquisas de clima e <i>workshops</i> . As contratações são por concurso público. Há práticas para elogio dos empregados.
	IS03	Empregados com idade entre 41 e 50 anos (%)	
	IS04	% mulheres/total de empregados	Participa do Programa Pró-Equidade de Gênero e Raça, para disseminar novas concepções na gestão de pessoas e na cultura organizacional, e para alcançar a igualdade entre mulheres e homens no mundo do trabalho.
	IS06	Benefícios com Educação	Além do benefício de auxílio educação, oferece subsídios aos empregados para a realização de cursos técnicos, de graduação e pós-graduação.
	IG01	Estrutura de governança da organização	Possui estrutura de governança com quatro órgãos deliberativos permanentes, divididos por níveis de responsabilidade e assessorados por comitês técnicos.
	IG08	Declarações de missão e princípios conduta	Inclui o desenvolvimento com sustentabilidade na declaração de missão.
	IG09	Meios para supervisionar a organização	Para aferir a maturidade e o nível da qualidade da governança corporativa, participa de índices e reconhecimentos de mercado que permitem a comparação com seus pares e empresas de diversos setores.

Fonte: Autor.

Quadro 37– Práticas de sustentabilidade ou desempenho da empresa E07 relacionadas aos indicadores indSU em destaque, ano 2021

Origem	Indicador	Descrição	Principal(is) prática(s) evidenciada(s)
indSU	IE01	Número de localidades atendidas (municípios)	Atendimento integral a 4,9 milhões de consumidores de energia distribuídos por 1.149 localidades, em 400 municípios.
	IE04	Energia vendida por empregado (MWh)	Cadeia de Valor de forma a entregar o melhor aos clientes. Forte normalização e certificação de procedimentos, de forma a garantir o adequado funcionamento dos processos.
	IS03	Número médio de atendentes	Diversos canais de atendimento: 0800, Ouvidoria, Site, APP e presencialmente em agências e postos de atendimento localizados em todos os municípios da área de concessão.
	IS06	TMA - Tempo Médio de Atendimento (s)	
	IA02	Número de alunos atendidos	Projeto Ciclo de Profissões: objetiva auxiliar estudantes do ensino médio de escolas públicas no planejamento de carreiras.
	IG01	Emissão de ações preferenciais	Migração para o nível 2 de Governança Corporativa da B3, que contempla práticas diferenciadas, como <i>tag along</i> de 100% para as ações ordinárias e preferenciais.
	IG06	Existência de um canal para comunicações anônimas	Existência de canal de comunicação confidencial.
	IG09	Aderência de compromissos voluntários	Programa de voluntariado corporativo para a realização de ações sociais.
	IG10	Existência de comitê de sustentabilidade	Coordenar e desenvolver ações para orientar e acompanhar o atendimento da política de sustentabilidade; promover a aderência dos relatórios de sustentabilidade e socioambiental; validar respostas de questionários; e definir os indicadores socioambientais.
	IG12	Compromisso com o combate à corrupção	Programa de Integridade, Código de Conduta, Política de Gestão Integrada de Riscos Corporativos, Política de Transações com Partes Relacionadas, Política do Canal de Denúncias, Norma Administrativa de Prevenção e Tratamento de Fraude e Corrupção.

Fonte: Autor.

Com relação à empresa E28, posição na matriz (I5, D4), aquela apresentou posição relativa maior que 0,5 ou 50% em relação aos seguintes indicadores:

dirSU: IS02, IS04, IS06, IG01, IG06, IG09;

indSU: IE01, IE04, IE05, IS03, IS05, IS06, IA02, IG06, IG09, IG10, IG12.

Os indicadores de governança corporativa são aqueles que possuem valor “1”. Para cada indicador analisado, evidenciou-se o trecho do RSA relacionado à alguma prática de sustentabilidade ou desempenho, conforme Quadros 38 e 39.

Quadro 38– Práticas de sustentabilidade ou desempenho da empresa E28 relacionadas aos indicadores dirSU em destaque, ano 2021

Origem	Indicador	Descrição	Principal(is) prática(s) evidenciada(s)
dirSU	IS02	Empregados com idade entre 31 e 40 anos (%)	Com a adoção de melhores práticas em recursos humanos, obteve adesão de 94% dos colaboradores, 89% de respostas favoráveis ao engajamento e 83% ao suporte organizacional, com alta adaptação ao papel de cada um na organização e confirmando as condições favoráveis para cada empregado exercer seu potencial.
	IS04	% mulheres/total de empregados	Promoção da igualdade de gênero por meio de oferecimento de turmas exclusivas para mulheres em escola de eletricitistas.
	IS06	Benefícios com Educação	Ofertas de programas de idiomas, incentivo educacional para cursos de graduação, técnico, MBA/Pós-graduação.
	IG01	Estrutura de governança da organização	Estrutura constituída de Conselho de Administração, Conselho Fiscal e Diretoria-Executiva.
	IG06	Garantia para evitar conflitos interesse	Manual de Patrocínios, Doações e Investimento Social Privado para orientar e auxiliar a conduta dos profissionais, com o objetivo de evitar situações de conflitos de interesses reais, potenciais ou percebidos e situações que possam caracterizar lavagem de dinheiro, financiamento de atividades ilícitas, suborno ou qualquer forma de corrupção pública ou privada, bem como reforçar os padrões éticos e de integridade presentes no Código de Ética.
IG09	Supervisão da organização	Conselho de Administração é responsável pelas partes interessadas na supervisão da organização.	

Fonte: Autor.

Quadro 39– Práticas de sustentabilidade ou desempenho da empresa E28 relacionadas aos indicadores indSU em destaque, ano 2021

Origem	Indicador	Descrição	Principal(is) prática(s) evidenciada(s)
indSU	IE01	Número de localidades atendidas (municípios)	Centro de Gerenciamento de Redes Inteligentes. Sistemas para manutenção preditiva, antecipando possíveis falhas e ocorrências.
	IE04	Energia vendida por empregado (MWh)	Assistente virtual com serviços de <i>chatbot</i> (uso de inteligência artificial) para WhatsApp e Facebook, plataforma de negociação digital, APP, serviços da agência virtual, novas formas de pagamento, cadastro e envio de fatura.
	IE05	Número de consumidores por empregado	
	IS03	Número médio de atendentes	
	IS05	% ICO - Índice de Chamadas Ocupadas	
	IS06	TMA - Tempo Médio de Atendimento (s)	
	IA02	Número de alunos atendidos	Capacitação de professores e alunos nos temas de combate ao desperdício de energia e eficiência energética.
	IG06	Existência de um canal para comunicações anônimas	Canais de reclamações e de denúncias.
	IG09	Aderência de compromissos voluntários	Programa de Voluntariado da empresa.
	IG10	Existência de comitê de sustentabilidade	Comitê como órgão estratégico e de engajamento do Conselho de Administração, bem como ser facilitador da integração de aspectos ESG à estratégia dos negócios da empresa.
	IG12	Compromisso com o combate à corrupção	Certificação do Sistema de Gestão Antissuborno. Integração à Plataforma Anticorrupção da Rede Brasil do Pacto Global da ONU.

Fonte: Autor.

Com relação à empresa E27, posição (I4, D4), aquela apresentou posição relativa maior que 0,5 ou 50% em relação aos seguintes indicadores:

dirSU: IE01, IE05, IS01, IS02, IS05, IG01, IG04, IG06;

indSU: IE01, IE02, IE03, IE05, IS05, IS06, IG06, IG09, IG12.

Os indicadores de governança corporativa são aqueles que possuem valor “1”. Para cada indicador analisado, evidenciou-se o trecho do RSA possivelmente relacionado à alguma prática de sustentabilidade ou desempenho, conforme Quadros 40 e 41.

Quadro 40– Práticas de sustentabilidade ou desempenho da empresa E27 relacionadas aos indicadores dirSU em destaque, ano 2021

Origem	Indicador	Descrição	Principal(is) prática(s) evidenciada(s)
dirSU	IE01	Número de consumidores atendidos	Oferta da melhor experiência aos clientes, com serviços confiáveis, inovadores e seguros.
	IE05	Capacidade instalada (MVA)	Clientes no centro da estratégia, que visa à máxima satisfação e à melhor experiência em toda a jornada de atendimento.
	IS01	Número de empregados próprios	Programas para o aperfeiçoamento de competências dos empregados e assistência para transição de carreira. Programas para avaliações regulares de desempenho e de desenvolvimento de carreira. Diversidade em órgãos de governança e empregados. Adoção de ideias de empregados em programas de inovação.
	IS02	Empregados com idade entre 31 e 40 anos (%)	
	IS05	% empregados programa contratação aprendizes	
	IG01	Estrutura de governança da organização	Avaliação e gerenciamento dos desempenhos econômico-financeiros, sociais e ambientais dos negócios, com vistas a assegurar a geração de valor em todas as atividades.
	IG04	Mecanismos para recomendações para governança	Adesão às recomendações do Instituto Brasileiro de Governança Corporativa (IBGC), às normas da Lei Sarbanes-Oxley (SOx) e do <i>Committee of Sponsoring Organizations (Coso)</i> .
	IG06	Garantia para evitar conflitos de interesse	Uso de critérios de comportamento para garantir a transparência e a proteção de interesses e estabelecer mecanismos para revelar potenciais conflitos de interesses dos responsáveis pela tomada de decisão.

Fonte: Autor.

Quadro 41– Práticas de sustentabilidade ou desempenho da empresa E27 relacionadas aos indicadores indSU em destaque, ano 2021

Origem	Indicador	Descrição	Principal(is) prática(s) evidenciada(s)
indSU	IE01	Número de localidades atendidas (municípios)	Aumento da capacidade de geração de energia limpa e renovável; desenvolvimento de processos cada vez mais sustentáveis nas áreas de construção; uso de equipamentos mais eficientes e econômicos; e acesso à energia em localidades remotas.
	IE02	Número de escritórios comerciais	
	IE03	Número de transformadores de distribuição	
	IE05	Número de consumidores por empregado	
	IS05	% ICO - Índice de Chamadas Ocupadas	Adoção de metodologias de mensuração e acompanhamento, como as definidas pelo Conselho Internacional de Relato Integrado ( <i>International Integrated Reporting Council – IIRC</i> ).
	IS06	TMA - Tempo Médio de Atendimento (s)	
	IG06	Existência de um canal para comunicações anônimas	Adoção do Canal Ético, por meio do qual podem ser comunicados práticas e comportamentos inadequados que violem os princípios do Código de Ética.
	IG09	Aderência de compromissos voluntários	Programa de Voluntariado da empresa.
	IG12	Compromisso com o combate à corrupção	Adesão ao Pacto Empresarial pela Integridade e Contra a Corrupção. Plano de Tolerância Zero com a Corrupção. Adoção de ferramentas para avaliação de riscos.

Fonte: Autor.

De forma complementar à etapa de Validação, investigou-se também, a existência de prováveis *outliers* frente aos indicadores de sustentabilidade ou desempenho apresentados pelas empresas em seus respectivos RSAs, visando descobrir a existência de boas práticas ou fatos relevantes.

Com bases nas Tabelas 4 e 5, que ilustram os indicadores mais referenciados pelas regras de decisão geradas pelo algoritmo de classificação VC-DomLEM, procedeu-se à análise dos três primeiros indicadores de cada tabela. Para a avaliação sobre a existência ou não de *outlier*, foi utilizado o intervalo interquartil  $(Q3 - Q1) \times 1,5$ , para mais ou para menos, e como referências, os quartis Q3 ou Q1, respectivamente (VINUTHA *et al.*, 2018). O Quadro 42 ilustra estes indicadores, empresas, valores *outliers*, períodos de referência e práticas/fatos relevantes.

Quadro 42– Indicadores mais referenciados pelas regras de decisão e respectivos valores *outliers*

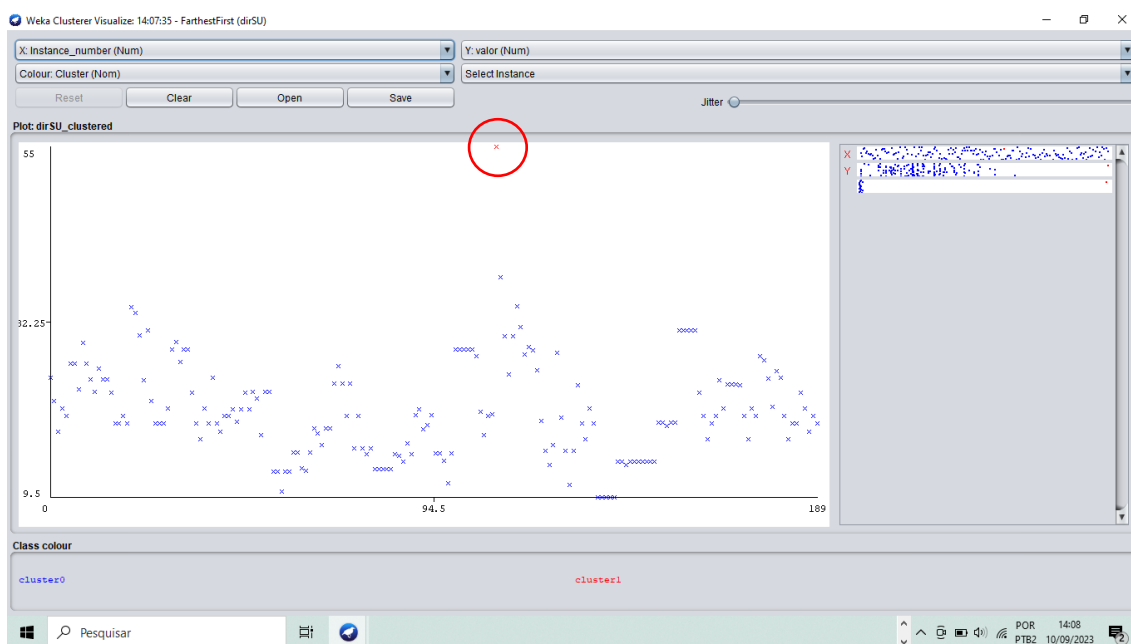
Origem	Indicador	Descrição	Empresa(s)	Valor outlier	Ano	Prática(s) ou fato(s) relevante(s)
dirSU	IS03	Empregados com idade entre 41 e 50 anos (%)	E23	55	2021	Há uma Política de liberdade de associação, pautada pelo respeito e reconhecimento da legitimidade das instituições como representantes dos trabalhadores, dentro dos princípios éticos e valores que norteiam as boas práticas sindicais, com garantia a todos os seus colaboradores o direito de livre associação, organização e mobilização sindical. Há canal interno de comunicação para relatar andamento das negociações. Há comprometimento com o cumprimento de legislações nacionais e internacionais sobre direitos humanos e trabalhistas.
	IA03	Consumo total de água, fonte Rede pública (em m <sup>3</sup> )	E08; E09	9.478.707	2019	Neste caso, trata-se de empreendimento termelétrico localizado em região de estresse hídrico, o que ocasionou na concorrência pelo consumo da água com outros usos locais, como o abastecimento da população.
	IS04	% mulheres/total de empregados	E25	33,06	2021	Há programas destinados especificamente para atendimento, atração e retenção de mulheres e, principalmente, com atenção para grávidas: importância da nutrição e das atividades físicas, e orientações sobre cuidados com recém-nascidos. Adota licença-maternidade de 180 dias.
indSU	IA01	Consumo de água por empregado (em m3)	E26	205,54	2015	De 2015 para 2021 houve redução para 1/10 do valor, aproximadamente.
	IE05	Número de consumidores por empregado	E24	2342	2021	Possui um Conselho de Consumidores, que propicia o protagonismo dos clientes na cadeia de valor. A empresa figurou como finalista do Prêmio de Satisfação do Cliente em 2020.
	IE03	Número de transformadores de distribuição	E04	925.468	2021	Implantação de controle de resíduos por sistema informatizado, com o objetivo de garantir maior rastreabilidade e disponibilidade dos dados de geração e destinação de resíduos. Comprometimento com a Política Nacional de Resíduos Sólidos. Possui certificação do Sistema de Gestão Ambiental Nível 1 e atendimento a NBR ISO 9001:2015. Segue Convenção de Basileia: não importa ou exporta resíduos perigosos; todo resíduo não recuperado é encaminhado para aterro ou tratamento, conforme caracterização.

Fonte: Autor.



Ademais, utilizou-se da técnica de *clustering*, por meio do *software* Weka, v. 3.8.4 (WITTEN; FRANK, 2005), algoritmo *FarthestFirst*, como auxílio na identificação de prováveis *outliers*. Neste algoritmo,  $k$  pontos são primeiro selecionados como centros de *cluster*: o primeiro centro é selecionado aleatoriamente; o segundo, é escolhido como o ponto mais distante do primeiro; e, cada centro restante é determinado selecionando o ponto mais distante do conjunto de centros já escolhidos. E, apresenta desempenho superior ao algoritmo clássico de *clustering*, *k-means* (VADEYAR; YOGISH, 2014). Como exemplo, a Figura 51 ilustra o resultado final da avaliação no período 2013-2021 para o indicador IS03, dirSU, Empregados com idade entre 41 e 50 anos (%), com o valor de 55% em destaque (circunferência em vermelho). Foi identificado um único *cluster*, *cluster 1*, com esse valor.

Figura 51– Uso da técnica de *clustering*, algoritmo *FarthestFirst*, para identificar valor *outlier*



Fonte: Autor, adaptado do software Weka.

Como exemplo da utilização de *outliers* para avaliações em desenvolvimento sustentável, há o estudo sobre detecção de *outliers* na poluição do ar em Calcutá, Índia (AHMAD *et al.*, 2023). Outras técnicas com o uso da RST poderiam ser utilizadas para detectar *clusters*, como *Rough K-Means* (LINGRAS; WEST, 2004), *Rough Clustering* (LINGRAS; PETERS, 2012) e *Fuzzy Rough* (VLUYMANS, 2019).

## 6. CONCLUSÕES

A presente Tese de Doutorado propõe um modelo integrado de avaliação da sustentabilidade e desempenho para empresas concessionárias de distribuição de energia elétrica, vis-à-vis os requisitos definidos pela Agência Nacional de Energia Elétrica, ANEEL.

Os indicadores definidos pela ANEEL constantes no Manual de Contabilidade do Setor Elétrico, MCSE, constituem meios para avaliar as empresas de energia com base em padrões internacionais, como o GRI, bem como, em necessidades específicas com o intuito de apreciar o desempenho daquelas empresas. Como exemplo, o consumo de água em m<sup>3</sup> por empregado, como necessidade da ANEEL e, o consumo total de água (m<sup>3</sup>), como padrão GRI. Assim, os indicadores propiciam formas de medição quando se analisa uma empresa individualmente e, comparativamente, quando se constitui o conjunto de empresas de distribuição de energia elétrica.

Porém, os indicadores estão distribuídos em suas respectivas dimensões, por exemplo, a dimensão ambiental. Analisar todos ou alguns indicadores de uma única dimensão se constitui em elemento importante de gestão, contudo, não permite uma percepção ampla frente ao conjunto de todos os indicadores que foram reunidos. Assim, realizar o *benchmarking* das empresas se constitui em uma útil ferramenta de gestão, principalmente, no que se refere à gestão de riscos mesmo que de forma individualizada.

Neste contexto, o modelo proposto de avaliação integrada das empresas permitiu realizar não somente um *benchmarking*, tendo em vista as várias formas existentes de construí-lo. Mas, aproveitar a oportunidade como experiência única de maravilhamento da retórica, qual seja, a capacidade de voltar a se surpreender com aquilo que se tornou comum; ou, romper com o *imprinting* cultural, que impede de ver diferentemente do que se mostra. É uma oportunidade para se perguntar: o que pode existir ou ser realizado além do já conhecido termo *benchmarking*? Assim, se constitui em momento singular para a inovação!

Consequentemente, este estudo visou reunir saberes interdisciplinares, tais como: desenvolvimento sustentável e sustentabilidade; Apoio Multicritério à Decisão; Teoria dos Conjuntos Aproximativos/princípio de Dominância, teoria matemática para tratamento da imprecisão e incerteza de dados; *Machine Learning* ou Aprendizado de Máquina, subcampo da Inteligência Artificial; Tecnologia da Informação, por meio de Mineração de textos e de dados, e Recuperação de informação; Ciência de Dados; desempenho; gestão de empresas; gestão de riscos; e auditoria.

A contribuição desta Tese está na avaliação integrada da sustentabilidade e do desempenho das empresas concessionárias de distribuição de energia elétrica no Brasil, por meio de *benchmarking*, com Apoio Multicritério à Decisão e modelo preditivo para a classificação e gestão de riscos das empresas. Este modelo visa: a) classificar as empresas sob as dimensões econômica, social, ambiental, de governança corporativa e geral, e posicioná-las em uma matriz de sustentabilidade; b) realizar *benchmarking* das empresas; c) conhecer boas práticas em sustentabilidade e desempenho, inclusive para fins de políticas públicas; d) efetuar predição das classificações por Aprendizado de Máquina com base em dados históricos de indicadores, regras de decisão inferidas e simulações; e) realizar efetiva gestão de riscos das empresas com base em simulações e predições, bem como, oportunidade para se atingir metas.

Ressalte-se que, a própria ANEEL e a ABRADDEE, Associação Brasileira de Distribuidores de Energia Elétrica, utilizam alguma forma de *benchmarking* das empresas de energia elétrica. Contudo, utiliza-se de um *benchmarking* por ranqueamento, muitas vezes restrito a determinados indicadores, o qual pode não traduzir de forma mais abrangente as classificações das empresas.

Para que estes resultados fossem alcançados nesta proposta de Tese, houve uma revisão sistemática da literatura visando identificar os estudos e *gaps* relacionados à questão de pesquisa: uso do método construtivista ProKnow-C, três Eixos de pesquisa, onze bases multidisciplinares da CAPES e período de pesquisa de janeiro/2000 a agosto/2023. Foram coletados 1.294 estudos; ao final, 67 estudos foram utilizados para construir a base de conhecimento sobre o tema pesquisado. Em seguida, foi construída a base de dados das empresas concessionárias distribuidoras de energia elétrica, tendo os Relatórios de Responsabilidade Socioambiental e Econômico-Financeiro (RSA) das empresas como fontes principais de pesquisa. Foram coletados os relatórios disponíveis de 41 empresas, de um total de 53, para o período (ano-base) 2013-2021, em formato PDF. Como há empresas que fazem parte de grupos econômicos, ao final, o conjunto foi reduzido para 38 empresas. Foram extraídas aproximadamente, 684.000 linhas de texto dos RSAs para que fosse possível coletar os indicadores de sustentabilidade e desempenho. Ao final, foram reunidos 62 indicadores, sendo 33 indicadores diretamente associados à sustentabilidade e, 29 indiretamente associados. Esses indicadores foram devidamente tratados e verificados, por meio da identificação de valores mínimos e máximos, amplitudes, variações, medianas e possíveis *outliers* para cada indicador, exceto os indicadores de governança corporativa. E, para valores ausentes, procedeu-se

ao seu preenchimento com o uso de medianas na horizontal e na vertical, por empresa e por ano, respectivamente. Após esta etapa, foi utilizada teoria em Apoio Multicritério à Decisão, especificamente, *Rough Sets Theory* e sua extensão, *Dominance Rough Set Approach* (DRSA), as quais permitiram o tratamento matemático da imprecisão e incerteza de dados, e análise dos indicadores, inclusive com a identificação de padrões de dados. No Aprendizado de Máquina, foram utilizados algoritmos indutores de regras de decisão com sensibilidade mínima de 60% para compor o método *ensemble bagging* de predição de classes.

Como limitações deste estudo, foi considerado um subconjunto do conjunto final de indicadores exigidos ou sugeridos pela ANEEL, devido à ausência de valores e inconsistências detectadas nos RSAs das empresas.

No que se refere a indicadores, torna-se condição *sine qua non*, possuir sensibilidade para interpretar valores numéricos. Até porque, imprecisões ou incertezas estão sempre presentes em qualquer atividade que visa algum tipo de medição. E, com isso, as classificações preditas pelos algoritmos devem ser sempre averiguadas, mesmo que a probabilidade final de sucesso seja alta. Por questões de ética, a decisão final sempre deverá caber a um gestor, público ou privado.

Ademais, como o modelo proposto permitiu o levantamento de boas práticas de sustentabilidade e desempenho a partir das empresas mais bem classificadas na matriz de sustentabilidade, o compartilhamento destas práticas torna-se elemento valioso para que as empresas avaliem o seu alinhamento frente aos Objetivos de Desenvolvimento Sustentável (ODS), da ONU. E, que também gestores públicos avaliem ou revisem as políticas públicas de desenvolvimento sustentável para o setor de energia elétrica visando possíveis aperfeiçoamentos.

Como estudos futuros, torna-se oportuno ampliar o conjunto de indicadores de sustentabilidade e desempenho, bem como, utilizar outras técnicas de Aprendizado de Máquina para fins de predição das classificações das empresas, como hotéis e de energia solar.

## 7. APRESENTAÇÕES E PUBLICAÇÕES

Como consequência da presente pesquisa, foram apresentados e publicados em anais de Simpósios, os seguintes artigos:

- 1- “A model for evaluating the sustainability and performance of companies in the Brazilian electric sector based on multicriteria decision and machine learning”, *27th International Sustainable Development Research Society Conference, Mid Sweden University*, 13 a 15/07/2021 (COUTO; RANGEL, 2021a) – Anexos I-a e I-b;
- 2- “Proposição de modelo de avaliação da sustentabilidade e desempenho de empresas do setor elétrico brasileiro com base em apoio multicritério e *machine learning*”, SBPO 2021, LIII Simpósio Brasileiro de Pesquisa Operacional, em 04/11/2021 (COUTO; RANGEL, 2021b) – Anexos II-a e II-b.

E foram publicados os seguintes artigos:

- 1- “Model proposition for predicting sustainability classes using multicriteria decision support and artificial intelligence”, *Gestão & Produção*, aceito em 10/10/2022, publicado em 2022 (COUTO; RANGEL, 2022) – Anexo III;
- 2- “Corporate governance aspects in sustainability and performance assessment models for energy companies: a systematic review of the literature”, *Gestão & Produção*, aceito em 18/12/2022, publicado em 2023 (COUTO; RANGEL, 2023) – Anexo IV.

## REFERÊNCIAS

- ABNT (2010), Associação Brasileira de Normas Técnicas, NBR ISO 26000, Diretrizes sobre responsabilidade social, 1ª ed.
- ADAMKAITE, J.; STREIMIKIENE, D.; RUDZIONIENE, K. (2022), The impact of social responsibility on corporate financial performance in the energy sector: Evidence from Lithuania, *Corp Soc Responsib Environ Manag.*, 1–14, DOI: 10.1002/csr.2340.
- AFONSO, M. H. F.; SOUZA, J. V.; ENSSLIN, S. R.; ENSSLIN, L. (2011), Como construir conhecimento sobre o tema de pesquisa? Aplicação do processo ProKnow-C na busca de literatura sobre avaliação do desenvolvimento sustentável, *Revista de Gestão Social e Ambiental - RGSA*, 5, 2, 47-62.
- AHMAD, M.; CHENG, W.; ZHAO, X. (2023), An Outlier Detection Study of Ozone in Kolkata India by the Classical Statistics, *Statistical Process Control and Functional Data Analysis, Sustainability*, 15, 12790, <https://doi.org/10.3390/su151712790>.
- AHMAD, M.; JAN, I.; JABEEN, G.; ALVARADO, R. (2021a), Does energy-industry investment drive economic performance in regional China: Implications for sustainable development, *Sustainable Production and Consumption*, 27, 176-192.
- AHMAD, T.; ZHANG, D.; HUANG, C.; ZHANG, H.; DAI, N.; SONG, Y.; CHEN, H. (2021b), Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities, *Journal of Cleaner Production, Artificial intelligence in sustainable*, 289, <https://doi.org/10.1016/j.jclepro.2021.125834>.
- ALI, S.; SMITH, K. A. (2006), On learning algorithm selection for classification, *Applied Soft Computing*, 6, 119-138.
- ALMEIDA, A. T. (2011), *O conhecimento e o uso de Métodos Multicritérios de Apoio à Decisão*, 2ª ed., Recife: Editora Universitária da UFPE.
- ALMEIDA, A. T. (2013), *Processo de Decisão nas Organizações: Construindo Modelos de Decisão Multicritério*, 1ª ed., São Paulo: Editora Atlas.
- ALMEIDA, A. T.; CAVALCANTE, C. A. V.; ALENCAR, M. H.; FERREIRA, R. J. P.; ALMEIDA-FILHO, A. T.; GARCEZ, T. V. (2015), *Multicriteria and Multiobjective Models for Risk, Reliability and Maintenance Decision Analysis*, International Series in Operations Research & Management Science, 231, Switzerland: Springer International Publishing.
- ANAND, G.; KODALI, R. (2008), Benchmarking the benchmarking models, *Benchmarking: An International Journal*, 15, 3, 257-291.
- ANDREOPOULOU, Z.; KOLIOUSKA, C. (2018), Benchmarking internet promotion of renewable energy enterprises: is sustainability present?, *Sustainability*, 2018, 10, 4187.
- ANEEL (2010), Agência Nacional de Energia Elétrica, Resolução Normativa nº 414/2010.

ANEEL (2014), Agência Nacional de Energia Elétrica, Resolução Normativa nº 605, de 11/03/2014.

ANEEL (2015), Agência Nacional de Energia Elétrica, Manual de Contabilidade do Setor Elétrico (MCSE).

ANEEL (2021), Agência Nacional de Energia Elétrica, Central de Informações Econômico-Financeiras, Acesso: < [https://www.aneel.gov.br/central-de-informacoes-economico-financeiras?p\\_p\\_id=ciefseuser\\_WAR\\_ciefseportlet&p\\_p\\_lifecycle=0&p\\_p\\_state=normal&p\\_p\\_mode=view&p\\_p\\_col\\_id=column-2&p\\_p\\_col\\_pos=1&p\\_p\\_col\\_count=3](https://www.aneel.gov.br/central-de-informacoes-economico-financeiras?p_p_id=ciefseuser_WAR_ciefseportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=3)>, em 30/03/2021.

ANEEL (2022a), Agência Nacional de Energia Elétrica, Central de conteúdos, relatórios e indicadores de distribuição, Acesso: < <https://www.gov.br/aneel/pt-br/centrais-de-conteudos/relatorios-e-indicadores/distribuicao>>, em 29/08/2022.

ANEEL (2022b), Agência Nacional de Energia Elétrica, Manual de Contabilidade do Setor Elétrico (MCSE).

AVEN, T. (2015), Risk Analysis, United Kingdom: John Wiley & Sons, Ltd, 2<sup>nd</sup> ed.

AVEN, T.; YLONEN, M. (2018), The Enigma of Knowledge in the Risk Field, in Aven, T.; Zio, E. (Eds.): Knowledge in Risk Assessment and Management, Oxford: John Wiley & Sons Ltd.

BAEZA-YATES, R.; RIBEIRO-NETO, B. (2013). Recuperação de Informação: conceitos e tecnologias das máquinas de busca. (2nd. Ed.), Porto Alegre: Bookman.

BANA E COSTA, C. A.; VANSNICK, J. C. (1993), Sur la quantification des jugements de valeur: L'approche MACBETH, Cahiers du LAMSADE, 117, Université Paris-Dauphine, Paris.

BARROS, J. J. C.; COIRA, M. L.; LOPEZ, M. P. C.; GOCHI, A. C.; SOARES, I. (2020), Probabilistic multicriteria environmental assessment of power plants: A global approach., Applied Energy, 260.a.

BLASZCZYNSKI, J.; ALMEIDA FILHO, A. T.; MATUSZYK, A.; SZELAG, M.; SLOWINSKI, R. (2021), Auto loan fraud detection using dominance-based rough set approach versus machine learning methods, Expert Systems with Applications, 163, 113740, <https://doi.org/10.1016/j.eswa.2020.113740>.

BLASZCZYNSKI, J.; GRECO, S.; MATARAZZO, B.; SLOWINSKI, R.; SZELAG, M. (2013a), Dominance-based Rough Set Approach Data Analysis Framework, User's guide, Acesso: < <http://www.cs.put.poznan.pl/jblaszczyński/Site/jRS.html> >, em 20/02/2013.

BLASZCZYNSKI, J.; GRECO, S.; MATARAZZO, B.; SLOWINSKI, R.; SZELAG, M. (2013b), jMAF - Dominance-based Rough Set Data Analysis Framework, Chapter 5 [In]:

A. Skowron, Z. Suraj (Eds.), *Rough Sets and Intelligent Systems - Professor Zdzisław Pawlak in Memoriam*, 1, Intelligent Systems Reference Library, 42, 185-209, Springer.

BLASZCZYNSKI, J.; SLOWINSKI, R.; SZELAG, M. (2009), VC-DomLEM: Rule induction algorithm for variable consistency rough set approaches. Technical Report RA-07/09, Poznań, University of Technology.

BLASZCZYNSKI, J.; SLOWINSKI, R.; SZELAG, M. (2011), Sequential covering rule induction algorithm for variable consistency rough set approaches. *Inf Sci*, 181:987–1002.

BOGETOFT, P. (2012), *Performance Benchmarking: Measuring and managing performance*, New York: Springer Science+Business Media.

CAIADO, R. G. G.; HEYMANN, M. C.; SILVEIRA, C. L. R.; MEZA, L. A.; QUELHAS, O. L. G. (2020), Measuring the eco-efficiency of Brazilian energy companies using DEA and directional distance function, *IEEE LATIN AMERICA TRANSACTIONS*, 18, 11, 1844-1852.

CHAMANDOUST, H.; DERAKHSHAN, G.; BAHRAMARA, S. (2020), Multi-objective performance of smart hybrid energy system with Multi-optimal participation of customers in day-ahead energy market, *Energy and Buildings*, 216.

CLARK, P.; NIBLETT, T. (1989), The CN2 induction algorithm, *Machine Learning*, 3, 261-283.

COHEN, W. W. (1995), Fast Effective Rule Induction, in *Proceedings of the Twelfth International Conference on Machine Learning*, 115–123, San Francisco: Morgan Kaufmann, <https://doi.org/10.1016/B978-1-55860-377-6.50023-2>.

COLLA, M.; IOANNOU, A.; FALCONE, G. (2020), Critical review of competitiveness indicators for energy projects., *Renewable & Sustainable Energy Reviews*, 125.

COUTO, A. B. G.; GOMES, L. F. A. M. (2010), A tomada de decisão em recursos humanos com dados replicados e inconsistentes: uma aplicação da teoria dos conjuntos aproximativos, *Pesquisa Operacional*, 30, 3, 657-686.

COUTO, A. B. G.; GOMES, L. F. A. M. (2020), Sovereign rating analysis through the dominance-based rough set approach, *Foundations of Computing and Decision Sciences*, 45, 1, 3-16.

COUTO, A. B. G.; RANGEL, L. A. D. (2021a), Sustainability and performance model of evaluation proposition for companies in the Brazilian electric sector based on multicriteria decision aiding and machine learning, 27<sup>th</sup> International Sustainable Development Research Society Conference, Mid Sweden University.

COUTO, A. B. G.; RANGEL, L. A. D. (2021b), Proposição de modelo de avaliação da sustentabilidade e desempenho de empresas do setor elétrico brasileiro com base em apoio multicritério e *machine learning*, LIII Simpósio Brasileiro de Pesquisa Operacional, SBPO 2021.



- COUTO, A. B. G.; RANGEL, L. A. D. (2022), Model proposition for predicting sustainability classes using multicriteria decision support and artificial intelligence, *Gestão & Produção*, 29, e6922, <https://doi.org/10.1590/1806-9649-2022v29e6922>.
- COUTO, A. B. G.; RANGEL, L. A. D. (2023), Corporate governance aspects in sustainability and performance assessment models for energy companies: a systematic review of the literature, *Gestão & Produção*, 30, e4322, <https://doi.org/10.1590/1806-9649-2022v29e4322>.
- DASHTPEYMA, M.; GHODSI, R. (2019), Developing the resilient solar energy management system: a hybrid qualitative-quantitative approach, *International Journal of Ambient Energy*, <https://doi.org/10.1080/01430750.2019.1630301>.
- DONG, R.; SHAO, C.; XIN, S.; LU, Z. (2023), A Sustainable Development Evaluation Framework for Chinese Electricity Enterprises Based on SDG and ESG Coupling. *Sustainability*, 15, 8960, <https://doi.org/10.3390/su15118960>.
- DU, W. S.; HU, B. Q. (2017), Dominance-based rough fuzzy set approach and its application to rule induction, *European Journal of Operational Research*, 261, 2, 690–703.
- EBRAHIMI, M.; RAHMANI, D. (2019), A five-dimensional approach to sustainability for prioritizing energy production systems using a revised GRA method: A case study, *Renewable Energy*, 135, 345-354.
- EKEL, P.; PEDRYCZ, W.; PEREIRA JR., J. (2020), *Multicriteria decision-making under conditions of uncertainty: A fuzzy set perspective*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 1<sup>st</sup> ed.
- ELKINGTON, J. (2020), *Green swans: The coming boom in regenerative capitalism*, Fast Company Press, New York.
- ENSSLIN, S. R.; ENSSLIN, L.; YAMAKAWA, E. K.; NAGAOKA, M. P. T.; AOKI, A. R.; SIEBERT, L. C. (2014), Processo estruturado de revisão da literatura e análise bibliométrica sobre avaliação de desempenho de processos de implementação de eficiência energética, *Revista Brasileira de Energia*, 20, 1, 21-50.
- EPSTEIN, M. J.; BUHOVAC, A. R. (2014), *Making sustainability work: Best practices in managing and measuring corporate social, environmental, and economic impacts*, Berrett-Koehler Publishers, Inc., San Francisco.
- ERSOY, N.; TASLAK, S. (2023), Comparative Analysis of MCDM Methods for the Assessment of Corporate Sustainability Performance in Energy Sector, *Ege Academic Review*, 23, 03, 341-362, doi: 10.21121/eab.986122.
- ERVURAL, B. C.; EVREN, R.; DELEN, D. (2018a), A multi-objective decision-making approach for sustainable energy investment planning, *Renewable Energy*, 126, 387-402.

ERVURAL, B. C.; ZAIM, S.; DEMIREL, O. F.; AYDIN, Z.; DELEN, D. (2018b), An ANP and fuzzy TOPSIS-based SWOT analysis for Turkey's energy planning, *Renewable and Sustainable Energy Reviews*, 82, 1538-1550.

FACELI, K.; LORENA, A. C.; GAMA, J.; ALMEIDA, T. A.; CARVALHO, A. C. P. L. F. (2021), *Inteligência artificial: uma abordagem de aprendizado de máquina*, Rio de Janeiro: LTC, 2ª ed.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. (1996a), The KDD Process for Extracting Useful Knowledge from Volumes of Data, *COMMUNICATIONS OF THE ACM*, 39, 11, 27-34.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. (1996b), From Data Mining to Knowledge Discovery in Databases, *American Association for Artificial Intelligence, AI Magazine*, 17, 3, 37-54.

FERNANDEZ, P.; MCCARTHY, I. P.; RAKOTOBE-JOEL, T. (2001), An evolutionary approach to benchmarking, *Benchmarking: An International Journal*, 8, 4, 281-305.

FERRETTI, P.; ZOLIN, M. B.; FERRARO, G. (2020), Relationships among sustainability dimensions: evidence from an Alpine area case study using Dominance-based Rough Set Approach, *Land Use Policy*, 92, 104457, <https://doi.org/10.1016/j.landusepol.2019.104457>.

FIGUEIRA, J.; GRECCO, S.; EHRGOTT, M. (2005), *Multiple Criteria Decision Analysis: State of the Art Surveys*, New York: Springer Science Business Media, Inc.

GARDAZI, S.S.N.; HASSAN, A.F.S.; JOHARI, J.B. (2020), Board of Directors Attributes and Sustainability Performance in the Energy Industry, *Journal of Asian Finance, Economics and Business*, 12, 317-328.

GEBRE, S. L.; CATTRYSSE, D.; VAN ORSHOVEN, J. (2021), Multi-Criteria Decision-Making Methods to Address Water Allocation Problems: A Systematic Review, *Water*, 13, 125. <https://doi.org/10.3390/w13020125>.

GOH, C. S.; CHONG, H.; JACK, L.; FARIS, A. (2020), Revisiting triple bottom line within the context of sustainable construction: A systematic review, *Journal of Cleaner Production*, 252, 1-10.

GOMES, L. F. A. M.; ARAYA, M. C. G.; CARIGNANO, C. (2004), *Tomada de decisões em cenários complexos*. São Paulo: Pioneira Thomson Learning.

GOMES, L. F. A. M. (2007), *Teoria da decisão*, São Paulo: Thomson Learning.

GOMES, L. F. A. M.; GOMES, C. F. S. (2019), *Princípios e métodos para tomada de decisão: enfoque multicritério*, São Paulo: Atlas, 6ª ed.

GOMES, L. F. A. M.; LIMA, M. M. P. P. (1992), From modeling individual preferences to multicriteria ranking of discrete alternatives: a look at prospect theory and the additive difference model. *Foundations of Computing and Decision Sciences*, 17(3), 171-184.

GOMES, L. F. A. M.; RANGEL, L. A. D. (2009a), An application of the TODIM method to the multicriteria rental evaluation of residential properties, *European Journal of Operational Research*, 193, 204–211.

GOMES, L. F. A. M.; RANGEL, L. A. D. (2009b), Determining the utility functions of criteria used in the evaluation of real estate, *Int. J. Production Economics*, 117, 420–426.

GOULART, D. D. (2013), Avaliação de índices de eficiência e de produtividade de distribuidoras de energia elétrica no Brasil aplicando análise envoltória de dados (DEA), Dissertação de Mestrado. Universidade Federal do Pampa.

GRECO, S.; INUIGUCHI, M.; SLOWINSKI, R. (2006), Fuzzy rough sets and multiple-premise gradual decision rules, *International Journal of Approximate Reasoning*, 41, 2, 179–211.

GRECO, S.; MATARAZZO, B.; SLOWINSKI, R. (2008), Fuzzy set extensions of the dominance-based rough set approach', in Bustince, H. *et al.* (Eds.): *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, Springer.

GRECO, S.; MATARAZZO, B.; SLOWINSKI, R.; STEFANOWSKI, J. (2001), An Algorithm for Induction of Decision Rules Consistent with the Dominance Principle, in Ziarko, W.; Yao, Y. (Eds.): *Rough Sets and Current Trends in Computing*, RSCTC 2000, Lecture Notes in Computer Science, v. 5, Springer. [https://doi.org/10.1007/3-540-45554-X\\_37](https://doi.org/10.1007/3-540-45554-X_37).

GRECO, S.; PAWLAK, Z.; SLOWINSKI, R. (2004), Can Bayesian confirmation measures be useful for rough set decision rules?, *Engineering Applications of Artificial Intelligence*, 17, 345–361, doi:10.1016/j.engappai.2004.04.008.

GRI (2000), Global Reporting Initiative, Sustainability Reporting Guidelines & Electric Utility Sector Supplement, RG version 3.0/EUSS.

GRZYMALA-BUSSE, J. W. (1988), Knowledge acquisition under uncertainty – a rough set approach, *Journal of Intelligent and Robotic Systems*, 1, 3-16.

GRZYMALA-BUSSE, J. W. (1997), A new version of the rule induction system LERS, *Fundamenta Informaticae*, 31, 27-39.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. (2009), *Análise multivariada de dados*, 6ª ed., Porto Alegre: Bookman.

HAN, J.; KAMBER, M.; PEI, J. (2012), *Data mining: concepts and techniques*. 3rd ed., Morgan Kaufmann Publishers, Elsevier, Waltham (USA).

HASSAN, A. (2019), Do renewable energy incentive policies improve the performance of energy firms? Evidence from OECD countries, *OPEC Energy Review*, 43, 168-192.

HASTIE, R.; DAWES, R. M. (2010), *Rational choice in an uncertain world: the psychology of judgment and decision making*, California: SAGE Publications, Inc., 2nd ed.

HAWKING, S. W. (2015), *Uma breve história do tempo*, Rio de Janeiro: Intrínseca, 1ª ed.

HENDIANI, S.; SHARIF, E.; BAGHERPOUR, M.; GHANNADPOUR, S. F. (2020), A multi-criteria sustainability assessment approach for energy systems using sustainability triple bottom line attributes and linguistic preferences, *Environment, Development and Sustainability*, 22:7771–7805, <https://doi.org/10.1007/s10668-019-00546-7>.

HERBOHN, K.; WALKER, J.; LOO, H.Y.M. (2014), Corporate Social Responsibility: The Link Between Sustainability Disclosure and Sustainability Performance, *Abacus*, 50, 422-459.

HIGUCHI, K. (2001), KH Coder, Acesso: < <http://khcoder.net/en/#>>, em 10/09/2019.  
IRFAN, M.; ZHAO, Z.-Y.; AHMAD, M.; BATOOL, K.; JAN, A.; MUKESHIMANA, M.C. (2019), Competitive assessment of Indian wind power industry: A five forces model, *Journal of Renewable and Sustainable Energy*, 11.

ILYAS, I. F.; CHU, X. (2019), *Data Cleaning*, Association for Computing Machinery, ACM Books series, University of Waterloo.

JARVIS, D.S.L.; SOVACOOOL, B.K. (2011), Conceptualizing and evaluating best practices in electricity and water regulatory governance, *Energy*, 36, 4340-4352.

JENSEN, R.; CORNELIS, C. (2008), A new approach to fuzzy-rough nearest neighbour classification, in Chan, C. C., Grzymala-Busse, J. W. and Ziarko, W. P. (Eds.): *Rough Sets and Current Trends in Computing*, RSCTC 2008, Lecture Notes in Computer Science, 5306, 310–319, Berlin: Springer.

JENSEN, R.; CORNELIS, C. (2010), Fuzzy-rough instance selection, WCCI IEEE World Congress on Computational Intelligence, 1776–1782.

JENSEN, R.; CORNELIS, C. (2011), Fuzzy rough nearest neighbour classification and prediction, *Theoretical Computer Science*, 412, 42, 5871–5884.

JENSEN, R.; CORNELIS, C.; SHEN, Q. (2009), Hybrid fuzzy-rough induction and feature selection, *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Korea, 1151–1156.

KAHNEMAN, D. (2012), *Rápido e devagar: duas formas de pensar*, Rio de Janeiro: Objetiva.

KAHNEMAN, D.; SIBONY, O.; SUNSTEIN, C. R. (2021), *Ruído: uma falha no julgamento humano*, Rio de Janeiro: Objetiva, 1ª ed.

KARNITIS, E.; BICEVSKIS, J.; KARNITIS, G. (2021), Measuring the implementation of the agenda 2030 vision in its comprehensive sense: Methodology and tool, *Energies*, 14, 856, <https://doi.org/10.3390/en14040856>.

KAROLAK, M.; PALKOWSKI, L.; KUBIAK, B.; BLASZCZYNSKI, J.; LUNIO, R.; SAWICKI, W.; SLOWINSKI, R.; KRYSINSKI, J. (2020), Application of dominance-based rough set approach for optimization of pellets tableting process, *Pharmaceutics*, 12, 1024, 1-14.

KEENEY, R. L.; RAIFFA, H. (1976), *Decision with multiples objectives: preferences and value tradeoffs*. New York: Wiley.

KOLOSOK, S.; PIMONENKO, T.; YEVDOKYMOVA, A.; HAJIYEV, N. O.; PALIENKO, M.; PRASOL, L. (2020), ENERGY EFFICIENCY POLICY: IMPACT OF GREEN INNOVATIONS, *Marketing and Management of Innovations*, 50-60.

KWAKKEL, J.H.; PRUYT, E. (2013), Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty, *Technological Forecasting and Social Change*, 80, 419-431.

KOHAVI, R. (1995), The power of decision tables, in Lavrac, N. and Wrobel, S. (Eds.): *Machine Learning: ECML-95, Lecture Notes in Computer Science*, 912, Berlin: Springer, [https://doi.org/10.1007/3-540-59286-5\\_57](https://doi.org/10.1007/3-540-59286-5_57).

KUMAR, A.; GUPTA, J.; DAS, N. (2022), Revisiting the influence of corporate sustainability practices on corporate financial performance: An evidence from the global energy sector, *Bus Strat Env.*, 31, 3231–3253, DOI: 10.1002/bse.3073.

LAMEIRA, V. J.; NESS, V. L. J.; QUELHAS, O. L. G.; PEREIRA, R. G. (2013), Sustentabilidade, valor, desempenho e risco no mercado de capitais brasileiro, *Revista Brasileira de Gestão de Negócios*, 15, 46, 76-90.

LANGLEY, P.; SIMON, H. A. (1995), Applications of machine learning and rule induction, *Communications of the ACM*, 38, 11, 55-64.

LAPES (Laboratório de Pesquisa em Engenharia de *Software*) (2021), Departamento de Computação, Universidade Federal de São Carlos (UFSCar), Acesso: <[http://lapes.dc.ufscar.br/tools/start\\_tool](http://lapes.dc.ufscar.br/tools/start_tool)>, em 20/02/2021.

LAPLACE, P-S. (2010), *Ensaio filosófico sobre as probabilidades*, Rio de Janeiro: Contraponto, 1ª ed.

LI, M.-J.; TAO, W.-Q. (2017), Review of methodologies and polices for evaluation of energy efficiency in high energy-consuming industry, *Applied Energy*, 187, 203-215.

LI, Y.; SHI, X.; YAO, L. (2016), Evaluating energy security of resource-poor economies: A modified principle component analysis approach, *Energy Economics*, 58, 211-221.

LI, Z.; DONG, Z. (2020), Comprehensive performance evaluation of Chinese nuclear power listed companies based on HSE perspective, *IOP Conference Series: Earth and Environmental Science*, 510.

LIMA, G. A. B. O. (2010), *Modelos de categorização: apresentando o modelo clássico*

e o modelo de protótipos, *Perspectivas em Ciência da Informação*, 15, 2, 108-122.

LINGRAS, P.; PETERS, G. (2012), Applying Rough Set Concepts to Clustering, in Peters, G., Lingras, P., Slezak, D., Yao, Y. (Eds.): *Rough Sets: Selected Methods and Applications in Management and Engineering*, Advanced Information and Knowledge Processing, London: Springer-Verlag.

LINGRAS, P.; WEST, C. (2014), Interval Set Clustering of Web Users with Rough K-Means, *Journal of Intelligent Information Systems*, 23:1, 5–16.

MAKRIDOU, G.; DOUMPOS, M.; LEMONAKIS, C. (2023), Relationship between ESG and corporate financial performance in the energy sector: empirical evidence from European companies, *International Journal of Energy Sector Management*, DOI 10.1108/IJESM-01-2023-0012.

MICHALSKI, R.S.; KAUFMAN, K.; WNEK, J. (1991), The AQ family of learning programs: a review of recent developments and an exemplary application, *Reports of Machine Learning and Inference Laboratory*, George Mason University.

MILOJEVIC, M.; URBANSKI, M.; TERZIC, I.; PRASOLOV, V. (2020), Impact of non-financial factors on the effectiveness of audits in energy companies, *Energies*, 13.

MORIARTY, J. P. (2011), A theory of benchmarking. *Benchmarking: An International Journal*, 18, 4, 588-611.

MORIN, E. (2013), *Os setes saberes necessários à educação do futuro*, São Paulo : Cortez Editora; Brasília, DF : UNESCO, 1ª ed.

MORIN, E. (2015), *Lições de um século de vida*, Rio de Janeiro: Bertrand Brasil, 1ª ed.

MORIN, E.; VIVERET, P. (2013), *Como viver em tempo de crise?*, Rio de Janeiro: Bertrand Brasil, 1ª ed.

MURPHY, K. P. (2012), *Machine Learning: a probabilistic perspective*, London: The MIT Press.

NAWAZ, W.; KOÇ, M. (2019), Exploring organizational sustainability: Themes, functional areas, and best practices, *Sustainability*, 11, 4307.

NOWICKI, R. K. (2019), *Rough Set–Based Classification Systems*, Studies in Computational Intelligence, 802, Switzerland: Springer.

ODS BRASIL (2022a). *Objetivos de desenvolvimento sustentável*. Acesso: <<https://odsbrasil.gov.br/home/agenda>>, em 25/01/2022.

ODS BRASIL (2022b). *Objetivos de desenvolvimento sustentável*. Acesso: <<https://odsbrasil.gov.br/relatorio/sintese>>, em 25/01/2022.

PACHECO, L. M.; ALVES, M. F. R.; LIBONI, L. B. (2018), Green absorptive capacity: A mediation-moderation model of knowledge for innovation, *Business Strategy and the Environment*, 27, 1502-1513.

PAPAGEORGIU, K.; CARVALHO, G.; PAPAGEORGIU, E.I.; BOCHTIS, D.; STAMOULIS, G. (2020), Decision-making process for photovoltaic solar energy sector development using fuzzy cognitive map technique, *Energies*, 13.

PAPAPOSTOULOU, A.; MEXIS, F. D.; KARAKOSTA, C.; PSARRAS, J. (2022), A Multicriteria Tool to Support Decision-Making in the Early Stages of Energy Efficiency Investments, 8th International Conference on Decision Support System Technology, ICDSST 2022, DOI: 10.1007/978-3-031-06530-9\_15.

PATARI, S.; ARMINEN, H.; TUPPURA, A.; JANTUNEN, A. (2014), Competitive and responsible? The relationship between corporate social and financial performance in the energy sector., *Renewable & Sustainable Energy Reviews*, 37, 142 - 154.

PATARI, S.; JANTUNEN, A.; KYLAHEIKO, K.; SANDSTROM, J. (2012), Does Sustainable Development Foster Value Creation? Empirical Evidence from the Global Energy Industry, *Corporate Social Responsibility and Environmental Management*, 19, 317-326.

PATYAL, V. S.; KUMAR, R.; LAMBA, K.; MAHESHWARI, S. (2023), Performance evaluation of Indian electricity distribution companies: An integrated DEA-IRP-TOPSIS approach, *Energy Economics*, 124, 106796, <https://doi.org/10.1016/j.eneco.2023.106796>.

TAHERDOOST, H.; MADANCHIAN, M. (2023), Multi-Criteria Decision Making (MCDM) Methods and Concepts, *Encyclopedia*, 3, 77–87. <https://doi.org/10.3390/encyclopedia3010006>.

PAWLAK, Z. (1982), Rough sets, *Int. J. Comput. Inf. Sci*, 11, 341-356.

PAWLAK, Z. (1991), *Rough sets. Theoretical aspects of reasoning about data*, Kluwer Academic Publishers, Dordrecht.

PAWLAK, Z. (2000), Rough sets and decision analysis, *Information Systems & Operational Research*, 38, 3, 132-144.

PAWLAK, Z. (2002), Rough sets, decision algorithms and Bayes' theorem, *European Journal of Operational Research*, 136, 181-189.

PAWLAK, Z.; GRZYMALA-BUSSE, J.; SLOWINSKI, R.; ZIARKO, W. (1995), Rough sets, *Communications of the ACM*, 38, 11, 88-95.

PAWLAK, Z.; SLOWINSKI, R. (1994), Rough set approach to multi-attribute decision analysis, *European Journal of Operational Research*, 72, 443-459.

PAZ, F. J.; KIPPER, L. M. (2016), Sustentabilidade nas organizações: vantagens e desafios, *Gestão da Produção, Operações e Sistemas*, 11, 2, 85-102.

PEREIRA NETO, F.; CÂNDIDO, G. A. (2020), Sustentabilidade corporativa: definição de indicadores para organizações do setor energético, *Revista de Gestão dos Países de Língua Portuguesa*, 19, 2, 104-126.

PHILLIS, A.; GRIGOROUDIS, E.; KOUIKOGLU, V.S. (2020), Assessing national energy sustainability using multiple criteria decision analysis, *International Journal of Sustainable Development and World Ecology*, 1-18.

PONOMARENKO, T.; RESHNEVA, E.; URBANO, A. P. M. (2022), Assessment of Energy Sustainability Issues in the Andean Community: Additional Indicators and Their Interpretation, *Energies*, 15, 3, DOI: 10.3390/en15031077.

PRODANOV, C. C.; FREITAS, E. C. (2013), *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico*, Universidade Feevale, 2ª ed.

PROVOST, F.; FAWCETT, T. (2013), *Data science for business: what you need to know about data mining and data-analytic thinking*. O'Reilly Media, Inc., Sebastopol (USA).

QI, W.; HUANG, Z.; DINÇER, H.; KORSAKIENE, R.; YUKSEL, S. (2020), Corporate governance-based strategic approach to sustainability in energy industry of emerging economies with a novel interval-valued intuitionistic fuzzy hybrid decision making model, *Sustainability (Switzerland)*, 12.

RAFAJ, P.; BARRETO, L.; KYPREOS, S. (2006), Combining policy instruments for sustainable energy systems: An assessment with the GMM model, *Environmental Modeling and Assessment*, 11, 277-295.

RAMAIAH, V.; JAYASANKAR, V. (2022), Performance Assessment of Indian Electric Distribution Utilities Using Data Envelopment Analysis (DEA), *International Journal of Electrical and Electronic Engineering & Telecommunications*, 11, 3, DOI: 10.18178/ijeetc.11.3.192-202.

RIGO, P.D.; REDISKE, G.; ROSA, C.B.; GASTALDO, N.G.; MICHELS, L.; JUNIOR, A.L.N.; SILUK, J.C.M. (2020), Renewable energy problems: Exploring the methods to support the decision-making process, *Sustainability (Switzerland)*, 12, 1-27.

RIZA, L. S.; JANUSZ, A.; BERGMEIR, C.; CORNELIS, C.; HERRERA, F.; SLEZAK, D.; BENITEZ, J. M. (2014), Implementing algorithms of rough set theory and fuzzy rough set theory in the R package "roughsets", *Information Sciences*, 287, 68-89.

RIZA, L.S.; JANUSZ, A.; SLEZAK, D.; CORNELIS, C.; HERRERA, F.; BENITEZ, J.M.; BERGMEIR, C.; STAWICKI, S. (2019), *Data Analysis Using Rough Set and Fuzzy Rough Set Theories*, Acesso: < <https://cran.r-project.org/web/packages/RoughSets/RoughSets.pdf>>, em 12/03/2021.

ROSA, C.B.; SILUK, J.C.M.; RIGO, P.D.; REDISKE, G.; BURIN, H.P.; MICHELS, L. (2020), Organizational Performances of Distributed Generation in Brazil Electric Utilities: A Balanced Scorecard Perspective, *Springer Proceedings in Mathematics and Statistics*, 337, 101-114.



ROY, B. (1968), Classement et choix en présence de points de vue multiples (la méthode ELECTRE). *La Revue d'Informatique et de Recherche Opérationnelle (RIRO)*, 8: 57–75.

ROY, B. (1989), Main sources of inaccurate determination, uncertainty and imprecision in decision models, *Mathematical and Computer Modelling*, 2, 1245-1254.

ROY, B., BOUYSSOU, D., (1991), Decision-aid: an elementary introduction with emphasis on multiple criteria, *Cahier du Lamsade, Université Paris-Dauphine*, 106, 18.

RUSSELL, S. (2021), *Inteligência artificial a nosso favor: como manter o controle sobre a tecnologia*, São Paulo: Companhia das Letras, 1ª ed.

RUSSELL, S.; NORVIG, P. (2010), *Artificial intelligence: A modern approach*, New Jersey: Prentice Hall, 3rd ed.

SAHABUDDIN, M.; KHAN, I. (2021), Multi-criteria decision analysis methods for energy sector's sustainability assessment: Robustness analysis through criteria weight change, *Sustainable Energy Technologies and Assessments*, 47, <https://doi.org/10.1016/j.seta.2021.101380>.

SALMI, S.; SUKUR, A.L.A.; NORLENA, H. (2019), The need of corporate social responsibility (CSR) implementation in energy industry: Proposition development, *International Journal of Recent Technology and Engineering*, 8, 201-207.

SARANGI, G.K.; MISHRA, A.; CHANG, Y.; TAGHIZADEH-HESARY, F. (2019), Indian electricity sector, energy security and sustainability: An empirical assessment, *Energy Policy*, 135.

SARAVANAN, R.; SUJATHA, P. (2018), A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification, *Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018)*, IEEE Xplore Compliant Part Number: CFP18K74-ART, 945-949.

SARTORI, S.; WITJES, S.; CAMPOS, L. M. S. (2017), Sustainability performance for Brazilian electricity power industry: an assessment integrating social, economic and environmental issues, *Energy Policy*, 111, 41–51.

SHAHBAZ, M.; KARAMAN, A.S.; KILIC, M.; UYAR, A. (2020), Board attributes, CSR engagement, and corporate performance: What is the nexus in the energy sector?, *Energy Policy*, 143.

SI, K.; XU, X.L.; CHEN, H.H. (2020), Examining the interactive endogeneity relationship between R&D investment and financially sustainable performance: Comparison from different types of energy enterprises, *Energies*, 13.

SILGE, J.; ROBINSON, D. (2017), *Text mining with R: A tidy approach*, O'Reilly Media, USA.

SILVEIRA, M. A. (2013), Strategic management of innovation towards sustainable development of brazilian electronic, *Journal Technologic Management & Innovation*, 8, s/n, 174-186.

SLACK, N.; BRANDON-JONES, A. (2021), *Operations and process management: Principles and practice for strategic impact*, United Kingdom: Pearson Education Limited, 6<sup>th</sup> ed.

SLOWINSKI, R.; GRECO, S.; MATARAZZO, B. (2012), Rough set and rule-based multicriteria decision aiding, *Pesquisa Operacional*, 32, 2, 213-269.

SUEYOSHI, T.; GOTO, M. (2014), Environmental assessment for corporate sustainability by resource utilization and technology innovation: DEA radial measurement on Japanese industrial sectors, *Energy Economics*, 46, 295-307.

SUEYOSHI, T.; GOTO, M. (2017), World trend in energy: an extension to DEA applied to energy and environment, *Journal of Economic Structures*, 6.

SUEYOSHI, T.; RYU, Y.; YUN, J-Y (2021), Covid-19 response and prospects of clean/sustainable energy transition in industrial nations: New environmental assessment, *Energies*, 14, 1174, <https://doi.org/10.3390/en14041174>.

SUN, J.; XU, S.; LI, G. (2020), Analyzing sustainable power supply chain performance: Evidence from China's provincial regions, *Journal of Enterprise Information Management*, DOI 10.1108/JEIM-09-2019-0296.

SWETS, J. A. (1988), Measuring the accuracy of diagnostic systems, *Science*, 3, 240 (4857), 1285-1293, doi: 10.1126/science.3287615.

TAJBAKSHI, A.; SHAMSI, A. (2019), A facility location problem for sustainability-conscious power generation decision makers, *Journal of Environmental Management*, 230, 319 - 334.

TALBOT, D.; BOIRAL, O. (2018), GHG reporting and impression management: An assessment of sustainability reports from the energy sector, *Journal of Business Ethics*, 147, 367-383.

TALEB, N. N. (2010), *A lógica do Cisne Negro: O impacto do altamente improvável*, Rio de Janeiro: Ed. Schwarcz S.A.

THAKKAR, J. J. (2021), *Multi-Criteria Decision Making*, *Studies in Systems, Decision and Control*, 336, Singapore: Springer Nature Singapore Pte Ltd.

VADEYAR, D. A.; YOGISH, H. K. (2014), Farthest First Clustering in Links Reorganization, *International Journal of Web & Semantic Technology (IJWesT)*, 5, 3.

VAN ECK, N. J.; WALTMAN, L. (2021), VOSviewer, Acesso: <<https://www.vosviewer.com/>>, em 11/12/2021.

VAN GERVEN, T.; BLOCK, C.; GEENS, J.; CORNELIS, G.; VANDECASTEELE, C. (2007), Environmental response indicators for the industrial and energy sector in Flanders, *Journal of Cleaner Production*, 15, 886-894.

VIANA, M. (2022), O que é a probabilidade?: interpretação mais popular tem problemas, *Folha de S. Paulo*, 05/10/2022, B8.

VIEIRA, R.; O'DWYER, B.; SCHNEIDER, R. (2017), Aligning Strategy and Performance Management Systems: The Case of the Wind-Farm Industry, *ORGANIZATION & ENVIRONMENT*, 30, 3-26.

VINUTHA, H. P.; POORNIMA, B.; SAGAR, B. M. (2018), Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset. In: Satapathy, S., Tavares, J., Bhateja, V., Mohanty, J. (Eds) *Information and Decision Sciences, Advances in Intelligent Systems and Computing*, 701, Singapore: Springer, [https://doi.org/10.1007/978-981-10-7563-6\\_53](https://doi.org/10.1007/978-981-10-7563-6_53).

VLUYMANS, S. (2019), *Dealing with Imbalanced and Weakly Labelled Data in Machine Learning using Fuzzy and Rough Set Methods*, Switzerland: Springer Nature.

WANG, D.; LI, S.; SUEYOSHI, T. (2014), DEA environmental assessment on U.S. industrial sectors: Investment for improvement in operational and environmental performance to attain corporate sustainability, *Energy Economics*, 45, 254 - 267.

WANG, Y.; LI, S.; WU, X.; ZHANG, Y.; LI, B.; GAO, L. (2021), Using sustainable performance prediction in data-scarce scenarios: A study of park-level integrated microgrid projects in Tianjin, China, *Journal of Cleaner Production*, 304, <https://doi.org/10.1016/j.jclepro.2021.127042>.

WANKE, P.; TAN, Y.; ANTUNES, J.; HADI-VENCHEH, A. (2020), Business environment drivers and technical efficiency in the Chinese energy industry: A robust Bayesian stochastic frontier analysis, *Computers and Industrial Engineering*, 144.

WCED (1987), *World Commission on Environment and Development, Our common future*. Oxford University Press, Oxford.

WITTEN, I. H.; FRANK, E. (2005), *Data mining: practical machine learning tools and techniques*, 2<sup>a</sup> ed., San Francisco: Elsevier.

XU, X.-L.; CHEN, H.H. (2020), Exploring the relationships between environmental management and financial sustainability in the energy industry: Linear and nonlinear effects, *Energy and Environment*, 31, 1281-1300.

XU, X.-L.; LIU, C.K. (2019), How to keep renewable energy enterprises to reach economic sustainable performance: From the views of intellectual capital and life cycle, *Energy, Sustainability and Society*, 9, <https://doi.org/10.1186/s13705-019-0187-2>.

YALCIN, A. S.; KILIC, H. S.; DELEN, D. (2022), The use of multi-criteria decision-making methods in business analytics: A comprehensive literature review,

Technological Forecasting & Social Change, 174, 121193.  
<https://doi.org/10.1016/j.techfore.2021.121193>.

YANG, L.; ZHANG, Y. (2020), An evaluation of the green performance of Chinese new energy enterprises from the perspective of social responsibility, *Nature Environment and Pollution Technology*, 19, 1527-1535.

YORO, K. O.; DARAMOLA, M. O.; SEKOAI, P. T.; WILSON, U. N.; ETERIGHO-IKELEGBE, O. (2021), Update on current approaches, challenges, and prospects of modeling and simulation in renewable and sustainable energy systems, *Renewable and Sustainable Energy Reviews*, 150, <https://doi.org/10.1016/j.rser.2021.111506>.

YUSIFBAYLI, N.; NASIBOV, V. (2020), Comparative analysis of Azerbaijan's energy sector efficiency trend at the current development stage, *E3S Web of Conferences*, 209.

ZAKI, M. J.; MEIRA JR., W. (2020), *Data mining and machine learning: Fundamental concepts and algorithms*, United Kingdom: TJ International Ltd., 2<sup>nd</sup> ed.

ZHAI, C.; MASSUNG, S. (2016), *Text data management and analysis: A practical introduction to information retrieval and text mining*, Association for Computing Machinery and Morgan & Claypool Publishers, University of Waterloo.

ZHAO, S. Y.; TSANG, E. C. C.; CHEN, D. G.; WANG, X. Z. (2010), Building a rule-based classifier – a fuzzy-rough set approach, *IEEE Transactions on Knowledge and Data Engineering*, 22, 5, 624–638.

ZHOU, P.; ZHOU, P.; YUKSEL, S.; DINCER, H.; ULUER, G. S. (2019), Balanced scorecard-based evaluation of sustainable energy investment projects with IT2 fuzzy hybrid decision making approach, *Energies*, 13.

ZIARKO, W. (1993), Analysis of uncertain information in the framework of variable precision rough sets, *Foundations of Computing and Decision Sciences*, 18, 3-4, 381-396.

ZIZKA, J.; DARENA, F.; SVOBODA, A. (2020), *Text mining with machine learning: principles and techniques*, Taylor & Francis Group, Boca Raton (USA).

## ANEXO I-a

27th International Sustainable Development Research Society Conference, Mid Sweden University, 13 – 15 July 2021

### **Sustainability and performance model of evaluation proposition for companies in the Brazilian electric sector based on multicriteria decision aiding and machine learning**

Ayrton Benedito Gaia do Couto<sup>1</sup>, Luís Alberto Duncan Rangel<sup>2</sup>

<sup>1</sup>Fluminense Federal University, [ayrtoncouto@gmail.com](mailto:ayrtoncouto@gmail.com)

<sup>2</sup>Fluminense Federal University, [luisduncan@id.uff.br](mailto:luisduncan@id.uff.br)

#### **Abstract**

This research aims to develop an evaluation model for electricity distribution concessionaires companies in Brazil, based on sustainability and performance indicators, aiming a company benchmarking. The proposed model is based on a multicriteria decision aiding with the Rough Sets Theory, which allows the mathematic treatment of inaccuracies originated from quantitative and qualitative measurements of the indicators. The indicators are extracted from Socio-environmental and Economic-Financial Responsibility Report, produced by each concessionaire company. Based on “decision tables” with condition attributes and decision classes from indicators and indexes, the decision rules “if... then...” are generated, in addition to the establishment of high, medium or low sustainability and performance classes. Furthermore, through Machine Learning, a subfield of Artificial Intelligence and classification algorithms, it will be possible to predict the sustainability and performance classes based on historical series of indicators. For each company a matching ordered pair will be formed “class-performance, class-sustainability”, to be pointed out in a “Matrix of 9 sectors”. This positioning will always be relative (or normalized), that is, it will have as reference the companies best positioned or classified in the matrix. The benchmarking will take place throughout a historical series of indicators and indexes of sustainability and performance for a single enterprise, as it will in relation to unidentified competitors. As benchmarking results, the use of the proposed model subsidizes an enterprise in establishing its goals, in its development of action plans, in its control of indicators and indexes, and in managing risks. For the public manager, it becomes an opportunity to better understand the best practices of sustainability and performance of enterprises considered benchmark, aiming to elaborate and/or improve public policy of sustainable development for the electricity sector. In the future, data about the companies will be assembled in a World Wide Web platform, as a mean to support the proposed model.

**Keywords:** Sustainability, Performance, Evaluation Model, Rough Sets Theory, Machine Learning

## ANEXO I-b



# Book of Abstracts

the 27th Annual Conference, International  
Sustainable Development Research Society



108

## A model for evaluating the sustainability and performance of companies in the Brazilian electric sector based on multicriteria decision and machine learning

PhD Student Ayrton Benedito Gaia do Couto, PhD Luis Alberto Duncan Rangel  
Federal Fluminense University, Niteroi-RJ, Brazil

### Abstract

The present study is aimed at proposing a model for joint assessment of sustainability and performance of companies in the Brazilian electricity sector, based on sustainability and performance indicators, aiming at benchmarking companies in relation to competing companies. As these are indicators based on quantitative and qualitative measurements and, most likely, they will incur inaccuracies in these measurements, the model to be proposed will be based on a multicriteria approach with the use of Rough Sets Theory, which allows the mathematical treatment of inaccuracies. From this theory, Decision Tables are then generated considering: condition attributes (sustainability or performance indicators) and decision attributes (sustainability or performance classes).

The sustainability or performance classes are obtained by "grouping" ("clustering"), considering three "clusters" or classes: "high, medium, low", based on sustainability or performance indexes, respectively. In addition, through Machine Learning (Artificial Intelligence subfield; and classification algorithms), it will be possible to predict the sustainability or performance classes (high, medium, low) based on historical series of indicators and indexes. For each company, an orderly pair (sustainability class, performance class) will be formed, to be pointed out in a "Matrix of 9 sectors". This positioning obtained by companies in the Brazilian electricity sector will always be relative, that is, it will have as reference the companies best positioned or classified in the matrix. Benchmarking will be carried out over a historical series of indicators and indexes of sustainability or performance for the same company, as well as, in relation to competing companies (not identified). Thus, the use of the proposed model will allow a company to identify and quantify its gaps in terms of indicators and indexes of sustainability and performance in a relative way compared to its competitors. This will provide inputs for strategic planning (action plan; monitoring of indicators and indexes) for the company under analysis, given the set of information gathered about companies in the electricity sector. Information about companies will be exported to a Web platform (World Wide Web), as a means of supporting the proposed model.

Full Paper upload pdf (if accepted as poster or paper presentation; for details see <https://2021.isdrconferences.org/call-for-papers/>)

[Uploaded file](#)

## ANEXO II-a

### PROPOSIÇÃO DE MODELO DE AVALIAÇÃO DA SUSTENTABILIDADE E DESEMPENHO DE EMPRESAS DO SETOR ELÉTRICO BRASILEIRO COM BASE EM APOIO MULTICRITÉRIO E MACHINE LEARNING

**Ayrton Benedito Gaia do Couto**

Universidade Federal Fluminense

R. Passo da Pátria, 156 - Bloco E, Sala 324 - São Domingos, Niterói - RJ  
ayrtoncouto@gmail.com

**Luís Alberto Duncan Rangel**

Universidade Federal Fluminense

R. Passo da Pátria, 156 - Bloco E, Sala 324 - São Domingos, Niterói - RJ  
luisduncan@id.uff.br

#### RESUMO

Esta pesquisa visa a elaboração de modelo para avaliação de empresas concessionárias de distribuição de energia elétrica no Brasil, com base em indicadores de sustentabilidade e desempenho, objetivando o *benchmarking* das empresas. O modelo proposto fundamenta-se em abordagem multicritério com a Teoria dos Conjuntos Aproximativos, que permite o tratamento matemático das imprecisões advindas de medições quantitativas e qualitativas dos indicadores. Com base em tabelas de decisão, são geradas regras de decisão (“se ... então ...”) a partir de indicadores e índices, além do estabelecimento de classes de sustentabilidade e desempenho alta, média ou baixa. Por meio de *Machine Learning*, será possível a predição das classes de sustentabilidade e de desempenho com base em séries históricas de indicadores. Para cada empresa, formar-se-á um par ordenado considerando: a classe-desempenho e a classe-sustentabilidade correspondente, a ser apontado em uma “Matriz de 9 setores”.

**PALAVRAS CHAVE:** Sustentabilidade, Modelo de avaliação, Teoria dos Conjuntos Aproximativos.

**Tópicos; Apoio à Decisão Multicritério**

#### ABSTRACT

This research aims to develop an evaluation model for electricity distribution concessionaires companies in Brazil, based on sustainability and performance indicators, aiming a company benchmarking. The proposed model is based on a multicriteria decision aiding with the Rough Sets Theory, which allows the mathematic treatment of inaccuracies originated from quantitative and qualitative measurements of the indicators. Based on decision tables, the decision rules are generated (“if... then...”) from indicators and indexes, in addition to the establishment of high, medium, or low sustainability and performance classes. Furthermore, through Machine Learning it will be possible to predict the sustainability and performance classes based on historical series of indicators. For each company, a matching ordered pair will be formed: a class-performance and a class-sustainability, to be pointed out in a “Matrix of 9 sectors”.

**KEYWORDS:** Sustainability, Evaluation Model, Rough Sets Theory.

**Paper topics: Multicriteria Decision Aiding**

## ANEXO II-b






## ANEXO III



GESTÃO  
&  
PRODUÇÃO

ISSN 0104-530X (Print)  
ISSN 1806-9649 (Online) 

ORIGINAL  
ARTICLE

## Model proposition for predicting sustainability classes using multicriteria decision support and artificial intelligence

*Proposição de modelo para a predição de classes de sustentabilidade por meio de apoio à decisão multicritério e inteligência artificial*

Ayrton Benedito Gaia do Couto<sup>1</sup> , Luís Alberto Duncan Rangel<sup>1</sup> 

<sup>1</sup>Universidade Federal Fluminense – UFF, Escola de Engenharia, Laboratório de Tecnologia e Gestão de Negócios, Programa de Doutorado em Sistemas de Gestão Sustentáveis, Niterói, RJ, Brasil. E-mail: ayrtoncouto@gmail.com; ayrtoncouto@id.uff.br; luisduncan@id.uff.br

**How to cite:** Couto, A. B. G., & Rangel, L. A. D. (2022). Model proposition for predicting sustainability classes using multicriteria decision support and artificial intelligence. *Gestão & Produção*, 29, e6922. <https://doi.org/10.1590/1806-9649-2022v29e6922>

**Abstract:** The current study proposes a novel prediction model of sustainability classes for electricity distribution companies in Brazil, based on sustainability indicators, aiming at a more effective risk management for a certain company among their competitors. Because such indicators are based on quantitative and qualitative measures and are very likely to incur imprecisions in their measures, the model to be proposed is based on a Multicriteria Decision Support, Rough Sets Theory, which allows the mathematical treatment of those imprecisions, and Artificial Intelligence, in this case, Machine Learning by rules inference. Consequently, decision tables are generated with condition attributes, sustainability indicators, and decision attributes, sustainability classes: high, medium or low. As a result, it is possible to predict sustainability classes based in temporal series of indicators and rules inference from decision tables, using RoughSets package in R and the jMAF software, demonstrating the use of five rule generation algorithms and their respective accuracies.

**Keywords:** Forecasting; Multicriteria decision; Rough sets theory; Artificial intelligence; Sustainability; Risk management.

**Resumo:** O presente estudo propõe um novo modelo de previsão de classes de sustentabilidade para empresas de distribuição de energia elétrica no Brasil, com base em indicadores de sustentabilidade, visando uma gestão de risco mais eficaz para uma determinada empresa frente a seus concorrentes. Como tais indicadores são baseados em medidas quantitativas e qualitativas e são muito propensos a incorrer em imprecisões em suas medidas, o modelo a ser proposto baseia-se em Apoio à Decisão Multicritério, a Teoria dos Conjuntos Aproximativos, que permite o tratamento matemático destas imprecisões, e a Inteligência Artificial, neste caso, Aprendizado de Máquina, por inferência de regras. Consequentemente, são geradas tabelas de decisão com atributos de condição, indicadores de sustentabilidade, e atributos de decisão, classes de sustentabilidade: alta, média ou baixa. Como resultado, é possível prever classes de sustentabilidade com base em séries temporais de indicadores e inferência de regras a partir das tabelas de decisão, utilizando-se o pacote RoughSets em R e o software jMAF, e demonstrando-se a aplicação de cinco algoritmos de geração de regras e suas respectivas precisões.

**Palavras-chave:** Previsão; Decisão multicritério; Teoria dos conjuntos aproximativos; Inteligência artificial; Sustentabilidade; Gestão de riscos.

Received Aug. 25, 2022 - Accepted Oct. 10, 2022  
Financial support: None.




This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ANEXO IV



GESTÃO  
&  
PRODUÇÃO

ISSN 0104-530X (Print)  
ISSN 1806-9649 (Online) 

ORIGINAL  
ARTICLE

## Corporate governance aspects in sustainability and performance assessment models for energy companies: a systematic review of the literature

*Aspectos de governança corporativa em modelos de avaliação da sustentabilidade e desempenho para empresas de energia: uma revisão sistemática da literatura*

Ayrton Benedito Gaia do Couto<sup>1</sup> , Luis Alberto Duncan Rangel<sup>1</sup> 

<sup>1</sup> Universidade Federal Fluminense – UFF, Programa de Pós-graduação em Sistemas de Gestão Sustentáveis, Escola de Engenharia, Laboratório de Tecnologia e Gestão de Negócios, Niterói, RJ, Brasil. E-mail: ayrtoncouto@gmail.com; ayrtoncouto@id.uff.br; luisduncan@id.uff.br

**How to cite:** Couto, A. B. G., & Rangel, L. A. D. (2023). Corporate governance aspects in sustainability and performance assessment models for energy companies: a systematic review of the literature. *Gestão & Produção*, 30, e4322. <https://doi.org/10.1590/1806-9649-2022v29e4322>

**Abstract:** The present study aims to map the sustainability and performance evaluation models for energy companies regarding their qualitative and quantitative aspects of corporate governance. This research was carried out through a systematic literature review and the establishment of connections between the issues raised and general aspects of corporate governance, vis-à-vis the sustainable development matter. Aspects of corporate governance, such as gender diversity in the creation of boards of directors and directors in companies, is relevant in view of the UN Sustainable Development Goals, as well as the ISO 26000 standard, regarding guidelines on corporate social responsibility. As an innovation in this mapping, a summary of the main qualitative and quantitative aspects of corporate governance is carried out using the clustering technique and co-occurrence map, mining and text analysis, compared to sustainability and performance evaluation models for energy companies. At the end, gaps in these models are pointed out referring to national and international organizations and studies regarding intersectoral principles, practices and tools of corporate governance; opportunities for improvement are also pointed out.

**Keywords:** Corporate governance; Sustainability; Evaluation model; Performance; Energy sector.

**Resumo:** O presente estudo visa mapear os modelos de avaliação da sustentabilidade e desempenho para empresas de energia quanto aos seus aspectos qualitativos e quantitativos de governança corporativa. Esta pesquisa foi realizada por meio de revisão sistemática da literatura e o estabelecimento de conexões entre os assuntos levantados e aspectos gerais de governança corporativa, vis-à-vis à questão do desenvolvimento sustentável. Aspecto de governança corporativa, como a diversidade de gênero na formação de conselhos de administração e diretorias em empresas, mostra-se relevante frente aos Objetivos de Desenvolvimento Sustentável da ONU, bem como, o padrão ISO 26000, a respeito das diretrizes sobre responsabilidade social corporativa. Como inovação neste mapeamento, realiza-se uma síntese visual dos aspectos qualitativos e quantitativos principais de governança corporativa com o uso da técnica de *clustering* e mapa de co-ocorrência, de mineração e análise de textos, frente a modelos de avaliação da sustentabilidade e desempenho para empresas de energia. Ao final, são apontadas lacunas desses modelos com referência a organismos nacionais e internacionais e estudos no que diz respeito a princípios, práticas e ferramentas intersectoriais de governança corporativa; são apontadas também, oportunidades de melhorias.

**Palavras-chave:** Governança corporativa; Sustentabilidade; Modelo de avaliação; Desempenho; Setor de energia.

Received Nov. 17, 2022 - Accepted Dec. 18, 2022  
Financial support: None.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ANEXO V

Algoritmo VC-DomLEM (BLASZCZYNSKI *et al.*, 2009)**Algorithm 1:** *VC-DomLEM*

**Input** : set  $\mathbf{X}$  of classes  $X_i \in U$ , upward unions of classes  $X_i^{\geq} \in U$  or downward unions of classes  $X_i^{\leq} \in U$ ,  
 set of attributes  $P \subseteq C$ ,  
 rule consistency measure  $\hat{\theta}$ ,  
 set of rule consistency measure thresholds  $\{\hat{\theta}_X : X \in \mathbf{X}\}$ ,  
 object covering option  $s$ .

**Output:** set of rules  $\mathbf{R}$ .

```

1  $\mathbf{R} := \emptyset$ ;
2 foreach element  $X \in \mathbf{X}$  do
3    $AO_P^{\theta_X}(X) := \text{AllowedObjects}(X, P, \theta_X, s)$ ;
4    $R_X^{\hat{\theta}_X} := \text{VC-SequentialCovering}^{\text{mix}}(\underline{P}^{\theta_X}(X), AO_P^{\theta_X}(X), P, \hat{\theta}, \hat{\theta}_X)$ ;
5    $\mathbf{R} := \mathbf{R} \cup R_X^{\hat{\theta}_X}$ ;
6    $\text{RemoveNonMinimalRules}(\mathbf{R})$ ;
7 end
```

**Algorithm 2:** *VC-SequentialCovering<sup>mix</sup>*

**Input** : set of positive objects  $\underline{P}^{\theta_X}(X) \subseteq U$ ,  
 set of objects that can be covered  $AO_P^{\theta_X}(X) \subseteq U$ ,  
 $AO_P^{\theta_X}(X) \supseteq \underline{P}^{\theta_X}(X)$   
 set of attributes  $P \subseteq C$ ,  
 rule consistency measure  $\hat{\theta}$ ,  
 rule consistency measure threshold  $\hat{\theta}_X$ .

**Output:** set of rules  $R_X^{\hat{\theta}_X}$  assigning objects to  $X$ .

```

1  $B := \underline{P}^{\theta_X}(X)$ ;
2  $R_X^{\hat{\theta}_X} := \emptyset$ ;
3 while  $B \neq \emptyset$  do
4    $r_X^{\hat{\theta}_X} := \emptyset$ ;
5    $EC := \text{ElementaryConditions}(B, P)$ ;
6   while ( $\hat{\theta}(r_X^{\hat{\theta}_X})$  does not satisfy  $\hat{\theta}_X$ ) or ( $\|\Phi_{r_X^{\hat{\theta}_X}}\| \not\subseteq AO_P^{\theta_X}(X)$ ) do
7      $ec := \text{BestElementaryCondition}(EC, r_X^{\hat{\theta}_X}, \hat{\theta}, \underline{P}^{\theta_X}(X))$ ;
8      $r_X^{\hat{\theta}_X} := r_X^{\hat{\theta}_X} \cup ec$ ;
9      $EC := \text{ElementaryConditions}(B \cap \text{supp}(r_X^{\hat{\theta}_X}), P)$ ;
10  end
11   $\text{RemoveRedundantElementaryConditions}(r_X^{\hat{\theta}_X}, \hat{\theta}, \hat{\theta}_X, AO_P^{\theta_X}(X))$ ;
12   $R_X^{\hat{\theta}_X} := R_X^{\hat{\theta}_X} \cup r_X^{\hat{\theta}_X}$ ;
13   $B := B \setminus \text{supp}(r_X^{\hat{\theta}_X})$ ;
14 end
15  $\text{RemoveRedundantRules}(R_X^{\hat{\theta}_X}, \hat{\theta}, \underline{P}^{\theta_X}(X))$ 
```

## ANEXO VI

### Regras de decisão dirSU inferidas pelo algoritmo VC-DomLEM, jMAF

#### [FILEINFO]

Program: jrs  
Algorithm: VC-DOMLEM  
DataFileDirectory: D:\2023-07-17\dirSU  
DataFileName: dados\_dirSU.isf  
Objects: 342  
Decision: CLASSE  
Duration: 2230 [ms]  
Comment: Rules generated with VC-DOMLEM algorithm

#### [PARAMETERS]

RulesFormat: INFORMATION\_TABLE\_PARETO  
ConsistencyLevel: 0.0  
Classes: D1, D2, D3, D4, D5

#### [ATTRIBUTES]

+ codigo\_empresa: (nominal), none, description  
+ ano: (nominal), none, description  
+ IE01: (continuous), gain  
+ IE02: (continuous), cost  
+ IE03: (continuous), cost  
+ IE04: (continuous), gain  
+ IE05: (continuous), gain  
+ IE06: (continuous), cost  
+ IE07: (continuous), gain  
+ IE08: (continuous), gain  
+ IE09: (continuous), gain  
+ IE10: (continuous), gain  
+ IS01: (continuous), gain  
+ IS02: (continuous), gain  
+ IS03: (continuous), gain  
+ IS04: (continuous), gain  
+ IS05: (continuous), gain  
+ IS06: (continuous), gain  
+ IS07: (continuous), gain  
+ IS08: (continuous), gain  
+ IS09: (continuous), gain  
+ IS10: (continuous), cost  
+ IA01: (continuous), cost  
+ IA02: (continuous), cost  
+ IA03: (continuous), cost  
+ IG01: (continuous), gain  
+ IG02: (continuous), gain  
+ IG03: (continuous), gain  
+ IG04: (continuous), gain  
+ IG05: (continuous), gain  
+ IG06: (continuous), gain  
+ IG07: (continuous), gain  
+ IG08: (continuous), gain  
+ IG09: (continuous), gain  
+ IG10: (continuous), gain  
+ CLASSE: [D1, D2, D3, D4, D5], gain, decision

decision: CLASSE

#### [PREFERENCES]

codigo\_empresa: none  
 ano: none  
 IE01: gain  
 IE02: cost  
 IE03: cost  
 IE04: gain  
 IE05: gain  
 IE06: cost  
 IE07: gain  
 IE08: gain  
 IE09: gain  
 IE10: gain  
 IS01: gain  
 IS02: gain  
 IS03: gain  
 IS04: gain  
 IS05: gain  
 IS06: gain  
 IS07: gain  
 IS08: gain  
 IS09: gain  
 IS10: cost  
 IA01: cost  
 IA02: cost  
 IA03: cost  
 IG01: gain  
 IG02: gain  
 IG03: gain  
 IG04: gain  
 IG05: gain  
 IG06: gain  
 IG07: gain  
 IG08: gain  
 IG09: gain  
 IG10: gain  
 CLASSE: gain

[RULES]

#Certain at least rules

- 1: (IA01 <= 0.204) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|
- 2: (IE07 >= 76328.0) & (IS04 >= 22.52) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|
- 3: (IE03 <= 276.7) & (IS10 <= 0.17) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|
- 4: (IE02 <= 25751.0) & (IE07 >= 71651.0) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|
- 5: (IS04 >= 30.26) & (IA02 <= 0.1) => (CLASSE >= D5) |CERTAIN, AT\_LEAST, D5|
- 6: (IG05 >= 1.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 7: (IS01 >= 5394.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 8: (IG04 >= 1.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 9: (IE07 >= 21565.0) & (IG08 >= 1.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 10: (IE04 >= 8.0) & (IA03 <= 1150.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 11: (IE01 >= 1158036.0) & (IS06 >= 2789.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 12: (IE03 <= 1669.0) & (IS10 <= 0.23) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 13: (IA01 <= 6.92) & (IA02 <= 9.89) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 14: (IS03 >= 25.85) & (IA01 <= 6.92) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 15: (IA03 <= 1216.0) & (IG06 >= 1.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 16: (IE02 <= 588.0) & (IA01 <= 6.69) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 17: (IS03 >= 32.78) & (IA01 <= 155945.44) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 18: (IE06 <= 362256.0) & (IS10 <= 0.27) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 19: (IS03 >= 19.32) & (IA02 <= 0.1) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 20: (IS08 >= 71666.0) & (IG01 >= 1.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|
- 21: (IS01 >= 1408.0) & (IS04 >= 30.6) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|

22: (IE01 >= 8337594.0) & (IS01 >= 4350.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|  
 23: (IE01 >= 6205000.0) & (IS08 >= 54423.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|  
 24: (IA03 <= 1838.0) & (IG09 >= 1.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|  
 25: (IE07 >= 475.0) & (IA03 <= 1838.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|  
 26: (IA02 <= 2459.36306880164) & (IA03 <= 2204.5) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|  
 27: (IE04 >= 66.0) & (IA03 <= 3384.0) => (CLASSE >= D4) |CERTAIN, AT\_LEAST, D4|  
 28: (IE07 >= 16510.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 29: (IS10 <= 0.83) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 30: (IA01 <= 7.53) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 31: (IS04 >= 30.26) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 32: (IE01 >= 1.6650141E7) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 33: (IS05 >= 19.54) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 34: (IS03 >= 27.81) & (IS06 >= 12.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 35: (IE10 >= 256.0) & (IS03 >= 27.81) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 36: (IS05 >= 0.8) & (IA03 <= 1949.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 37: (IS07 >= 56829.0) & (IG01 >= 1.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 38: (IE04 >= 8.0) & (IA03 <= 1350.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 39: (IS01 >= 1123.0) & (IS06 >= 2603.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 40: (IS03 >= 20.27) & (IS06 >= 2615.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 41: (IS06 >= 916.0) & (IG10 >= 1.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 42: (IS03 >= 24.2) & (IA03 <= 2204.5) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 43: (IS10 <= 7.22) & (IA02 <= 0.11) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 44: (IA02 <= 0.1) & (IG06 >= 1.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 45: (IE07 >= 317.0) & (IA03 <= 3384.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 46: (IS08 >= 28694.0) & (IG09 >= 1.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 47: (IS05 >= 1.7) & (IS09 >= 2769.0) & (IG08 >= 1.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 48: (IE03 <= 4.54) & (IE04 >= 67.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 49: (IE01 >= 8695421.0) & (IS01 >= 3955.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 50: (IE06 <= 207835.0) & (IS09 >= 2083.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 51: (IE05 >= 6773.33) & (IE06 <= 531624.0) & (IS03 >= 24.71) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 52: (IE01 >= 12104.0) & (IE03 <= 4.87) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 53: (IS01 >= 3614.0) & (IS07 >= 39816.0) & (IG01 >= 1.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 54: (IE02 <= 15646.0) & (IS01 >= 3708.0) => (CLASSE >= D3) |CERTAIN, AT\_LEAST, D3|  
 55: (IE10 >= 473.08) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 56: (IS03 >= 25.08) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 57: (IA02 <= 157.05) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 58: (IE06 <= 13745.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 59: (IA01 <= 2550.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 60: (IG09 >= 1.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 61: (IE03 <= 4.87) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 62: (IS05 >= 6.5) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 63: (IE05 >= 7042.73) & (IE07 >= 2563.495) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 64: (IS05 >= 2.86) & (IA03 <= 3887.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 65: (IE03 <= 17.0) & (IA03 <= 17065.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 66: (IS03 >= 15.4) & (IS10 <= 1.21) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 67: (IS02 >= 50.42) & (IS10 <= 28.33) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 68: (IE02 <= 215.0) & (IA01 <= 54803.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 69: (IE07 >= 310.9) & (IS06 >= 1323.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 70: (IE01 >= 1404298.0) & (IG01 >= 1.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 71: (IE07 >= 960.0) & (IG01 >= 1.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 72: (IE10 >= 460.0) & (IS10 <= 8.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 73: (IS05 >= 3.1) & (IA03 <= 6523.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 74: (IS08 >= 4718.0) & (IA03 <= 5500.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 75: (IE07 >= 735.0) & (IS02 >= 44.1) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 76: (IS09 >= 361.0) & (IA03 <= 6727.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 77: (IE07 >= 130.91) & (IS04 >= 19.09) & (IA03 <= 6674.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 78: (IS06 >= 723.0) & (IG06 >= 1.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|



79: (IS06 >= 46.0) & (IS10 <= 6.18) & (IA03 <= 7320.0) => (CLASSE >= D2) |CERTAIN, AT\_LEAST, D2|  
 #Certain at most rules  
 80: (IE01 <= 3049218.0) & (IE06 >= 747127.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 81: (IS10 >= 39.09) & (IA03 >= 13860.0) & (IG06 <= 0.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 82: (IE07 <= 391.0) & (IE09 <= 1.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 83: (IE05 <= 2716.1) & (IS07 <= 25.06) & (IS10 >= 9.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 84: (IE05 <= 3829.0) & (IE06 >= 654010.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 85: (IS03 <= 12.9) & (IS04 <= 16.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 86: (IE01 <= 37712.0) & (IE06 >= 547713.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 87: (IE03 >= 59.0) & (IE05 <= 65.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 88: (IE07 <= 163.16) & (IS04 <= 9.92) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 89: (IS03 <= 13.5) & (IS04 <= 15.6) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 90: (IS03 <= 16.63) & (IS04 <= 10.29) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 91: (IE04 <= 7.0) & (IS05 <= 0.01) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 92: (IE10 <= 115.0) & (IS02 <= 37.8) & (IS06 <= 0.01) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 93: (IE01 <= 767611.0) & (IS04 <= 14.4) & (IS06 <= 0.01) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 94: (IS04 <= 14.04) & (IS09 <= 137.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 95: (IS02 <= 31.8) & (IS09 <= 137.0) & (IS10 >= 8.12) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 96: (IS05 <= 2.8) & (IS09 <= 146.0) & (IS10 >= 9.865) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 97: (IS03 <= 18.4) & (IS09 <= 146.0) & (IA01 >= 382368.0) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 98: (IS03 <= 16.36) & (IS04 <= 13.32) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 99: (IE01 <= 1365659.0) & (IS03 <= 15.22) & (IS04 <= 13.61) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 100: (IE08 <= 117.0) & (IS03 <= 16.5) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 101: (IE02 >= 1511.2) & (IE07 <= 127.68) & (IS05 <= 1.85) => (CLASSE <= D1) |CERTAIN, AT\_MOST, D1|  
 102: (IE09 <= 2.43) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 103: (IE07 <= 5.24) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 104: (IE01 <= 1.6367069E7) & (IS04 <= 2.93) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 105: (IE06 >= 747127.0) & (IS01 <= 3761.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 106: (IS05 <= 2.6) & (IA03 >= 129340.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 107: (IS02 <= 25.45) & (IG01 <= 0.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 108: (IE06 >= 654010.0) & (IS02 <= 25.45) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 109: (IE03 >= 283.4) & (IS02 <= 38.6) & (IS06 <= 0.01) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 110: (IE10 <= 249.0) & (IS05 <= 3.3) & (IS06 <= 0.01) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 111: (IE02 >= 28747.0) & (IE04 <= 108.0) & (IG09 <= 0.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 112: (IE02 >= 28747.0) & (IE04 <= 108.0) & (IS06 <= 862.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 113: (IS03 <= 13.5) & (IS09 <= 471.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 114: (IE05 <= 1294.0) & (IS10 >= 20.0) & (IA03 >= 9671.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 115: (IE08 <= 315.0) & (IS10 >= 18.11) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 116: (IE02 >= 1513.87) & (IS02 <= 31.8) & (IS09 <= 137.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 117: (IS04 <= 14.37) & (IS09 <= 191.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 118: (IS04 <= 14.22) & (IS09 <= 202.0) & (IA03 >= 3770.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 119: (IE05 <= 4106.0) & (IE06 >= 438113.0) & (IS07 <= 43704.17) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 120: (IE08 <= 259.0) & (IS04 <= 13.19) & (IA03 >= 7389.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|

121: (IE05 <= 4010.0) & (IS05 <= 1.4) & (IS10 >= 9.0) & (IA03 >= 18505.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 122: (IS03 <= 16.5) & (IS04 <= 15.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 123: (IE10 <= 223.0) & (IS03 <= 16.255) & (IS06 <= 213.43) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 124: (IS03 <= 17.76) & (IS09 <= 361.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 125: (IS03 <= 18.37) & (IS09 <= 359.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 126: (IS06 <= 72.0) & (IA01 >= 365369.0) & (IG08 <= 0.0) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 127: (IE02 >= 10727.86) & (IE04 <= 14.0) & (IS05 <= 2.84) => (CLASSE <= D2) |CERTAIN, AT\_MOST, D2|  
 128: (IE01 <= 34373.0) & (IS01 <= 123.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 129: (IS05 <= 3.4) & (IS06 <= 10.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 130: (IS10 >= 38.15) & (IA03 >= 8201.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 131: (IE06 >= 747127.0) & (IE07 <= 2463.7) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 132: (IE04 <= 7.0) & (IG06 <= 0.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 133: (IE08 <= 561.0) & (IS09 <= 3.44) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 134: (IA03 >= 129340.0) & (IG04 <= 0.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 135: (IS02 <= 24.63) & (IG09 <= 0.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 136: (IS02 <= 25.45) & (IA03 >= 59045.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 137: (IS03 <= 21.66) & (IS09 <= 191.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 138: (IE01 <= 6105000.0) & (IA01 >= 5.0636541E7) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 139: (IS09 <= 202.0) & (IS10 >= 7.99) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 140: (IE09 <= 186.0) & (IS04 <= 9.06) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 141: (IE08 <= 316.0) & (IS04 <= 8.79) & (IS06 <= 123.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 142: (IE03 >= 3705.0) & (IE08 <= 283.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 143: (IE03 >= 3859.0) & (IS08 <= 15596.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 144: (IS03 <= 18.37) & (IS05 <= 1.85) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 145: (IE07 <= 378.0) & (IS03 <= 25.9) & (IS05 <= 1.7) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 146: (IS05 <= 1.4) & (IA02 >= 1246.795) & (IG08 <= 0.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 147: (IE02 >= 11365.0) & (IE03 >= 3590.5) & (IE08 <= 273.5) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 148: (IS03 <= 17.76) & (IS05 <= 1.88) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 149: (IE06 >= 206084.0) & (IS10 >= 12.7) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 150: (IS03 <= 17.9) & (IS04 <= 12.02) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 151: (IE07 <= 38.25) & (IA02 >= 336.471153611159) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 152: (IE03 >= 3415.0) & (IS08 <= 8478.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 153: (IS03 <= 16.5) & (IS09 <= 535.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 154: (IE05 <= 1231.0) & (IS06 <= 72.0) & (IA01 >= 365369.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 155: (IE09 <= 187.0) & (IS04 <= 13.48) & (IA03 >= 6674.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 156: (IE08 <= 227.0) & (IS04 <= 13.48) & (IA03 >= 9751.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 157: (IE08 <= 259.0) & (IS03 <= 20.0) & (IS04 <= 13.19) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 158: (IE04 <= 14.0) & (IA01 >= 24489.07) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 159: (IE01 <= 1080899.0) & (IA02 >= 307888.96) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 160: (IS06 <= 111.0) & (IA03 >= 20525.0) & (IG08 <= 0.0) => (CLASSE <= D3) |CERTAIN, AT\_MOST, D3|  
 161: (IS03 <= 19.3) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
 162: (IA02 >= 20847.0) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
 163: (IS04 <= 14.37) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
 164: (IA03 >= 83328.0) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
 165: (IE08 <= 5.56) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
 166: (IE04 <= 3.0) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
 167: (IS03 <= 19.8) & (IG08 <= 0.0) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|



168: (IS03 <= 24.0) & (IA02 >= 336.471153611159) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
169: (IE01 <= 254198.0) & (IS03 <= 24.0) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
170: (IE05 <= 309.2) & (IS03 <= 20.9) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
171: (IE03 >= 3415.0) & (IA01 >= 19745.89) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|  
172: (IS03 <= 26.8) & (IA02 >= 1583.61) => (CLASSE <= D4) |CERTAIN, AT\_MOST, D4|

## ANEXO VII

### Regras de decisão indSU inferidas pelo algoritmo VC-DomLEM, jMAF

#### [FILEINFO]

Program: jrs  
Algorithm: VC-DOMLEM  
DataFileDirectory: D:\2023-07-17\indSU  
DataFileName: dados\_indSU.isf  
Objects: 342  
Decision: CLASSE  
Duration: 2262 [ms]  
Comment: Rules generated with VC-DOMLEM algorithm

#### [PARAMETERS]

RulesFormat: INFORMATION\_TABLE\_PARETO  
ConsistencyLevel: 0.0  
Classes: I1, I2, I3, I4, I5

#### [ATTRIBUTES]

+ codigo\_empresa: (nominal), none, description  
+ ano: (nominal), none, description  
+ IE01: (continuous), gain  
+ IE02: (continuous), gain  
+ IE03: (continuous), gain  
+ IE04: (continuous), gain  
+ IE05: (continuous), gain  
+ IS01: (continuous), cost  
+ IS02: (continuous), cost  
+ IS03: (continuous), gain  
+ IS04: (continuous), cost  
+ IS05: (continuous), cost  
+ IS06: (continuous), cost  
+ IS07: (continuous), cost  
+ IS08: (continuous), cost  
+ IS09: (continuous), gain  
+ IS10: (continuous), cost  
+ IA01: (continuous), cost  
+ IA02: (continuous), gain  
+ IG01: (continuous), gain  
+ IG02: (continuous), gain  
+ IG03: (continuous), gain  
+ IG04: (continuous), cost  
+ IG05: (continuous), gain  
+ IG06: (continuous), gain  
+ IG07: (continuous), gain  
+ IG08: (continuous), gain  
+ IG09: (continuous), gain  
+ IG10: (continuous), gain  
+ IG11: (continuous), gain  
+ IG12: (continuous), gain  
+ CLASSE: [I1, I2, I3, I4, I5], gain, decision

decision: CLASSE

#### [PREFERENCES]

codigo\_empresa: none  
ano: none  
IE01: gain

IE02: gain  
 IE03: gain  
 IE04: gain  
 IE05: gain  
 IS01: cost  
 IS02: cost  
 IS03: gain  
 IS04: cost  
 IS05: cost  
 IS06: cost  
 IS07: cost  
 IS08: cost  
 IS09: gain  
 IS10: cost  
 IA01: cost  
 IA02: gain  
 IG01: gain  
 IG02: gain  
 IG03: gain  
 IG04: cost  
 IG05: gain  
 IG06: gain  
 IG07: gain  
 IG08: gain  
 IG09: gain  
 IG10: gain  
 IG11: gain  
 IG12: gain  
 CLASSE: gain

[RULES]

#Certain at least rules

1: (IA02 >= 365685.0) => (CLASSE >= I5) |CERTAIN, AT\_LEAST, I5|  
 2: (IS03 >= 404.0) & (IS08 <= 8.21) => (CLASSE >= I5) |CERTAIN, AT\_LEAST, I5|  
 3: (IE01 >= 216.0) & (IA01 <= 3.64) => (CLASSE >= I5) |CERTAIN, AT\_LEAST, I5|  
 4: (IE04 >= 5600.0) & (IS03 >= 332.0) => (CLASSE >= I5) |CERTAIN, AT\_LEAST, I5|  
 5: (IE03 >= 884294.0) & (IS02 <= 9357062.0) => (CLASSE >= I5) |CERTAIN, AT\_LEAST, I5|  
 6: (IE03 >= 897084.0) & (IS06 <= 288.0) => (CLASSE >= I5) |CERTAIN, AT\_LEAST, I5|  
 7: (IE05 >= 2037.0) & (IS09 >= 1776426.0) => (CLASSE >= I5) |CERTAIN, AT\_LEAST, I5|  
 8: (IA01 <= 4.1) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 9: (IA02 >= 49518.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 10: (IE03 >= 156785.0) & (IS09 >= 1361782.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 11: (IE01 >= 224.0) & (IS03 >= 335.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 12: (IE04 >= 5480.0) & (IE05 >= 2143.08) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 13: (IE03 >= 182079.0) & (IS07 <= 199.32) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 14: (IE01 >= 228.0) & (IS03 >= 250.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 15: (IE03 >= 68572.0) & (IS07 <= 4.34) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 16: (IA01 <= 5.3) & (IA02 >= 600.0) & (IG01 >= 1.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 17: (IS07 <= 208.07) & (IA02 >= 17845.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 18: (IE04 >= 6200.0) & (IE05 >= 1935.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 19: (IA01 <= 6.2) & (IG02 >= 1.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 20: (IS01 <= 12.1) & (IG02 >= 1.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 21: (IE03 >= 45295.0) & (IA02 >= 9114.0) & (IG02 >= 1.0) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 22: (IS07 <= 5.0) & (IA01 <= 15.24) => (CLASSE >= I4) |CERTAIN, AT\_LEAST, I4|  
 23: (IE03 >= 177471.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 24: (IE05 >= 1850.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 25: (IS01 <= 6.29) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 26: (IS08 <= 0.01) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 27: (IE01 >= 28.0) & (IA02 >= 8300.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|

28: (IS01 <= 247.3) & (IS03 >= 211.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 29: (IE01 >= 139.0) & (IS01 <= 9.28) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 30: (IS01 <= 13.32) & (IA01 <= 12.58) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 31: (IS01 <= 10.75) & (IS09 >= 1015577.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 32: (IS03 >= 167.0) & (IA02 >= 5120.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 33: (IA01 <= 8.07) & (IG10 >= 1.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 34: (IE03 >= 52946.0) & (IA01 <= 8.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 35: (IE05 >= 1615.39) & (IS07 <= 157.1) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 36: (IE03 >= 169863.5) & (IG10 >= 1.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 37: (IE01 >= 273.0) & (IA02 >= 2331.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 38: (IE05 >= 1799.0) & (IG10 >= 1.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 39: (IE01 >= 62.0) & (IA02 >= 7656.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 40: (IA01 <= 17.0) & (IG02 >= 1.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 41: (IE03 >= 65919.0) & (IA01 <= 15.24) & (IA02 >= 7554.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 42: (IE04 >= 6278.52) & (IE05 >= 1190.44) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 43: (IE05 >= 1581.17) & (IS02 <= 1360866.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 44: (IE05 >= 854.0) & (IS02 <= 162004.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 45: (IE02 >= 79.0) & (IA01 <= 21.9) & (IA02 >= 4064.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 46: (IS03 >= 152.0) & (IA01 <= 15.55) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 47: (IE01 >= 82.0) & (IS08 <= 3.68) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 48: (IS04 <= 0.44) & (IG01 >= 1.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 49: (IE02 >= 83.0) & (IS10 <= 4153.0) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 50: (IS05 <= 0.01) & (IS06 <= 174.0) & (IA01 <= 13.35) => (CLASSE >= I3) |CERTAIN, AT\_LEAST, I3|  
 51: (IE03 >= 79857.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 52: (IG01 >= 1.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 53: (IE05 >= 1292.84) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 54: (IS03 >= 190.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 55: (IE04 >= 6278.52) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 56: (IS08 <= 2.75) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 57: (IE01 >= 72.0) & (IS04 <= 0.01) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 58: (IS04 <= 1.76) & (IG10 >= 1.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 59: (IE04 >= 1815.0) & (IS08 <= 3.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 60: (IE05 >= 1163.85) & (IG06 >= 1.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 61: (IE03 >= 68751.5) & (IG09 >= 1.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 62: (IA01 <= 15.0) & (IG09 >= 1.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 63: (IE02 >= 52.0) & (IA02 >= 2571.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 64: (IE02 >= 63.0) & (IG11 >= 1.0) => (CLASSE >= I2) |CERTAIN, AT\_LEAST, I2|  
 #Certain at most rules  
 65: (IS07 >= 959.07) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 66: (IS04 >= 8.57) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 67: (IA01 >= 205.54) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 68: (IE05 <= 350.0) & (IS09 <= 36979.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 69: (IE03 <= 2426.0) & (IS06 >= 192.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 70: (IE04 <= 1560.0) & (IE05 <= 350.5) & (IS08 >= 3.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 71: (IE05 <= 390.451612903226) & (IS07 >= 128.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 72: (IE05 <= 1086.0) & (IA01 >= 31.15) & (IG09 <= 0.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 73: (IE02 <= 3.5) & (IS02 >= 288542.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 74: (IE02 <= 3.5) & (IS03 <= 66.0) & (IS07 >= 150.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 75: (IS08 >= 6.66) & (IG06 <= 0.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 76: (IE03 <= 4444.0) & (IS10 >= 20455.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 77: (IE03 <= 70232.0) & (IS02 >= 4281166.0) => (CLASSE <= I1) |CERTAIN, AT\_MOST, I1|  
 78: (IS09 <= 28455.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|  
 79: (IE03 <= 3626.0) & (IS02 >= 162004.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|  
 80: (IE01 <= 273.0) & (IS04 >= 5.84) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|  
 81: (IE03 <= 73226.0) & (IA01 >= 31.15) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|  
 82: (IS01 >= 175.67) & (IS10 >= 88315.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|

- 83: (IS08 >= 17.22) & (IA01 >= 17.23) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 84: (IE04 <= 2940.0) & (IS03 <= 17.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 85: (IE04 <= 6.42) & (IS02 >= 1384180.0) & (IS09 <= 510119.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 86: (IE02 <= 3.5) & (IE05 <= 586.29) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 87: (IS04 >= 2.97) & (IS08 >= 15.8) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 88: (IE05 <= 473.0) & (IS09 <= 305478.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 89: (IS03 <= 139.0) & (IG06 <= 0.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 90: (IE02 <= 77.0) & (IS07 >= 540.19) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 91: (IE05 <= 683.0) & (IS08 >= 15.17) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 92: (IE01 <= 72.0) & (IG10 <= 0.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 93: (IE05 <= 562.0) & (IA01 >= 13.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 94: (IE05 <= 562.0) & (IS09 <= 161573.0) & (IG02 <= 0.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 95: (IS08 >= 13.17) & (IA01 >= 25.1) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 96: (IE01 <= 259.0) & (IE05 <= 846.0) & (IS01 >= 175.67) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 97: (IE03 <= 28148.0) & (IE05 <= 1190.44) & (IS06 >= 221.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 98: (IE02 <= 83.0) & (IS03 <= 33.5) & (IA01 >= 15.925) & (IG02 <= 0.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 99: (IE04 <= 1953.0) & (IS09 <= 155545.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 100: (IE03 <= 31357.0) & (IA01 >= 6.485) & (IG10 <= 0.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 101: (IE03 <= 46895.0) & (IS06 >= 257.0) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 102: (IE05 <= 879.0) & (IS03 <= 54.0) & (IA01 >= 22.12) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 103: (IS03 <= 70.0) & (IS09 <= 225041.0) & (IA01 >= 21.94) => (CLASSE <= I2) |CERTAIN, AT\_MOST, I2|
- 104: (IE04 <= 6080.0) & (IS01 >= 215.04) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 105: (IS03 <= 35.0) & (IA02 <= 4064.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 106: (IS03 <= 30.0) & (IG01 <= 0.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 107: (IS09 <= 777089.0) & (IA01 >= 28.34) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 108: (IE02 <= 104.0) & (IS08 >= 15.17) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 109: (IS05 >= 1.16) & (IA01 >= 11.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 110: (IE04 <= 7.08) & (IA01 >= 15.77) & (IG11 <= 0.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 111: (IE05 <= 562.0) & (IA01 >= 4.5) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 112: (IS03 <= 40.0) & (IA01 >= 16.1) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 113: (IE03 <= 4722.0) & (IS06 >= 189.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 114: (IS09 <= 65053.0) & (IA02 <= 228.0) & (IG02 <= 0.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 115: (IE03 <= 31357.0) & (IA01 >= 6.485) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 116: (IS09 <= 210123.0) & (IA01 >= 12.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 117: (IE05 <= 778.0) & (IS09 <= 254534.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 118: (IS02 >= 872092.0) & (IS09 <= 225651.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 119: (IS09 <= 217797.0) & (IS10 >= 20415.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 120: (IS09 <= 216960.0) & (IA01 >= 17.76) & (IG02 <= 0.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 121: (IS05 >= 0.455) & (IA01 >= 19.05) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 122: (IS07 >= 540.19) & (IA01 >= 17.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 123: (IS03 <= 50.0) & (IS08 >= 6.02) & (IS09 <= 259237.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 124: (IS01 >= 142.34) & (IS06 >= 204.0) & (IG12 <= 0.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 125: (IE01 <= 72.0) & (IS08 >= 13.44) & (IS10 >= 6073.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 126: (IE05 <= 637.0) & (IS10 >= 37092.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 127: (IS06 >= 238.0) & (IS09 <= 529149.0) => (CLASSE <= I3) |CERTAIN, AT\_MOST, I3|
- 128: (IE01 <= 186.0) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|
- 129: (IE05 <= 633.0) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|

130: (IE04 <= 8.02812982998454) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
131: (IS01 >= 175.67) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
132: (IS04 >= 4.93) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
133: (IS05 >= 1.55) & (IA02 <= 3120.0) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
134: (IS06 >= 310.0) & (IA02 <= 0.01) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
135: (IS03 <= 54.0) & (IA02 <= 358.0) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
136: (IE02 <= 117.0) & (IA01 >= 20.48) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
137: (IE05 <= 706.0) & (IS01 >= 82.0) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|  
138: (IE05 <= 735.0) & (IA02 <= 0.01) => (CLASSE <= I4) |CERTAIN, AT\_MOST, I4|

## ANEXO VIII

**Procedimento em R, *package RoughSets*, para predição de classe dirSU na simulação de ascensão, E25 para E27**

**Obs.: a tabela de decisão na etapa 1 foi reduzida para fins de exemplificação.**

```
#####
#
# Predição de classe dirSU - v. 09/2023
#
#####
library(RoughSets)
library(tidyverse)
library(modelr)
library(dplyr)

#####
# 1- importa e prepara tabela de decisão; dados de treinamento
#####
dt.ex1 <- data.frame (
c( 953647, 136949.5, 3228811, 8882293, 38835, 22104, 4519630, 1511594, 1804425, 263115, 209981,
1015526, 1365659, 445557, 105555, 1404298, 645131, 761924, 775837.5, 573855, 198742, 1197777,
1791090, 2810489, 2629206, 1365097, 17580637, 6352000, 3888000, 1526000, 2833000, 1143000,
35328.5, 4288505, 12104, 118969, 155000, 7311, 1045326, 136949.5, 2899170, 8695421, 38175, 22104,

"D3", "D5", "D2", "D2", "D2", "D4", "D3", "D3", "D3", "D4", "D1", "D3", "D3", "D4", "D1", "D3", "D1",
"D3", "D2", "D4", "D5", "D2", "D4", "D3", "D2", "D5", "D4", "D1", "D1", "D2", "D3", "D2", "D1", "D1",
"D1", "D1", "D5", "D3", "D3", "D3", "D2", "D2", "D2", "D4", "D3", "D4", "D3", "D4", "D1", "D3", "D3",
"D4", "D1", "D3" ))

colnames(dt.ex1) <- c("IE01", "IE02", "IE03", "IE04", "IE05", "IE06", "IE07", "IE08", "IE09", "IE10",
"IS01", "IS02", "IS03", "IS04", "IS05", "IS06", "IS07", "IS08", "IS09", "IS10", "IA01", "IA02", "IA03",
"IG01", "IG02", "IG03", "IG04", "IG05", "IG06", "IG07", "IG08", "IG09", "IG10", "CLASSE" )

attributes <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27,
28, 29, 30, 31, 32, 33, 34)

conditional.attr <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,
27, 28, 29, 30, 31, 32, 33)
decision.attr <- c(34)

decision.table1 <- SF.asDecisionTable(dataset = dt.ex1, decision.attr = 34, indx.nominal = 34)

#####
# 2- importa e prepara tabela p/ predição da classe; dados de teste E25
#####

dt.ex2 <- data.frame (
c( 17580637 ),
c( 4773 ),
c( 842),
c( 844 ),
c( 31349 ),
c( 103814 ),
c( 378 ),
c( 339 ),
c( 204 ),
c( 255 ),
```

```

c( 8870),
c( 45.3),
c(25.9),
c( 33.06 ),
c( 4 ),
c( 213.43 ),
c( 15068 ),
c( 13942 ),
c( 1473 ),
c( 15.74 ),
c( 328797.19 ),
c( 319904.17 ),
c( 34449 ),
c( 1 ),
c( 0 ),
c( 0 ),
c( 1 ),
c( 0 ),
c( 1 ),
c( 0 ),
c( 1 ),
c( 0 ),
c( 1 ),
c( 0 ),
c( 0 ))

```

```

colnames(dt.ex2) <- c("IE01" ,"IE02" ,"IE03" ,"IE04" ,"IE05" ,"IE06" ,"IE07" ,"IE08" ,"IE09" ,"IE10"
,"IS01" ,"IS02" ,"IS03" ,"IS04" ,"IS05" ,"IS06" ,"IS07" ,"IS08" ,"IS09" ,"IS10" ,"IA01" ,"IA02" ,"IA03"
,"IG01" ,"IG02" ,"IG03" ,"IG04" ,"IG05" ,"IG06" ,"IG07" ,"IG08" ,"IG09" ,"IG10")

```

```

attributes <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27,
28, 29, 30, 31, 32, 33)
conditional.attr <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,
27, 28, 29, 30, 31, 32, 33)

```

```

decision.table2 <- SF.asDecisionTable(dataset = dt.ex2)

```

```

#####
# 3- discretiza tabelas; pesquisa relações de indiscernibilidade
#####

```

```

cut.values <- D.discretize.equal.intervals.RST(decision.table1, nOfIntervals = 7)

```

```

## generate a new decision table

```

```

dectab.discretized <- SF.applyDecTable(decision.table1, cut.values)
decision.table1A <- SF.asDecisionTable(dataset = dectab.discretized, decision.attr = 34)

```

```

dectab.discretized <- SF.applyDecTable(decision.table2, cut.values)
decision.table2A <- SF.asDecisionTable(dataset = dectab.discretized)

```

```

IND.relacoes = BC.IND.relation.RST(dectab.discretized, feature.set = NULL)

```

```

#####
# 4- pesquisa as aproximações inferior e superior
#####

```

```

#aproxim <- BC.LU.approximation.RST(dectab.discretized, IND.relacoes)

```

```

#####
# 5- pesquisa região positiva
#####

```



```

#pos.region = BC.positive.reg.RST(dectab.discretized, aproxim)

#####
# 6- pesquisa matriz de discernibilidade/reduo(s)
#####

#disc.matrix <- BC.discernibility.mat.RST(decision.table1A, return.matrix = TRUE)

#res.1 <- FS.quickreduct.RST(decision.table1A)

#####
# 8- realiza prediçao a partir das regras geradas
#####

rules2 <- RI.CN2Rules.RST(decision.table1A, K = 5)
pred.vals2 <- predict(rules2, decision.table2A)

#####
# 9- imprime resultados
#####

print("Prediçoes:")

print("RI.CN2Rules.RST")
pred.vals2

## fim

> print("RI.CN2Rules.RST")
[1] "RI.CN2Rules.RST"
> pred.vals2
predictions
  D4

```

## ANEXO IX

### Procedimento em VBA/Microsoft Excel, para análise da sensibilidade de indicadores

```

Sub h_análise()
'
' analisa variações nos valores de indicadores visando verificar se houve alteração na classe
'
' entrada: _Indicadores4_ESAG 6.xlsm ; saída: _Indicadores4_ESAG 7.xlsm
'
Dim x1 As New Excel.Application
Dim xlw1, xlw2 As Excel.Workbook
Dim ws As Worksheet

Dim ano, li_empresa, li_relatao, li_indicador, c_cl_atual, c_un, i, n1, n2, nn, n_ind_dir, n_ind_ind,
n_codemp, col_ind, variacao, variacao_ant As Integer
Dim ano_s, cod_empresa, aba1, aba2, pasta1, pasta2, cod_indicador, dimensao, xorigem, origem,
nom_ind, nom_ind_esag, classe_inicial_dir, classe_inicial_ind, classe_final_dir, classe_final_ind,
classe_atual As String
Dim g_c, classe_inicial, classe_final As String
Dim fim, var_maior As Boolean

Dim valor, valor_ref, xvalor As Double

' ===== PARÂMETROS =====
pasta2 = "C:\Users\ayrto\OneDrive\Documents\UFF_Doutorado\Pesquisa RSA\planilhas geradas_23-
11-22 (indicadores E-S-A-G)\Resultados\"
' =====
'
ano = "2021"
empresa = "E07"
classe_final_dir = "D4"
classe_final_ind = "I4"
'
' =====

'Abre o arquivo do Excel
Set xlw1 = x1.Workbooks.Open(pasta2 + "_Indicadores4_ESAG 6.xlsm")

xlw1.Worksheets("Empresas").Cells(1, 10).Value = Now

For i = 1 To 2 ' A1

li_relatao = 6
li_empresa = 3

If i = 1 Then
aba1 = ano + "_dirSU"
aba2 = "análise sensibilidade_dir"
origem = "dirSU"
c_cl_atual = 7
c_un = 4
n_ind = 33

cod_empresa = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 1).Value)
While (cod_empresa <> empresa)

```

```

    li_empresa = li_empresa + 1
    cod_empresa = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 1).Value)
Wend

    classe_final = classe_final_dir
    classe_inicial = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 35).Value)
Else
    aba1 = ano + "_indSU"
    aba2 = "análise sensibilidade_ind"
    origem = "indSU"
    c_cl_atual = 6
    c_un = 3
    n_ind = 29

    cod_empresa = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 1).Value)
    While (cod_empresa <> empresa)
        li_empresa = li_empresa + 1
        cod_empresa = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 1).Value)
    Wend

    classe_final = classe_final_ind
    classe_inicial = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 31).Value)
End If

xlw1.Worksheets(aba2).Cells(3, 2).Value = "EMPRESA: " + empresa
xlw1.Worksheets(aba2).Cells(3, 3).Value = "ANO: " + ano

xlw1.Worksheets(aba2).Cells(3, c_cl_atual).Value = "CLASSE ATUAL: " + classe_inicial
xlw1.Worksheets(aba2).Cells(3, (c_cl_atual + 2)).Value = "CLASSE A PROSPECTAR: " +
classe_final

For n = 1 To n_ind ' A2

    col_ind = n + 1
    cod_indicador = Trim(xlw1.Worksheets(aba1).Cells(1, col_ind).Value)
    g_c = Trim(xlw1.Worksheets(aba1).Cells(2, col_ind).Value) ' ganho ou custo

    If g_c = "G" Then
        valor_ref = xlw1.Worksheets(aba1).Cells(41, col_ind).Value ' valor máximo
    Else
        valor_ref = xlw1.Worksheets(aba1).Cells(42, col_ind).Value ' valor mínimo
    End If

    valor = xlw1.Worksheets(aba1).Cells(li_empresa, col_ind).Value

    ' pesquisa descrição, unidade, GRI, G-C
    li_indicador = 1
    xcod = ""
    xorigem = ""

    While Not ((xcod = cod_indicador) And (xorigem = origem))
        li_indicador = li_indicador + 1
        xcod = Trim(xlw1.Worksheets("Indicadores_natureza").Cells(li_indicador, 2).Value)
        xorigem = Trim(xlw1.Worksheets("Indicadores_natureza").Cells(li_indicador, 3).Value)
    Wend
'

    If (xcod = cod_indicador) And (xorigem = origem) Then

```

```

xlw1.Worksheets(aba2).Cells(li_relata, 2).Value =
Trim(xlw1.Worksheets("Indicadores_natureza").Cells(li_indicador, 8).Value) ' descrição
If i = 1 Then
    xlw1.Worksheets(aba2).Cells(li_relata, 3).Value =
Trim(xlw1.Worksheets("Indicadores_natureza").Cells(li_indicador, 7).Value) ' GRI
End If
xlw1.Worksheets(aba2).Cells(li_relata, c_un).Value =
Trim(xlw1.Worksheets("Indicadores_natureza").Cells(li_indicador, 5).Value) ' unidade
xlw1.Worksheets(aba2).Cells(li_relata, (c_un + 1)).Value =
Trim(xlw1.Worksheets("Indicadores_natureza").Cells(li_indicador, 6).Value) ' G-C
End If
'

variacao = 0 '%'
var_maior = False
classe_atual = ""
fim = False

While (classe_final <> classe_atual) And (Not fim)

If Left(cod_indicador, 2) <> "IG" Then
    variacao = variacao + 1

If classe_final > classe_inicial Then
    If g_c = "G" Then
        xvalor = valor * ((100 + variacao) / 100)
    Else
        xvalor = valor * ((100 - variacao) / 100)
    End If
Else
If classe_final < classe_inicial Then
    If g_c = "G" Then
        xvalor = valor * ((100 - variacao) / 100)
    Else
        xvalor = valor * ((100 + variacao) / 100)
    End If
End If
End If
Else ' indicadores de governança (0 ou 1)
If classe_final > classe_inicial Then
    If g_c = "G" Then
        xvalor = 1
    Else
        xvalor = 0
    End If
Else
If classe_final < classe_inicial Then
    If g_c = "G" Then
        xvalor = 0
    Else
        xvalor = -1
    End If
End If
End If
fim = True
End If

xlw1.Worksheets(aba1).Cells(li_empresa, col_ind).Value = xvalor ' atualiza planilha para obter
classe_atual

```

```

If i = 1 Then
  classe_atual = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 35).Value) ' aba aaaa_dirSU
Else
  classe_atual = Trim(xlw1.Worksheets(aba1).Cells(li_empresa, 31).Value) ' aba aaaa_indSU
End If

If variacao >= 100 Then
  var_maior = True
  fim = True
End If

Wend

If (classe_final = classe_atual) Then
  xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 2)).Value = classe_atual
  xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 3)).Value = valor
  xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 4)).Value = xvalor
  If Left(cod_indicador, 2) <> "IG" Then
    xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 5)).Value = variacao
    xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 6)).Value = valor_ref
    If g_c = "G" Then
      xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 7)).Value = (xvalor / valor_ref) * 100
    Else
      If xvalor = 0 Then
        xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 7)).Value = 0
      Else
        xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 7)).Value = (xvalor / valor_ref) * 100
      End If
    End If
  Else
    xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 5)).Value = ""
  End If
Else
  If var_maior Then
    xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 5)).Value = "" 'até 100"
    xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 6)).Value = ""
    xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 7)).Value = ""
  Else
    ' xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 2)).Value = classe_atual
    ' xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 3)).Value = valor
    ' xlw1.Worksheets(aba2).Cells(li_relat, (c_un + 4)).Value = xvalor
  End If
End If

'retorna valor inicial do indicador
xlw1.Worksheets(aba1).Cells(li_empresa, col_ind).Value = valor

li_relat = li_relat + 1

Next ' A2

Next ' A1

xlw1.Worksheets("Empresas").Cells(1, 11).Value = Now

MsgBox ("===== CONCLUÍDO =====")

End Sub

```