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[1], Luís Alberto Duncan Rangel[1]

[1]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.

1 Introduction

The current study is aimed at the proposition of a model for sustainability classes prediction for electricity distribution companies in Brazil, based on sustainability indicators and indexes, aiming at a more effective risk management for a certain company in relation to their competitors. The model to be proposed is based in a multicriteria approach (Gomes & Rangel, 2009a, b; Slowinski et al., 2012) with the use of Rough Sets Theory (RST), a mathematical theory for treating imprecise data, and Machine Learning, an Artificial Intelligence subfield, in order to infer decision rules, *if … then …*, in the prediction of sustainability classes based on progress and sustainability indicators, using RoughSets package in R and the jMAF software.

According to Wu & Wu (2012), Elkington (2020) and Pereira & Cândido (2020), the sustainable development and sustainability involve interdisciplinary themes and is present in discussions across many science areas, in public and private organizations, non-governmental ones and in society as a whole. Sustainability is related not only to ecologic aspects, but it also relates to economic, political, cultural, social, temporal, and spatial aspects, deeming it essential the creation of measurement instruments, such as sustainability indicators, tools made up of one or more variables, that can be related in many ways. Consequently, establishing goals and creating instruments are fundamental steps in making possible the corporate sustainability measuring. There are many corporate sustainability indicator systems being used in Brazil, such as the Brazilian Institute of Social and Economic Analysis' (IBASE) model; the Ethos model; the Corporate Sustainability Index (CSI); Dow Jones Sustainability Index (DJSI); and the international model Global Reporting Initiative (GRI). Paz & Kipper (2016) point towards other models to measure sustainability in public and private organizations from a variety of fields: Pressure-State-Response (PSR), environmental dimension; Driving Forces-State-Response (DSR), social, environmental, institutional and economic dimensions; Genuine Progress Indicator (GPI), social and economic dimensions; World Bank, social, environmental, economic and cultural dimensions; Human Development Index (HDI), social, economic, cultural and political dimensions; Barometer of Sustainability, social and environmental dimensions; Sustainability Panel, social, environmental, institutional and economic dimensions; etc. According to Elkington (2020), all this effort must be aligned with the United Nations' Sustainable Development Goals: a set of 17 ambitious goals and 169 related targets. The sustainable development goals must be the "north" for the sustainability.

Furthermore, the indicators can be unique when they represent a specific dimension, or aggregated, when dimensions of the process or part of it are represented by a set of indicators that are frequently aggregated in other indicators. According to Wu & Wu (2012) and Franceschini et al. (2019), in order to reduce the number of indicators or to reflect integrations in a system, indicators are mathematically combined to produce indexes, aggregate of two or more indicators.

Regarding Figure 1, the emphasis in identification and use of indicators in a predictive way is relatively new. Predictive indicators are suitable when the main interest is to prevent the occurrence of problems, instead of fixing them, e.g.: financial flow through time (Franceschini et al., 2019).

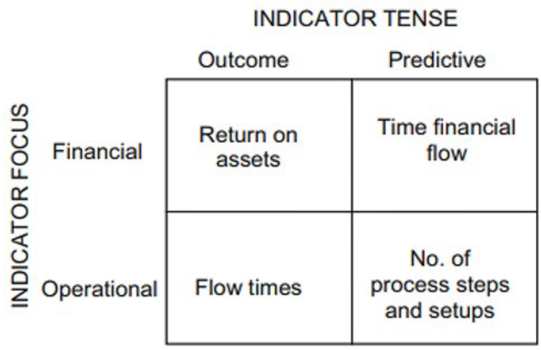

Figure 1. Predictive indicators for financial and operational use. Source: Franceschini et al. (2019).

Furthermore, ISO 26000, Social Responsibility Guidance Standard (ABNT, 2010), established that "An organization can exert influence over others to strengthen the positive impacts on sustainable development or to minimize the negative impact, or both cases". Among the proposed methods to exert influence, there is promoting best practices. In September of 2015, the United Nations General Assembly, through the 2030 Agenda, established a collection of 17 global goals (Global Goals for a Sustainable Development) (ONU, 2015). In that Agenda, regarding Goal "12", Ensure sustainable consumption and production patterns, you can read: Encourage companies, especially large and transnational companies, to adopt sustainable practices and to integrate sustainability information into their reporting cycle.

In Brazil, the National Electric Energy Agency (ANEEL), through Normative Resolution n. 605, from 03/11/2014, approved the Electricity Sector Accounting Manual (ESAM), where it was established, among other goals, contributing to the optimization of social environmental performance through explicitly showing costs originated from compliance with the National Environment Policy, necessary to the environmental conformity and the sustainability of concessions attributed by the Federal Union, aiming at the elaboration of the Report on Socio-Environmental and Economic-Financial e (RSA) (Brasil, 2014; ANEEL, 2015). The ESAM relates more than two hundred quantitative indicators related to the general, corporate governance, economic-financial, social and sectorial, and environmental dimensions for granted companies, generation, transmission and distribution of electricity, demanded and/or suggested by ANEEL, beyond environmental performance indicators that are specific to generation and transmission and/or distribution companies (ANEEL, 2015). The ESAM still mentions qualitative indicators, e.g.: aspects of corporate governance.

The contribution**s** of this study are in the model proposition that: a) uses Rough Sets Theory/Dominance principle and Machine Learning to extract decision rules and infer sustainability classes from historical series of indicators and simulated values, aiming to obtain better risk management in the economic, social, environmental and corporate governance dimensions of a company against competitors; b) establishes sustainability classes for companies in order to identify possible links on aspects of sustainability and performance between companies belonging to the same class, as opposed to simple ranking; c) allows to relate condition and decision criteria in decision rules, for example, by coverage factor, and consequently, obtain patterns in data, without referring to a priori and posterior probabilities, as in Bayesian analysis (Pawlak, 2002).

This study is currently divided into the following additional sections: 2, Literature review; 3, Methodology, composed of 3.1, Rough Sets Theory and Dominance principle, and 3.2, Machine Learning; 4, Results and discussions; and 5, Conclusion and future studies.

2 Literature review

A systematic literature review was carried out in May/2022, in the databases Scopus, Web of Science, Compendex, IEEE Xplore, Emerald Insight, Scielo, ACM Digital Library, EBSCO and Wiley Online, for the period 2000-2022, using the ProKnow-C methodology, Knowledge Development Process – Constructivist (Afonso et al., 2011; Ensslin et al., 2014). Table 1 illustrates the sequence of procedures in eight phases adapted from the ProKnow-C methodology, which was used to build the systematic literature review. The following search string (1) was used for the abstract field: *("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")*. There was a return of 689 articles (bibliographic portfolio BP0, Phase 4, according to Table 1) which, after excluding repeated publications, reading the titles and abstracts, using a "cut line" (above, at least 85% of the total citations on Google Scholar; below, most recent articles, published in the last 3 years), full reading of the remainder and verification of adherence to the research, resulted in 53 studies. To this set of 53 studies, the following search string (2) was applied to the abstract field: *(("multi-criteria" OR "multicriteria" OR "multi-objective" OR "multiobjective") OR "machine learning").* From this set, 14 studies resulted. Complementarily, for the period January-May 2022, the junction "and" of the previous strings 1 and 2 was applied to search for new studies. It resulted in 7 studies. Thus, there are 14 + 7 studies. Table 2 summarizes the results of the systematic literature review.

Table 1. Systematic literature review, ProKnow-C methodology. Source: Authors, adapted from Ensslin et al. (2014).

PHASE 1	PHASE 2	PHASE 3	PHASE 4	PHASE 5	PHASE 6	PHASE 7	PHASE 8
Set keywords for the search	Define bases aligned with research	Perform title, abstract and keyword research of publications in each database	Count the number of publications per axis and base in order to build the initial bibliographic portfolio (BP0)	Raise number of BP1 citations on Google Scholar	Select most cited publications (up to 85% of total citations)	Select least cited publications (remaining 15% of total citations)	Gather publications BP2, BP3 and BP4
Define research axes; assemble boolean expressions ("and"; "or")	Set search filters (time period, subject, publication type, etc.)	Review whether searched keywords adhere to the keywords of the publications	Select publications adhering to the research (reading the title); exclude repeated posts (BP1)	Sort publications in descending order of citations	Select publications adhering to the research (abstract reading) (BP2)	Select publications in the last 2 years adhering to the research (abstract reading) (BP3); Select remaining publications that have authors in BP2 and adherent to the research (abstract reading) (BP4)	Select publications adhering to the research (full reading) to assemble the final bibliographic portfolio (BP5)

The Table 1 shows the sequence of procedures divided into eight phases adapted from the ProKnow-C methodology, aiming to simplify its use in systematic literature review works.

Table 2. Results of the systematic literature review. Source: Authors.

Author(s)	Objective	Method(s)	Result(s)
Ahmad et al. (2021)	A baseline that allows researchers and readers to compare their AI efforts, new state-of-the-art applications and global roles in policymaking of energy industry	AI techniques	The energy industry, utilities, power system operators, and independent power producers may need to focus more on AI technologies if they want meaningful results to remain competitive.
Ahmadi et al. (2022)	Predicting energy consumption patterns in urban buildings	Deep neural network with fuzzy wavelets	This study shows that the presented method provides high-performance prediction at a lower level of complexity.
Al-Barakati et al. (2022)	Evaluating the renewable energy sources	Extended interval-valued Pythagorean fuzzy WASPAS	The evaluation results showed that the wind energy with a maximum assessment score degree using the proposed method was found the best option for selecting renewable energy sources over diverse criteria.
Buțurache & Stancu (2022)	Building energy consumption prediction	Neural-based models	Neural-based models possess the capability of learning and generalizing from different datasets having different patterns.
Caiado et al. (2017)	Proposing a novel model for solving decision-making problems in the evaluation of electrical energy companies	TOPSIS method	Robust evaluation and ranking of the energy companies with respect to the observed aspects of sustainability.
Chamandoust et al. (2020)	Performance assessment of smart hybrid energy system (SHES)	Shuffled frog leaping algorithm (SFLA)	Optimal scheduling of SHES with acceptable levels of operation costs, emission pollution and customer satisfaction.
Colla et al. (2020)	Critically review of multi-disciplinary KPIs, allowing a holistic comparison across different types of energy projects	A structured evaluation framework based on the identified set of indicators	Integrated framework and a fairer assessment of competing energy projects by relevant stakeholders.
Daugavietis et al. (2022)	A Comparison of Multi-Criteria Decision Analysis Methods for sustainability assessment of District Heating systems	WSM, TOPSIS, PROMETHEE, ELECTRE and DEA methods	The results of sensitivity analysis along with literature investigation shows that all methods are suitable for sustainability analyses of District Heating systems while also having differences in the calculation process and in the interpretation of results.
Ervural et al. (2018a)	Energy investment planning	TOPSIS and AHP	The renewable energy investment plan of a power company in Turkey is evaluated with a newly developed integral approach.
Ervural et al. (2018b)	Energy planning	ANP, fuzzy TOPSIS and SWOT analysis	Integrated framework for the Turkey's energy sector to prioritize alternative energy strategies.
Kwakkel & Pruyt (2013)	An approach for model-based foresight under deep uncertainty	Exploratory Modeling and Analysis	Multiplicity assessment of deep uncertainties in the analysis of decision-making problems in the electricity sector.
Panchal et al. (2022)	A novel structured framework for analyzing sustainable operational performance-related issues of ash handling unit (AHU) under vague/uncertain information	Integrated fuzzy lambda–tau and fuzzy multicriteria decision-making methods	The current work discussed a framework for evaluating performance issues of the thermal power industry for sustainable and environmental friendly operation for emission control.
Qi et al. (2020)	Defining a set of criteria and dimensions for analyzing the corporate governance-based strategic	IVIF DEMATEL; IVIF VIKOR	Extending investigations on corporate governance and sustainable production in energy industry.
Rigo et al. (2020)	Systematic literature review of renewable energy problems associated with MCDM methods	MCDM methods	Improving their ability to choose the proper MCDM methods to solve energy problems.
Rolnick et al. (2022)	How machine learning can be a powerful tool in reducing greenhouse gas emissions and helping society adapt to a changing climate	Machine Learning techniques	Climate change solution domains with selected areas of ML that are relevant.

Table 2. Continued…

Author(s)	Objective	Method(s)	Result(s)
Sahabuddin & Khan (2021)	Energy sector's sustainability assessment	Multi-criteria decision analysis methods	The analysis revealed that COPRAS is the most robust MCDA method, followed by WPM.
Tajbakhsh & Shamsi (2019)	Analyzing a comprehensive objective function in the facility location context in the presence of variable costs, fixed costs, and sustainability consideration	Double bootstrap data envelopment analysis	The proposed approach substantially diminishes greenhouse gas emissions at the cost of slight increases in total expense.
Vargas-Solar et al. (2022)	Providing data analytical tools for metering household energy consumption and CO_2 footprint under different perspectives	GREENHOME environment toolkit	The article reports on experiments conducted for modelling and forecasting energy consumption and CO_2 footprint in the context of the Triple-A European project.
Wang et al. (2021)	Using sustainable performance prediction in data-scarce scenarios	Integrated energy efficiency system	Strengthening the real-time control of integrated energy projects and for effectively promoting the sustainable development of the integrated energy industry.
Wanke et al. (2020)	Improving the technical efficiency of the energy industry in China	Bayesian stochastic frontier analysis	Increases in the efficiency of the Chinese energy industry can be achieved by increasing the level of inventories and fixed assets.
Zhou et al. (2019)	Evaluation of sustainable energy investment projects	Balanced Scorecard, IT2 Fuzzy DEMATEL, IT2 Fuzzy QUALIFLEX	Which issues are effective in financial institutions' lending process of large-scale energy projects.

There are studies with the application of several multicriteria decision support methods, in individual or hybrid form, as well as proposals for specific models aimed at solving electrical energy problems. This reveals trends to use models with Machine Learning and Neural Networks, for example, to infer results on production, efficiency and consumption of electricity (Ahmad et al., 2021; Ahmadi et al., 2022; Buțurache & Stancu, 2022; Kwakkel & Pruyt, 2013; Rolnick et al., 2022; Vargas-Solar et al., 2022). In addition, there are proposals for models for analyzing problems using Fuzzy logic, a theory for the mathematical treatment of data imprecision (Al-Barakati et al., 2022; Ervural et al., 2018a, b; Panchal et al., 2022; Qi et al., 2020; Zhou et al., 2019).

Furthermore, there was an extension of the systematic literature review in the Web of Science and Scopus bases, in the period 2020-2022, using the following search string: "*TITLE-ABS-KEY (("multicriteria" OR "multi-criteria" OR "multiobjective" OR "multi-objective") AND "decision making" AND ("predicting" OR "forecasting"))*", "*decision*" and "*artificial intelligence*" subareas. There was a return of 36 (thirty-six) articles. However, of these studies, only 12 are related to energy, classification or machine learning. There are studies with the joint application of the Promethee multicriteria method and Machine Learning prediction in financial decision making (Mousavi & Lin, 2020); and a new hybrid fuzzy prediction method is introduced by combining the Fuzzy Analytic Hierarchy Process (FAHP) and machine learning model (Ozdemir et al., 2021). For applications to energy or classification problems: fuzzy interval time series energy and financial forecasting model (Liu et al., 2020); flood hazards susceptibility mapping using statistical, fuzzy logic and MCDM methods (Akay, 2021); assessment of a failure prediction model in the energy sector with multicriteria discrimination approach, Promethee based classification (Angilella & Pappalardo, 2021); hybrid neurofuzzy investigation of short-term variability of wind resource (Adedeji et al., 2021); TOPSIS-ELM framework for stock index price movement prediction (Samal & Dash, 2021); cost-sensitive business failure prediction when misclassification costs are uncertain (Bock et al., 2020); multi objective optimization of crude oil supply portfolio based on interval prediction data (Sun et al., 2022);

optimization of integrated fuzzy decision tree and regression models for selection of oil spill response (Mohammadiun et al., 2021); use of PairCode algorithm for ordinal classification based on pairwise comparison (Yang et al., 2020); and client profile prediction using convolutional neural networks (Nedjah et al., 2022). In addition, there are studies that show the application of the PROMETHEE-SAPEVO-M1 multi-criteria method to the analysis of OECD countries (Pereira et al., 2022) and multicriteria analysis applied to aircraft selection, case in Brazilian Navy (Maêda et al., 2021).

Consequently, a gap is identified in the literature regarding the proposition of a model with the use of Rough Sets Theory/Dominance principle, theory for mathematical treatment of data imprecision and generation of decision rules, and Machine Learning for the inference of sustainability classes for electric power companies.

3 Methodology

3.1 Rough sets theory and dominance principle

According to Pawlak (1982, 1991), Pawlak & Sowinski (1994), Pawlak et al. (1995) and Slowinski et al. (2012), the Rough Sets Theory (RST) was originated with Zdzislaw Pawlak, at the beginning of the 1980s, as a mathematical tool to treat imprecision and uncertainty of data. The approach made possible with RST is of fundamental importance for Artificial Intelligence (AI) and the cognitive sciences, especially in the Machine Learning areas, knowledge acquisition, decision analysis, knowledge discovery in databases, specialist systems, decision making support systems, inductive reasoning and pattern recognition. RST does not compete with Fuzzy Logic, with which it is frequently compared, but it complements it.

In any case, these theories are independent approaches of imperfect knowledge. One of the main advantages RST offers is that it doesn't need preliminary or additional information about data, the way probability distribution or pertinence level need in the fuzzy sets theory. There is a concept of Information Systems with: K = (U, A), where U is a finite and nonempty set of objects; A is a finite and nonempty set of attributes, such that a: U → V_a for every a ϵ A, where V_a is a set of values that can be attributed to attribute *a*.

There is also the concept about indiscernibility relation: given an information system (U,A) and for any B ⊆ A, an equivalence relation (or classification, indistinctively) R_B is defined as (Riza et al., 2014), Equation 1:

$$R_B\left(x,y\right) = \{\left(x,y\right)\epsilon U^2 \mid \forall a\epsilon B,\ a\left(x\right) = a\left(y\right)\} \tag{1}$$

If (x,y) ϵ R_B(x,y), then x and y have exactly the same values for attributes in B. According to Pawlak (1991), it is not always possible to express exactly a certain set of objects with the available knowledge. Consequently, it is possible to express a set of objects by other two subsets: lower ($\underline{R}$) and upper ($\overline{R}$) approximations, given that X ⊆ U and an equivalence relation R, Equation 2:

$$\underline{R}X = \cup\{Y\epsilon U \mid R:\ Y \subseteq X\};\overline{R}X = \cup\{Y\epsilon U \mid R:\ Y \cap X \neq \varnothing\} \tag{2}$$

where $\underline{R}X$ is the subset of U elements, that can certainly be classified as X elements; RX is the subset of U elements, that can possibly be classified as X elements. A set is rough with relation to R, if and only if, $\underline{R}X \neq \overline{R}X$.

From the Information System concept, K = (U, A), it can be obtained the concept of decision table: when A is formed by two subsets: C and D, attributes of condition and decision, respectively, (C, D ⊂ A); or, T = (U, A, C, D). The equivalence classes of R_B and R_C relations are known as condition and decision classes, respectively. For each x ϵ U, a function d_x: A → V is associated, such that $d_x(a) = a(x)$, for every a ϵ C ∪ D; the d_x function is known as decision rule in T (Pawlak, 1991). According to Riza et al. (2014), a superreduct is a set of attributes B ⊂ A such that $R_B = R_A$, where R_B and R_A are indiscernibility relations defined by B and A, respectively. If this relation is also minimum, it is a reduct. The intersection of all reducts is a core of an information system. In Pawlak (2000), it is found an example of a Decision Table – Table 3. This table is composed of six stores and four attributes (quantitative or qualitative aspects): E, sellers' autonomy; Q, merchandise quality; L, location with intense traffic; P, result profit or loss. E, Q and L represent condition attributes; P represents a decision attribute.

Table 3. Example table. Source: Pawlak (2000).

Store	E	Q	L	P
1	High	Good	No	Profit
2	Medium	Good	No	Loss
3	Medium	Good	No	Profit
4	None	Medium	No	Loss
5	Medium	Medium	Yes	Loss
6	High	Medium	Yes	Profit

The attributes E, Q and L, it is possible to affirm the following: stores 1 and 6 achieved profit, stores 4 and 5 had losses and stores 2 and 3 cannot be classified (in profit or loss), because they are indiscernible to these attributes. Being so, employing attributes E, Q and L, it can deduce that: stores 1 and 6 certainly made profit, that is, possibly belong to the set {1,3,6}; while stores 1, 2, 3 and 6 possibly had profit, that is, possibly belong to set {1,3,6}. The sets {1,6} and {1,2,3,6} represent, respectively, the lower and upper approximations. Possible rules extracted from Decision Table (Equations 3-5):

$$(E,\ medium)\ and\ (Q,\ good) \rightarrow (P,\ loss) \tag{3}$$

$$(E,\ none) \rightarrow (P,\ loss)\ (or\ if\ E\ =\ none\ then\ loss) \tag{4}$$

$$(E,\ medium)\ and\ (Q,\ medium) \rightarrow (P,\ loss) \tag{5}$$

Being so, a decision rule in S is an expression Φ → Ψ, reading if Φ then Ψ, where Φ ∈ For(C), Ψ ∈ For(D), C and D are attributes of condition and decision, respectively.

Consequently, the RST ignores not only the order of preference in the set of attribute values, but also the "monotonic" relationship of object evaluations regarding the values of condition attributes and the order of preference of the values of decision attributes (classification or degree of preference). Slowinski et al. (2012) presents the Dominance principle – Dominance Rough Sets Approach, DRSA: "objects that possess a better evaluation or that possess at the minimum the same evaluation (decision class), cannot be associated to a worse decision class, all decision criteria considered". The indiscernibility relations are replaced by dominance relations in the decision class approximations. By DRSA, due to the order of preference between the decision classes, the sets become approximations and are known as unions of decision classes: upward and downward classes. The decision rules can be considered under 5 types:

1- certain decision rules-$D_{\geq}$:

   if $f(x,q_1) \geq r_{q1}$ and $f(x,q_2) \geq r_{q2}$ and ... $f(x,q_p) \geq r_{qp}$, then x $Cl_t^{\geq}$;

2- possible decision rules-$D_{\geq}$:

   if $f(x,q_1) \geq r_{q1}$ and $f(x,q_2) \geq r_{q2}$ and ... $f(x,q_p) \geq r_{qp}$, then x possibly belongs to $Cl_t^{\geq}$;

3- certain decision rules-$D_{\leq}$:

   if $f(x,q_1) \leq r_{q1}$ and $f(x,q_2) \leq r_{q2}$ and ... $f(x,q_p) \leq r_{qp}$, then $x \in Cl_t^{\leq}$;

4- possible decision rules-$D_{\leq}$:

   if $f(x,q_1) \leq r_{q1}$ and $f(x,q_2) \leq r_{q2}$ and ... $f(x,q_p) \leq r_{qp}$, then x possibly belongs to $Cl_t^{\leq}$
   where $P = \{q_1, ..., q_p\} \subseteq C$, $(r_{q1}, ..., r_{qp}) \in V_{q1} \times V_{q2} \times ... \times V_{qp}$ and $t \in T$;

5- approximate decision rules-$D_{\leq\geq}$:

   if $f(x,q_1) \geq r_{q1}$ and $f(x,q_2) \geq r_{q2}$ and ... $f(x,q_k) \geq r_{qk}$ and $f(x,q_{k+1}) \leq r_{qk+1}$ and $f(x,q_p) \leq r_{qp}$, then

   $x \in Cl_s \cup Cl_{s+1} \cup ... \cup Cl_t$.

The rules of type 1) and 3) represent certain knowledge extracted from data (ordinal classification, for example); the rules of type 2) and 4) represent possible knowledge; and the rules of type 5) represent doubtful knowledge (are supported by inconsistent objects only).

3.2 Machine learning

Machine Learning refers to the subfield of Artificial Intelligence that aims to project algorithms and allow computers to elaborate behaviours based on empirical data. And, as an instrument, for example, it can find the induction of rules: condition-action rules, decision trees or similar knowledge structures (Langley & Simon, 1995; Pawlak et al., 1995). Softwares learn automatically to recognize complex patterns and to take intelligent decisions based in data. The Machine Learning methods are divided in: supervised (classification cases; training data has a label), non-supervised (clustering case; training data has no label) and semi-supervised (a combination of both previous methods) (Russell & Norvig, 2010; Han et al., 2012).

4 Results and discussion

Initially, there was a collection and treatment of information related to constant indicators in the Reports of Socio Environmental and Economic-Financial Responsibility (RSA), according to the Electricity Sector Accounting Manual (ESAM) (ANEEL, 2015). The ESAM distinguishes indicators regarding sustainability in: directly related (filled column, with reference to the GRI pattern, Global Reporting Initiative/Sustainability Reporting Guidelines & Electric Unity Sector Supplement), in this study identified by the initials “Su”. In March 2021, from a total of 63 (sixty-three) electricity distribution companies, there were 26 (twenty-six) RSA reports regarding the fiscal year of 2019 (ANEEL, 2021). The RSA reports are annual, with demonstrations from the past three fiscal years. For the current study, the 10 biggest electricity distribution companies were considered (EPE, 2020). However, in case of lack of information, the function RANDOM()*(MAXIMUM(interval)-MINIMUM(interval))+MINIMUM(interval)” was used, in Microsoft Excel, to simulate the values of the absent indicators. The indicators were then distributed by environmental, economic, social and corporate governance dimensions, according to the classification established on ESAM. For this study, 5 (five) indicators of each environmental, economic and social dimensions were used, except for corporate governance, where twenty indicators were used. Following, there was an identification as to the type of indicator: if it is “G”, gain (the bigger the value, the best); or if it is “C”; cost or loss (the smaller the value, the best), as shown in the Tables 4-5.

Table 4. Indicators directly related to sustainability – part 1. Sources: GRI (2000), ANEEL (2015), authors.

Indicator Description	Reference	Dimension	Unit	Initials	(G)ain/(C)ost
Direct energy consumption per primary energy source	ANEEL (2015); GRI (2000): EN3	Environmental	kWh	IA1	C
Total water consumption	ANEEL (2015); GRI (2000): EN8	Environmental	m^3	IA2	C
Annual volume of greenhouse gases (CO2, CH4, N2O, HFC, PFC, SF6) emitted into the atmosphere (in tons of equivalent CO2)	ANEEL (2015); GRI (2000): EN16, EN17, EN18	Environmental	t CO2	IA3	C
Total disposal of water, by quality and destination	ANEEL (2015); GRI (2000): EN21	Environmental	m^3	IA4	C
Quantity of residues contaminated by PCB (Ascarel) destined	ANEEL (2015); GRI (2000): EN24	Environmental	t	IA5	C
Number of end users served	ANEEL (2015); GRI (2000):EU3	Economic	un	IE1	G
Number of end users served – Free	ANEEL (2015); GRI (2000): EU3	Economic	un	IE2	G
Energy bought (GWh)	ANEEL (2015); GRI (2000): EU10	Economic	GWh	IE3	C
Energy sold (GWh)	ANEEL (2015); GRI (2000): EU3, 2.7	Economic	GWh	IE4	G
Substations (in units)	ANEEL (2015); GRI (2000): EU1	Economic	un	IE5	G
Number of own employees	ANEEL (2015); GRI (2000): LA1	Social	un	IS5	G
Equivalent Duration Interruption per Consumer Unity “DEC”, company-wide – Calculated value	ANEEL (2015); GRI (2000): EU29	Social	h	IS1	C
Equivalent Duration Interruption per Consumer Unity “DEC”, company-wide – Limit	ANEEL (2015); GRI (2000): EU29	Social	h	IS2	C
Equivalent Frequency Interruption per Consumer Unity “FEC”, company-wide – Calculated value	ANEEL (2015); GRI (2000): EU28	Social	un	IS3	C
Equivalent Frequency Interruption per Consumer Unity “FEC”, company-wide – Limit	ANEEL (2015); GRI (2000): EU28	Social	un	IS4	C

The Table 4 shows the indicators directly related to sustainability in the environmental, economic and social dimensions, considering five indicators of each dimension.

Table 5. Indicators directly related to sustainability, dimension Governance – part 2. Sources: GRI (2000), Pereira & Cândido (2020), authors.

Indicator Description	Reference	Unidade	Initials	(G)ain/(C)ost
Preferencial acts broadcaster.	Pereira & Cândido (2020)	0 or 1	IG01	G
Preferencial shareholders have the right to vote in relevant matters.	Pereira & Cândido (2020)	0 or 1	IG02	G
Disclosure mechanisms about the deliberated themes in the assemblies.	Pereira & Cândido (2020)	0 or 1	IG03	G
Prohibition of loans and guarantees in favour of the controller, the managers and other related parties.	Pereira & Cândido (2020)	0 or 1	IG04	G
Existence of a channel dedicated to anonymous communications that is intended to receive reports, complaints and suggestions.	Pereira & Cândido (2020)	0 or 1	IG05	G
Annual and/or sustainability report with accessibility for people with disabilities.	Pereira & Cândido (2020)	0 or 1	IG06	G
Programs for education on sustainability and the audiences reached.	Pereira & Cândido (2020)	0 or 1	IG07	G
Adherence of voluntary commitments related to sustainable development.	Pereira & Cândido (2020)	0 or 1	IG08	G
Existence of a sustainability committee.	Pereira & Cândido (2020)	0 or 1	IG09	G
Independent audit opinion.	Pereira & Cândido (2020)	0 or 1	IG10	G
Commitment with the fight against corruption that encompasses the internal audience.	Pereira & Cândido (2020)	0 or 1	IG11	G
Commitment with the fight against corruption of company partners.	Pereira & Cândido (2020)	0 or 1	IG12	G
Organizations' governance structure, including Committees under the highest governance body responsible by specific tasks, such as strategic configuration or organizational supervision. Describe the mandate and the composition (including number of independent and/or non-executive members) of such committees and assign any direct responsibility by economic, social and environmental performance matters.	GRI (2000)	0 or 1	IG13	G
For organizations that have a unitary council structure, indicate the number of members from the highest governance body who are independent and/or non-executive members. Declare how the organization defines "independent" and "non-executive". This element applies only to organizations that have unitary council structures.	GRI (2000)	0 or 1	IG14	G
Mechanisms that enable shareholders and employees to supply recommendations or orientations for the highest governance body. It includes a reference to processes related to: • The use of shareholder resolutions or other mechanisms in order to allow the shareholder minority to express opinions to the highest governance body; and • Report and consult employees about work relationships with the representative bodies, such as a "work council" level of organization, and representation of employees of the highest governance body. Identify topics related to economic, environmental, and social performance aspects generated by these mechanisms throughout the period covered by the report.	GRI (2000)	0 or 1	IG15	G
Processes in effect for the highest governance body in order to guarantee that conflicts of interest are avoided.	GRI (2000)	0 or 1	IG16	G

Table 5. Continued…

Indicator Description	Reference	Unidade	Initials	(G)ain/(C)ost
Process to determine the qualifications and experience from the members of the highest level of governance body in order to guide the organization regarding the economic, environmental, and social strategies.	GRI (2000)	0 or 1	IG17	G
Mission statements developed internally or values, codes of conduct and relevant principles for the economic, environmental, and social performance and the status of its implementation. Explain the level in which these: • Are applied in the entire organization in different regions and departments/units; and • Relate to the internationally accepted standards.	GRI (2000)	0 or 1	IG18	G
Procedures from the highest governance body to supervise the organization: economic, environmental, and social performance management, including relevant risks and opportunities for adherence or conformity to internationally accorded standards, code of conduct and principles. It includes the frequency with which the highest governance body evaluated the sustainability performance.	GRI (2000)	0 or 1	IG19	G
Processes to evaluate the highest governance, particularly regarding the economic, environmental and social performance.	GRI (2000)	0 or 1	IG20	G

Specifically, for governance (corporate) dimension indicators, the RSA reports' contents were analysed (read) according to the text mining technique, researched the frequency of words, according to contents from the columns Keyword, Table 6. The reading of these reports was aided with the use of mining text software, KH Coder, version 3.Beta.01a (Higuchi, 2001), through the use of options (Word) Frequency list, searched words and its respective frequency of occurrence, Word association, words strongly associated with other words, KWIC (Key Words in Context) concordance, how the searched words are used in the text, and Co-occurrence network, a net with syntagmatic relations between words (Higuchi, 2001; Zhai & Massung, 2016). This way, in case it found a keyword associated to the indicator in a certain paragraph of the RSA report, and if this paragraph was associated to the description of the indicator, "1" was computed to the indicator (Andreopoulou & Koliouska, 2018). At the end, it obtained a general corporate governance indicator through adding the values 0 or 1 of the respective indicators.

Table 6. Corporate governance indicators. Source: authors.

Indicator Description	Initials	Keywords
Preferencial acts broadcaster.	IG01	preferencials
Preferencial shareholders have the right to vote in relevant matters.	IG02	preferencials
Disclosure mechanisms about the deliberated themes in the assemblies.	IG03	disclosure
Prohibition of loans and guarantees in favour of the controller, the managers and other related parties.	IG04	prohibition
Existence of a channel dedicated to anonymous communications that is intended to receive reports, complaints and suggestions.	IG05	channel
Annual and/or sustainability report with accessibility for people with disabilities.	IG06	report; sustainability
Programs for education on sustainability and the audiences reached.	IG07	education
Adherence of voluntary commitments related to sustainable development.	IG08	adherence
Existence of a sustainability committee.	IG09	committee
Independent audit opinion.	IG10	independent audit

Table 6. Continued…

Indicator Description	Initials	Keywords
Commitment with the fight against corruption that encompasses the internal audience.	IG11	corruption
Commitment with the fight against corruption of company partners.	IG12	corruption
Organizations' governance structure, including Committees under the highest governance body responsible by specific tasks, such as strategic configuration or organizational supervision. Describe the mandate and the composition (including number of independent and/or non-executive members) of such committees and assign any direct responsibility by economic, social and environmental performance matters.	IG13	committee; independent; independence
For organizations that have a unitary council structure, indicate the number of members from the highest governance body who are independent and/or non-executive members. Declare how the organization defines "independent" and "non-executive". This element applies only to organizations that have unitary council structures.	IG14	independent; independence
Mechanisms that enable shareholders and employees to supply recommendations or orientations for the highest governance body. It includes a reference to processes related to: • The use of shareholder resolutions or other mechanisms in order to allow the shareholder minority to express opinions to the highest governance body; and • Report and consult employees about work relationships with the representative bodies, such as a "work council" level of organization, and representation of employees of the highest governance body. Identify topics related to economic, environmental, and social performance aspects generated by these mechanisms throughout the period covered by the report.	IG15	recommendations; recommendation; orientation; orientations
Processes in effect for the highest governance body in order to guarantee that conflicts of interest are avoided.	IG16	conflict; interest
Process to determine the qualifications and experience from the members of the highest level of governance body in order to guide the organization regarding the economic, environmental, and social strategies.	IG17	qualifications; experience
Mission statements developed internally or values, codes of conduct and relevant principles for the economic, environmental, and social performance and the status of its implementation. Explain the level in which these: • Are applied in the entire organization in different regions and departments/units; and • Relate to the internationally accepted standards.	IG18	declarations; declaration; values; conduct; principles
Procedures from the highest governance body to supervise the organization: economic, environmental, and social performance management, including relevant risks and opportunities for adherence or conformity to internationally accorded standards, code of conduct and principles. It includes the frequency with which the highest governance body evaluated the sustainability performance.	IG19	supervision; supervise; risks; adherence; conformity; standars; standard; performance; sustainability
Processes to evaluate the highest governance, particularly regarding the economic, environmental and social performance.	IG20	evaluation; evaluate

Each value of the indicator (v_i) was then normalized (v_{in}), that is, it is found in the interval [0; 1], and aims to obtain a percentage in relation to the maximum found value (v_{max}), if gain or profit indicator, or minimum (v_{min}), if cost or loss indicator: $v_{in} = v_i / v_{max}$ or $v_{in} = 1 / (v_i / v_{min})$, respectively. This way, for each company and for each indicator, it is obtained a relative position (%) to the company that can be considered a paradigm in that indicator, that obtained the maximum or minimum values, according to the type of indicator – gain or cost, respectively. For each company and for each set of indicators of the same dimension (example, social), it was computed an index by simple arithmetic average of the normalized values. At the end. It was obtained indexes in the environmental InA, economic InE, social InS and corporate governance InG dimensions; the simple arithmetic average of these 4 (four) indexes results in the final index InSu, for ten companies, E01 to E10. Class "L", low, "M", medium or "H", high, is obtained considering the index value in the intervals: (0; 0.3333], (0.3333; 0.6666] and (0.6666; 1], respectively, Table 7. There is a crescent order of preference for classes: low, medium and high.

Table 7. Companies, indexes and its respective sustainability classes (Su). Source: authors.

Year	Company	InA	InE	InS	InG	Class A	Class E	Class S	Class G	In Su	Class Su
2019	E01	0.3579	0.8804	0.9320	0.3333	M	H	H	M	0.6259	M
	E02	0.7087	0.6271	0.7807	0.6667	H	M	H	H	0.6958	H
	E03	0.6135	0.6458	0.7865	1.0000	M	M	H	H	0.7615	H
	E04	0.5101	0.7868	0.9566	0.8333	M	H	H	H	0.7717	H
	E05	0.4298	0.6691	0.8427	0.5000	M	H	H	M	0.6104	M
	E06	0.2934	0.7630	0.9051	0.5000	L	H	H	M	0.6154	M
	E07	0.3077	0.7778	0.3908	0.5000	L	H	M	M	0.4941	M
	E08	0.3578	0.5806	0.8075	1.0000	M	M	H	H	0.6865	H
	E09	0.4761	0.4555	0.8119	0.5000	M	M	H	M	0.5609	M
	E10	0.4493	0.6715	0.8557	1.0000	M	H	H	H	0.7441	H
2018	E01	0.4255	0.8646	0.9663	0.3333	M	H	H	M	0.6474	M
	E02	0.7267	0.5913	0.7833	0.6667	H	M	H	H	0.6920	H
	E03	0.6458	0.6352	0.7518	1.0000	M	M	H	H	0.7582	H
	E04	0.5448	0.7604	0.9389	0.8333	M	H	H	H	0.7693	H
	E05	0.4502	0.6211	0.7931	1.0000	M	M	H	H	0.7161	H
	E06	0.4275	0.3845	0.8698	0.5000	M	M	H	M	0.5455	M
	E07	0.4878	0.4053	0.3780	0.1667	M	M	M	L	0.3594	M
	E08	0.4411	0.6381	0.8306	1.0000	M	M	H	H	0.7274	H
	E09	0.5364	0.5886	0.8473	0.5000	M	M	H	M	0.6181	M
	E10	0.4757	0.6904	0.8856	1.0000	M	H	H	H	0.7629	H
2017	E01	0.6396	0.8502	0.9461	0.3333	M	H	H	M	0.6923	H
	E02	0.6730	0.5228	0.7449	0.6667	H	M	H	H	0.6518	M
	E03	1.0000	0.5682	0.6544	1.0000	H	M	M	H	0.8056	H
	E04	0.5287	0.7141	0.9360	0.8333	M	H	H	H	0.7530	H
	E05	0.5594	0.5383	0.3198	0.0000	M	M	L	L	0.3544	M
	E06	0.6396	0.5720	0.7372	0.6667	M	M	H	H	0.6538	M
	E07	0.5999	0.6250	0.3685	0.5000	M	M	M	M	0.5234	M
	E08	0.6273	0.6936	0.6907	1.0000	M	H	H	H	0.7529	H
	E09	0.5879	0.5877	0.8214	0.5000	M	M	H	M	0.6243	M
	E10	0.6936	0.5887	0.7809	1.0000	H	M	H	H	0.7658	H

The collected and simulated values of the indicators for each company, and its respective final classes, Su Class, are found on Table 8, and represents a Decision Table, considering the indicators as attributes or condition criteria and Su class as an attribute or decision criteria.

Table 8. Decision table: sustainability indicators and classes Su. Source: authors.

Year	Company	IA1	IA2	IA3	IA4	IA5	IE1	IE2	IE3	IE4	IE5	IS1	IS2	IS3	IS4	IS5	IG	Class Su
2019	E01	37827660	157795	2548765	126236	80	8537040	1441	47463	25562	409	10.62	10.53	5.05	7.24	4203	0.3333	M
	E02	22979922	81631	141311	235301	17	3049218	1149	19087	14454	170	10.86	11.21	7.52	8.94	3404	0.6667	H
	E03	158055	95513	465755	191577	68	6103473	359	22132	17166	299	12.19	16.46	5.91	8.31	4212	1.0000	H
	E04	22721951	100022	214073	212786	36	4713240	1389	26163	19784	374	9.11	10.06	6.02	7.67	4964	0.8333	H
	E05	24315161	98134	2349742	138027	51	8273647	409	39019	15579	407	10.16	12.43	6.73	7.63	4012	0.5000	M
	E06	19378319	132500	1470540	245000	73	8281213	1426	31595	15425	265	10.69	10.32	5.14	8.06	4055	0.5000	M
	E07	10949199	124483	2116566	218290	76	8364068	1031	28848	23211	255	10.02	28.33	24.32	28.22	1120	0.5000	M
	E08	25211945	107962	1932805	186122	62	3741228	515	40940	24405	281	10.82	14.57	7.24	8.38	4679	1.0000	H
	E09	18723750	90508	958443	145935	37	255045	454	34106	16216	302	11.59	14.82	7.16	7.33	4476	0.5000	M
	E10	16948845	90964	489174	168420	56	7553881	569	39476	17516	372	11.36	12.66	6.19	7.46	4446	1.0000	H

Table 8. Continued…

Year	Company	IA1	IA2	IA3	IA4	IA5	IE1	IE2	IE3	IE4	IE5	IS1	IS2	IS3	IS4	IS5	IG	Class Su
2018	E01	41244038	130514	3391332	104411	80	8409044	1138	45245	25271	400	10.05	10.58	5.06	7.26	4530	0.3333	M
	E02	4272084	59104	159462	171306	50	2975190	947	18572	14251	168	10.68	11.72	7.33	9.39	3333	0.6667	H
	E03	101955	84277	1843294	136922	74	5993104	302	21128	16522	341	14.44	14.5	6.44	8.42	3729	1.0000	H
	E04	10474221	99614	244660	132672	73	4637804	1121	25751	19594	369	10.31	10.44	6.22	8.01	5364	0.8333	H
	E05	39339116	84471	1316268	133178	78	5434772	352	28064	17782	370	11.47	13.56	7.32	7.60	3610	1.0000	H
	E06	39847252	121715	2927338	117504	71	312745	118	35326	22022	199	13.10	12.88	5.77	7.87	5217	0.5000	M
	E07	8726764	87185	571635	170945	58	1503806	194	14609	4452	200	11.86	28.33	24.32	28.22	1120	0.1667	M
	E08	15002688	117408	1057711	120331	74	6627398	890	27592	15994	183	11.05	13.96	6.51	7.49	4016	1.0000	H
	E09	26809396	59417	1039882	125106	72	255045	1006	42668	18277	385	12.50	13.73	7.23	7.35	5278	0.5000	M
	E10	30247811	109055	297812	155485	80	4990851	690	21375	19863	313	10.91	10.46	6.33	7.90	4235	1.0000	H
2017	E01	37476519	76060	133286	60848	90	8337594	989	33173	25091	404	10.83	10.80	5.44	7.60	4394	0.3333	H
	E02	37476519	64789	133286	60848	80	2899170	823	18290	14132	167	12.35	12.11	8.35	10.00	3298	0.6667	M
	E03	71430	26754	133286	60848	76	5900258	246	21383	16264	341	19.83	15.01	8.23	8.84	2897	1.0000	H
	E04	30361008	96237	316296	60848	81	4560493	991	25983	19743	369	10.46	10.88	6.83	8.53	5746	0.8333	H
	E05	30911709	61688	281906	60848	86	211000	203	8400	22003	236	18.74	22.47	25.94	32.04	652	0.0000	M
	E06	7007485	61085	157249	60848	84	4804650	806	24801	17659	173	14.92	12.49	7.77	9.58	3596	0.6667	M
	E07	32264140	78233	190602	60848	80	5958615	837	32437	19500	214	13.52	28.33	24.32	28.22	1120	0.5000	M
	E08	22379199	61181	179791	60848	80	6676763	804	30723	21741	289	17.00	12.90	7.95	9.67	3052	1.0000	H
	E09	9169908	63299	234148	60848	81	255045	817	30250	22332	370	11.85	12.19	8.13	9.08	4781	0.5000	M
	E10	22435685	28000	246230	60848	79	8200036	284	18484	14561	258	18.13	10.88	6.99	9.52	4357	1.0000	H

However, when it is desired to simulate a sustainability class for a certain company, for a future situation and/or when there is not yet all the sustainability indicators set for all the companies taken into account, by means Rough Sets and Machine Learning (ML), it becomes doable the use of past absolute values of those indicators for sustainability class prediction – Figure 2.

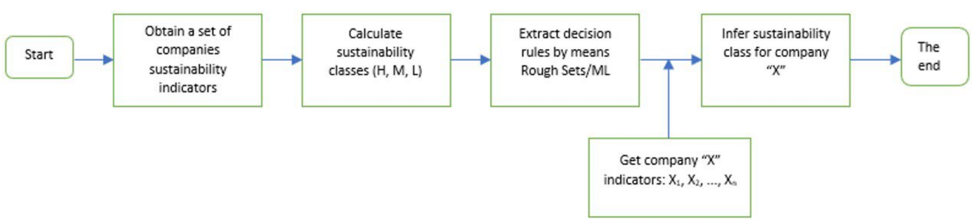

Figure 2. Prediction model for sustainability class by means Rough Sets and Machine Learning. Source: authors.

Once it is had a temporal series of indicators and its respective sustainability classes for each company, the class can be obtained through prediction for a certain company "X". The values for X_1, X_2, ..., X_n can be obtained by regression, for example, or can be directly informed or suggested.

For this research proposal, it was used the RoughSets package in R and it is available on CRAN at http://cran.r-project.org/package=RoughSets. This package makes available four algorithms for inference or generation of decision rules using a script in R: AQRules, CN2Rules, indiscernibilityBasedRules and LEM2Rules (Clark & Niblett, 1989; Michalski et al., 1991; Grzymala-Busse, 1997; Riza et al., 2014, 2019).

In order to calculate the accuracy, correct results / total of instances, of the inference of decision rules algorithms, it was used the information in Table 8 and the method Holdout: randomly, 1/3 of the thirty existing records were chosen as test records (column "#", Table 9); and the rest 2/3 were used as training records (Han et al., 2012). The algorithms in highlight use these training records to infer or generate the decision rules. These rules are then used to predict sustainability classes based on absolute values of indicators (test records) (Riza et al., 2014, 2019). The estimated values for accuracy are found on Table 9. The lowest accuracy 0.6 is a result of indiscernibilityBasedRules and LEM2Rules algorithms, 0.7 of AQRules and the highest 0.9, of CN2Rules algorithm.

Table 9. Use of the Holdout method to determine the accuracy of rule inference algorithms. Source: authors.

Year	Company	#	Class Su	Rule inference algorithms			
				AQRules	**CN2Rules**	**indiscernibilityBasedRules**	**LEM2Rules**
2019	E05	5	M	M	M	H	M
	E08	8	H	M	H	H	H
2018	E01	11	M	M	M	M	M
	E02	12	H	H	H	H	H
	E05	15	H	H	H	H	H
	E07	17	M	H	H	H	H
	E09	19	M	M	M	H	M
2017	E03	23	H	H	H	M	M
	E08	28	H	M	H	H	M
	E10	30	H	H	H	H	M
			Accuracy	**0.7**	**0.9**	**0.6**	**0.6**

And, following, some examples of multicriteria rules that were inferred by the CN2Rules algorithm and its respective rules quality indexes.

1: IF IG is [0.833, Inf] THEN class is H; (supportSize=8; laplace=0.9);

2: IF IA3 is [8.02e+05, Inf] THEN class is M; (supportSize=5; laplace=0.8571);

3: IF IA4 is [-Inf,7.97e+04) and IS3 is [7.4, Inf] THEN class is M; (supportSize=5; laplace=0.8571);

4: IF IA2 is [-Inf,8.25e+04) THEN class is H; (supportSize=2; laplace=0.75).

```
> RI.laplace(rules2)
Rule_1 Rule_2 Rule_3 Rule_4
0.9000000 0.8571429 0.8571429 0.7500000
> RI.support(rules2)
Rule_1 Rule_2 Rule_3 Rule_4
0.40 0.25 0.25 0.10
> RI.confidence(rules2)
Rule_1 Rule_2 Rule_3 Rule_4
1 1 1 1
> RI.lift(rules2)
Rule_1 Rule_2 Rule_3 Rule_4
1.800000 1.714286 1.714286 1.500000
```

By the previous rule "1", for example, and exclusively considering the training instances set, 2/3 of a total of 30, or 20 instances, there are the following quality indexes on this rule: a) supportSize = 8, which means there are 8 or 8/20 instances or companies that satisfy this rule; b) laplace = 0.9, which means this rule has an adjusted accuracy of 90%; c) confidence = 1, which means there is a probability of 100% of the instances that the class is an "H", given that the governance index IG is within the interval [0.833;1]; d) lift = 1.8, which means a company classed as "H" bumps up in almost twice (1.8) the chance of the governance index IG being in the interval [0.833;1] (Dzeroski et al., 1993; Słowiński et al., 2008; Han et al., 2012;

Provost & Fawcett, 2013). There was also, reduct suggestion: a feature subset consisting of 3 attributes: IA2, IE2 e IG.

Based on the proposed model for predicting sustainability classes, Figure 2, there are two cases to analyze, as examples.

a) Case 1

In this first situation, company E05 was used as reference. According to Figure 3, with relation to the year 2019, it was estimated a situation: for cost indicators, it was simulated an increase of 40% and 50% and, for gain indicators, a reduction of similar percentages. Following, the script in R was executed and it was obtained the inference of the following sustainability classes:

			Rule inference algorithms			
Year	**Year/2019**	**Class Su**	**AQRules**	**CN2Rules**	**indiscernibility BasedRules**	**LEM2Rules**
n+1	50%		M	M	H	M
n	40%		H	M	H	M
2019		M				
2018		H				
2017		M				

Figure 3. Inference of sustainability class for company E05, with variations of 40% and 50% in the gain and cost indicators. Source: authors.

Considering it was simulated an increase of 40% and 50% for cost indicators and a reduction for gain indicators with both percentages, the CN2Rules algorithm with higher accuracy shows that the class will probably remain "M".

b) Case 2

In this second case, for company E07 and in relation to the year 2019, it was simulated a favourable situation: ascending the sustainability class from "M" to "H", however, establishing a variation for indicators sets, environmental, economic, social and governance, with the goal of identifying the set that could influence the most the class change – Figure 4. The indicators of a cost nature had a reduction and the ones of gain nature had a raise. Both of 50%.

								Rule inference algorithms			
Year	**Year/2019**	**IA1**	**IA2**	**IA3**	**IA4**	**IA5**	**Class Su**	**AQRules**	**CN2Rules**	**indiscernibility BasedRules**	**LEM2Rules**
n	**50%**	5474599	62241	1058283	109145	38		M	M	H	M
2019		10949199	124483	2116566	218290	76	M				

Figure 4. Inference of sustainability class for company E07, with variation of 50% by indicators set and for the best case, from "M" to "H". Source: authors.

In this case, it was identified that the environmental indicators set, IA1 to IA5, presents the biggest influence over a positive change of sustainability class, probably, from "M" to "H", according to the indiscernibilityBasedRules algorithm, accuracy = 0.6.

Furthermore, the data from Tables 8-9, 2/3 or 20 training records were submitted to the jMAF software, Dominance-Based Rough Set Data Analysis Framework, it was used in order to support the multicriteria analysis, provided by the Institute of Computing Science, Poznan University of Technology (Błaszczyński et al., 2013). There was a suggestion of 250 reducts, Cross validation (DRSA; 10 folds; standard) for training records with 85% or 17/20 accuracy and the confusion matrix – Figure 5. The following sequence was followed in the jMAF software, option Calculate: P-Dominance sets, D_P^+ Calculate dominating set, D_P^- 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.

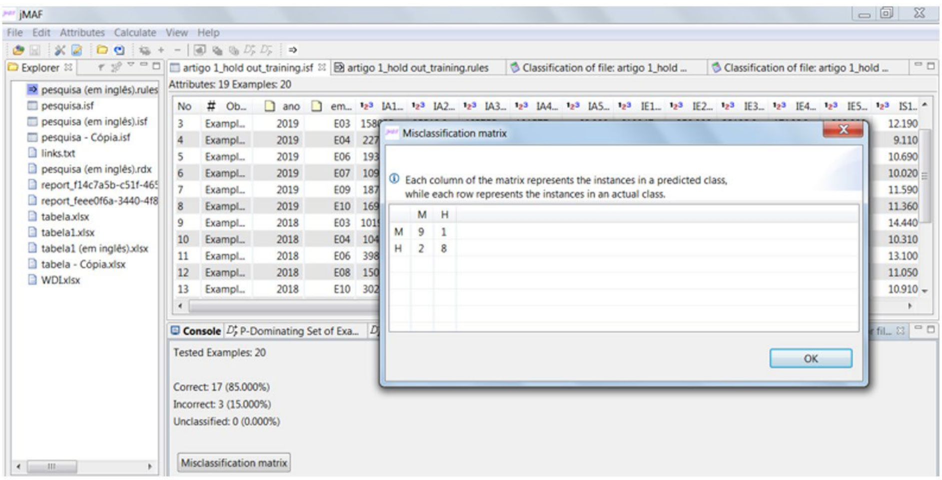

Figure 5. jMAF software, Cross validation with 2/3 training records. Source: authors, adapted from Błaszczyński et al. (2013).

And the multicriteria rules generated with VC-DomLEM algorithm (Błaszczyński et al., 2009, 2011):

#Certain at least rules

1: (IG >= 0.8333) => (class >= H) |CERTAIN, AT_LEAST, H| (Support:8; CoverageFactor: 0.8)

2: (IA5 <= 17.0) => (class >= H) |CERTAIN, AT_LEAST, H| (Support:1; CoverageFactor: 0.1)

3: (IA3 <= 133286.0) & (IE1 >= 8337594.0) => (class >= H) |CERTAIN, AT_LEAST, H| (Support:1; CoverageFactor: 0.1)

#Certain at most rules

4: (IS3 >= 7.77) => (class <= M) |CERTAIN, AT_MOST, M| (Support:6; CoverageFactor: 0.6)

5: (IA2 >= 121715.0) => (class <= M) |CERTAIN, AT_MOST, M| (Support:4; CoverageFactor: 0.4)

6: (IE1 <= 255045.0) => (class <= M) |CERTAIN, AT_MOST, M| (Support:3; CoverageFactor: 0.3)

As a way of interpreting the previous decision rules, according to rule 1, the CoverageFactor of 0.8 indicates that, given that the class is H, there is a conditional probability of 80% that the companies have a governance index (IG) greater than or equal to 0.8333; and, by rule 4, given that the class is at most M, there is a conditional probability of 60% or CoverageFactor 0.6 that, the social indicator 3 (IS3) of the companies is greater than or equal to 7.77 (Pawlak, 2002).

The set of 1/3 or 10 test records was submitted to classification using the previous rules, as shown in Figure 6.

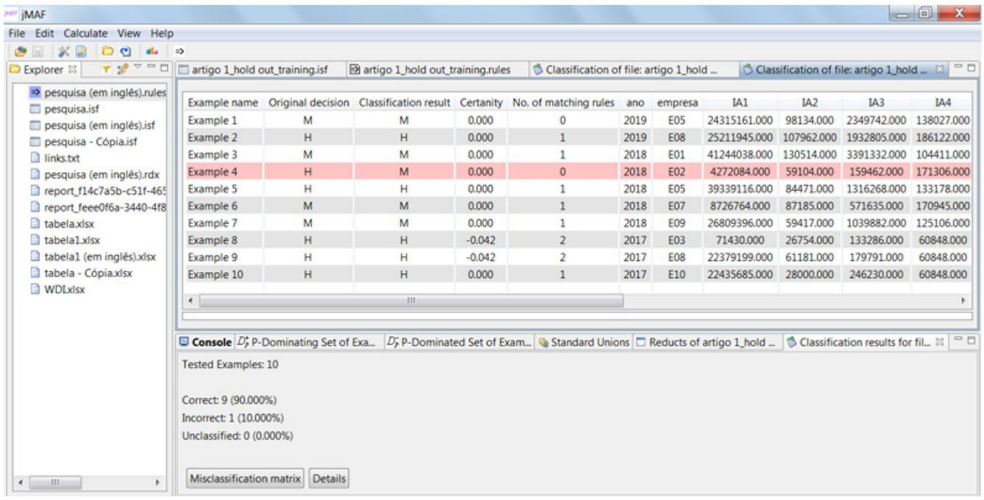

Figure 6. jMAF software, classifying 1/3 test records. Source: authors, adapted from Błaszczyński et al. (2013).

There was a 90% or 9/10 accuracy in classifying these records by VC-DRSA method, Variable Consistency Dominance-based Rough Set Approaches (Greco et al., 2005): example 4, the original decision was “H” and the classification result was “M”. For the others, there was agreement in the classification of records.

5 Conclusion and future studies

The sustainability indicators continue to still be the best way of monitoring and controlling economic, social, environmental, corporate governance impacts etc. accrued from development in its broadest form. As an example of indicators and indexes use, there is the HDI (Human Development Index) to measure the progresses of a population regarding its life expectancy, level of schooling and per capita income; the Covid-19 transmission rate, from Imperial College of London, to evaluate if it is appropriate or not to adopt restrictive measures of mobility in a certain region; index Risk-country; stock exchange index etc. In other words, we use indicators and indexes to guide our decision making in the most diverse circumstances and needs.

In the context of company management and, specifically in this research proposal, electricity distribution companies in Brazil, from a certain universe of indicators, it was presented a proposal for sustainability classes prediction based on past values of indicators directly related to sustainability for a certain company, before a situation in which it is not available the whole indicators set of the other electricity distribution companies.

The two cases that were addressed in the study show that sustainability class prediction based on indicators can collaborate in the anticipation of conditions and/or situations of risk, as well as the opportunity of improvements in a future moment and/or when it is not yet available the entire sustainability indicators set for all the companies considered. Beyond that, the study shows the use of the AQRules, CN2Rules, indiscernibility BasedRules, LEM2Rules and VC-DomLEM algorithms for generation of decision rules in the sustainability classes prediction, based on historical absolute values of the indicators.

Consequently, the contributions of this study are in the model proposition that: a) uses Rough Sets Theory/Dominance principle and Machine Learning to extract decision rules and infer sustainability classes from historical series of indicators and simulated values, aiming to obtain better risk management in the economic, social, environmental and corporate governance dimensions of a company against competitors; b) establishes sustainability classes for companies in order to identify possible links on aspects of sustainability and performance between companies belonging to the same class, as opposed to simple ranking; c) allows to relate condition and decision criteria in decision rules, for example, by coverage factor, and consequently, obtain patterns in data, without referring to a priori and posterior probabilities, as in Bayesian analysis.

As future study, there is a possibility of broadening this set of indicators, vis-à-vis with the existence of other quantitative and qualitative indicators on the electric sector.

Acknowledgements

The authors are grateful to the Institute of Computing Science of Poznan University of Technology, Poland, for permission of using the jMAF software.

References

Adedeji, P. A., Akinlabi, S. A., Madushele, N., & Olatunji, O. O. (2021). Hybrid neurofuzzy investigation of short-term variability of wind resource in site suitability analysis: a case study in South Africa. *Neural Computing & Applications*, 33(19), 13049-13074. http://dx.doi.org/10.1007/s00521-021-06001-x.

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*, 5(2), 47-62. http://dx.doi.org/10.24857/rgsa.v5i2.424.

Agência Nacional de Energia Elétrica – ANEEL. (2015). *Manual de contabilidade do setor elétrico*. Brasília: ANEEL.

Agência Nacional de Energia Elétrica – ANEEL. (2021). *Distribuição*. Retrieved in 2021, March 30, from https://www.gov.br/aneel/pt-br/centrais-de-conteudos/relatorios-e-indicadores/distribuicao.

Ahmad, T., Zhang, D., Huang, C., Zhang, H., Dai, N., Song, Y., & Chen, H. (2021). Artificial intelligence in sustainable energy industry: status quo, challenges and opportunities. *Journal of Cleaner Production*, 289, 125834. http://dx.doi.org/10.1016/j.jclepro.2021.125834.

Ahmadi, M., Soofiabadi, M., Nikpour, M., Naderi, H., Abdullah, L., & Arandian, B. (2022). Developing a deep neural network with fuzzy wavelets and integrating an inline PSO to predict energy consumption patterns in urban buildings. *Mathematics*, 10(8), 1270. http://dx.doi.org/10.3390/math10081270.

Akay, H. (2021). Flood hazards susceptibility mapping using statistical, fuzzy logic, and MCDM methods. *Soft Computing*, 25(14), 9325-9346. http://dx.doi.org/10.1007/s00500-021-05903-1.

Al-Barakati, A., Mishra, A. R., Mardani, A., & Rani, P. (2022). An extended interval-valued Pythagorean fuzzy WASPAS method based on new similarity measures to evaluate the renewable energy sources. *Applied Soft Computing*, 120, 108689. http://dx.doi.org/10.1016/j.asoc.2022.108689.

Andreopoulou, Z., & Koliouska, C. (2018). Benchmarking internet promotion of renewable energy enterprises: is sustainability present? *Sustainability*, 10(11), 4187. http://dx.doi.org/10.3390/su10114187.

Angilella, S., & Pappalardo, M. R. (2021). Assessment of a failure prediction model in the European energy sector: a multicriteria discrimination approach with a PROMETHEE based classification. *Expert Systems with Applications*, 184, 115513. http://dx.doi.org/10.1016/j.eswa.2021.115513.

Associação Brasileira de Normas Técnicas – ABNT. (2010). *NBR ISO 26000-diretrizes sobre responsabilidade social*. São Paulo: ABNT.

Błaszczyński, J., Greco, S., Matarazzo, B., Słowinski, R., & Szelag, M. (2013). jMAF-dominance-based rough set data analysis framework. In A. Skowron & Z. Suraj (Eds.), *Rough sets and intelligent systems - Professor Zdzisław Pawlak in memoriam* (vol. 1, pp. 185-209, Intelligent Systems Reference Library, no. 42). Berlin: Springer. http://dx.doi.org/10.1007/978-3-642-30344-9_5.

Błaszczyński, J., Słowinski, R., & Szelag, M. (2009). *VC-DomLEM: rule induction algorithm for variable consistency rough set approaches*. Poznań: University of Technology. Technical Report RA-07/09.

Błaszczyński, J., Słowinski, R., & Szelag, M. (2011). Sequential covering rule induction algorithm for variable consistency rough set approaches. *Information Sciences*, 181(5), 987-1002. http://dx.doi.org/10.1016/j.ins.2010.10.030.

Bock, K. W., Coussement, K., & Lessmann, S. (2020). Cost-sensitive business failure prediction when misclassification costs are uncertain: a heterogeneous ensemble selection approach. *European Journal of Operational Research*, 285(2), 612-630. http://dx.doi.org/10.1016/j.ejor.2020.01.052.

Brasil. Agência Nacional de Energia Elétrica – ANEEL. (2014). *Resolução normativa nº 605, de 11/03/2014* (seção 1, nº 53, p. 43). Brasília, DF: Diário Oficial da República Federativa do Brasil.

Buțurache, A.-N., & Stancu, S. (2022). Building energy consumption prediction using neural-based models. *International Journal of Energy Economics and Policy*, 12(2), 30-38. http://dx.doi.org/10.32479/ijeep.12739.

Caiado, R. G. G., Lima, G. B. A., Gavião, L. O., Quelhas, O. L. G., & Paschoalino, F. F. (2017). Sustainability analysis in electrical energy companies by similarity technique to ideal solution. *IEEE Latin America Transactions*, 15(4), 675-681. http://dx.doi.org/10.1109/TLA.2017.7896394.

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 Building*, 216, 109964. http://dx.doi.org/10.1016/j.enbuild.2020.109964.

Clark, P., & Niblett, T. (1989). The CN2 induction algorithm. *Machine Learning*, 3(4), 261-283. http://dx.doi.org/10.1007/BF00116835.

Colla, M., Ioannou, A., & Falcone, G. (2020). Critical review of competitiveness indicators for energy projects. *Renewable & Sustainable Energy Reviews*, 125, 109794. http://dx.doi.org/10.1016/j.rser.2020.109794.

Daugavietis, J. E., Soloha, R., Dace, E., & Ziemele, J. (2022). A comparison of multi-criteria decision analysis methods for sustainability assessment of district heating systems. *Energies*, 15(7), 2411. http://dx.doi.org/10.3390/en15072411.

Dzeroski, S., Cestnik, B., & Petrovski, I. (1993). Using the m-estimate in rule induction. *Journal of Computing and Information Technology*, 1, 37-46.

Elkington, J. (2020). *Green swans: the coming boom in regenerative capitalism*. New York: Fast Company Press.

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.

Empresa de Pesquisa Energética – EPE. Ministério de Minas e Energia. (2020). *Anuário Estatístico de Energia Elétrica 2020, ano base 2019*. Brasília: EPE. Retrieved in 2021, March 30, from https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-160/topico-168/Anu%C3%A1rio%20Estat%C3%ADstico%20de%20Energia%20El%C3%A9trica%202020.pdf.

Ervural, B. C., Evren, R., & Delen, D. (2018a). A multi-objective decision-making approach for sustainable energy investment planning. *Renewable Energy*, 126, 387-402. http://dx.doi.org/10.1016/j.renene.2018.03.051.

Ervural, B. C., Zaim, S., Demireld, O. F., Aydin, Z., & Delen, D. (2018b). An ANP and fuzzy TOPSIS-based SWOT analysis for Turkey's energy planning. *Renewable & Sustainable Energy Reviews*, 82, 1538-1550. http://dx.doi.org/10.1016/j.rser.2017.06.095.

Franceschini, F., Galetto, M., & Maisano, D. (2019). *Designing performance measurement systems: theory and practice of key performance indicators*. Cham: Springer, http://dx.doi.org/10.1007/978-3-030-01192-5.

Global Reporting Initiative – GRI. (2000). *Sustainability reporting guidelines & electric utility sector supplement, RG version 3.0/EUSS*. Amsterdam: GRI.

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(1), 204-211. http://dx.doi.org/10.1016/j.ejor.2007.10.046.

Gomes, L. F. A. M., & Rangel, L. A. D. (2009b). Determining the utility functions of criteria used in the evaluation of real estate. *International Journal of Production Economics*, 117(2), 420-426. http://dx.doi.org/10.1016/j.ijpe.2008.12.006.

Greco, S., Matarazzo, B., Slowinsk, R., & Stefanowski, J. (2005). Variable consistency model of dominance-based rough sets approach. In W. Ziarko & Y. Yao (Eds.), *Rough sets and current trends in computing: second international conference, RSCTC 2000 Banff, Canada, October 16-19, 2000 revised papers* (pp. 170-181, Lecture Notes in Artificial Intelligence). Berlin: Springer.

Grzymala-Busse, J. W. (1997). A new version of the rule induction system LERS. *Fundamenta Informaticae*, 31(1), 27-39. http://dx.doi.org/10.3233/FI-1997-3113.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: concepts and techniques* (3rd ed.). Waltham: Morgan Kaufmann Publishers/Elsevier.

Higuchi, K. (2001). *KH coder*. Retrieved in 2019, October 9, from http://khcoder.net/en/#

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(3), 419-431. http://dx.doi.org/10.1016/j.techfore.2012.10.005.

Langley, P., & Simon, H. A. (1995). Applications of machine learning and rule induction. *Communications of the ACM*, 38(11), 54-64. http://dx.doi.org/10.1145/219717.219768.

Liu, G., Xiao, F., Lin, C.-T., & Cao, Z. (2020). A fuzzy interval time series energy and financial forecasting model using network-based multiple time-frequency spaces and the induced ordered weighted averaging aggregation operation. *IEEE Transactions on Fuzzy Systems*, 28(11), 2677-2690. http://dx.doi.org/10.1109/TFUZZ.2020.2972823.

Maêda, S. M. N., Costa, I. P. A., Castro, M. A. P. Jr., Fávero, L. P., Costa, A. P. A., Corriça, J. V. P., Gomes, C. F. S., & Santos, M. (2021). Multi-criteria analysis applied to aircraft selection by Brazilian Navy. *Production*, 31, e20210011. http://dx.doi.org/10.1590/0103-6513.20210011.

Michalski, R. S., Kaufman, K., & Wnek, J. (1991). *The AQ family of learning programs: a review of recent developments and an exemplary application*. Fairfax: George Mason University. Reports of Machine Learning and Inference Laboratory.

Mohammadiun, S., Hu, G., Gharahbagh, A. A., Mirshahi, R., Li, J., Hewage, K., & Sadiq, R. (2021). Optimization of integrated fuzzy decision tree and regression models for selection of oil spill response method in the Arctic. *Knowledge-Based Systems*, 213, 106676. http://dx.doi.org/10.1016/j.knosys.2020.106676.

Mousavi, M. M., & Lin, J. (2020). The application of PROMETHEE multi-criteria decision aid in financial decision making: case of distress prediction models evaluation. *Expert Systems with Applications*, 159, 113438. http://dx.doi.org/10.1016/j.eswa.2020.113438.

Nedjah, N., Azevedo, V. R., & Mourelle, L. D. M. (2022). Client profile prediction using convolutional neural networks for efficient recommendation systems in the context of smart factories. *Enterprise Information Systems*, 16(10-11), 1653-1693. http://dx.doi.org/10.1080/17517575.2020.1856423.

Organização das Nações Unidas – ONU. (2015). *Transformando nosso mundo: a agenda 2030 para o desenvolvimento sustentável*. Retrieved in 2020, March 5, https://www.mds.gov.br/webarquivos/publicacao/Brasil_Amigo_Pesso_Idosa/Agenda2030.pdf.

Ozdemir, C., Onar, S. C., Bagriyanik, S., Kahraman, C., Akalin, B. Z., & Öztayşi, B. (2021). Estimating shopping center visitor numbers based on a new hybrid fuzzy prediction method. *Journal of Intelligent & Fuzzy Systems*, 42(1), 63-76. http://dx.doi.org/10.3233/JIFS-219175.

Panchal, D., Chatterjee, P., Pamucar, D., & Yazdani, M. (2022). A novel fuzzy-based structured framework for sustainable operation and environmental friendly production in coal-fired power industry. *International Journal of Intelligent Systems*, 37(4), 2706-2738. http://dx.doi.org/10.1002/int.22507.

Pawlak, Z. (1982). Rough sets. *IInternational Journal of Computer & Information Sciences*, 11(5), 341-356. http://dx.doi.org/10.1007/BF01001956.

Pawlak, Z. (1991). *Rough sets, theoretical aspects of reasoning about data*. Dordrecht: Kluwer Academic Publishers.

Pawlak, Z. (2000). Rough sets and decision analysis. *Information Systems & Operational Research*, 38(3), 132-144. http://dx.doi.org/10.1080/03155986.2000.11732405.

Pawlak, Z. (2002). Rough sets, decision algorithms and Bayes' theorem. *European Journal of Operational Research*, 136(1), 181-189. http://dx.doi.org/10.1016/S0377-2217(01)00029-7.

Pawlak, Z., & Sowinski, R. (1994). Rough set approach to multi-attribute decision analysis. *European Journal of Operational Research*, 72(3), 443-459. http://dx.doi.org/10.1016/0377-2217(94)90415-4.

Pawlak, Z., Grzymala-Busse, J., Slowinski, R., & Ziarko, W. (1995). Rough sets. *Communications of the ACM*, 38(11), 88-95. http://dx.doi.org/10.1145/219717.219791.

Paz, F. J., & Kipper, L. M. (2016). Sustentabilidade nas organizações: vantagens e desafios. *Revista Gestão da Produção Operações e Sistemas*, 11(2), 85-102.

Pereira, D. A. M., Santos, M., Costa, I. P. A., Moreira, M. A. L., Terra, A. V., Rocha, C. S. Jr., & Gomes, C. F. S. (2022). Multicriteria and statistical approach to support the outranking analysis of the OECD countries. *IEEE Access: Practical Innovations, Open Solutions*, 10, 69714-69726. http://dx.doi.org/10.1109/ACCESS.2022.3187001.

Pereira, F. No., & 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. http://dx.doi.org/10.12660/rgplp.v19n2.2020.80610.

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

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*, 12(8), 3307. http://dx.doi.org/10.3390/su12083307.

Rigo, P. D., Rediske, G., Rosa, C. B., Gastaldo, N. G., Michels, L., Neuenfeldt, A. L. Jr., & Siluk, J. C. M. (2020). Renewable energy problems: exploring the methods. *Sustainability*, 12(23), 10195.

Riza, L. S., Janusz, A., Bergmeir, C., Cornelis, C., Herrera, F., Slezak, D., & Benitez, J. M. (2014). Implemeting algorithms of rough set theory and fuzzy rough set theory in the R package “roughsets”. *Information Sciences*, 287, 68-89. http://dx.doi.org/10.1016/j.ins.2014.07.029.

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*. Retrieved in 2021, December 3, from https://cran.r-project.org/web/packages/RoughSets/RoughSets.pdf

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J., & Bengio, Y. (2022). Tackling climate change with machine learning. *ACM Computing Surveys*, 55(2), 42. http://dx.doi.org/10.1145/3485128.

Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: a modern approach*. Hoboken: Prentice Hall.

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, 101380. http://dx.doi.org/10.1016/j.seta.2021.101380.

Samal, S., & Dash, R. (2021). A TOPSIS-ELM framework for stock index price movement prediction. *Intelligent Decision Technologies*, 15(2), 201-220. http://dx.doi.org/10.3233/IDT-200013.

Słowiński, R., Greco, S., & Matarazzo, B. (2008). Dominance-based rough set approach to multiple criteria decision support. *Troina*, 2, 9-56.

Slowinski, R., Greco, S., & Matarazzo, B. (2012). Rough set and rule-based multicriteria decision aiding. *Pesquisa Operacional*, 32(2), 213-270. http://dx.doi.org/10.1590/S0101-74382012000200001.

Sun, X., Hao, J., & Li, J. (2022). Multi-objective optimization of crude oil-supply portfolio based on interval prediction data. *Annals of Operations Research*, 309(2), 611-639. http://dx.doi.org/10.1007/s10479-020-03701-w.

Tajbakhsh, A., & Shamsi, A. (2019). A facility location problem for sustainability-conscious power generation decision makers. *Journal of Environmental Management*, 230, 319-334. http://dx.doi.org/10.1016/j.jenvman.2018.09.066. PMid:30293017.

Vargas-Solar, G., Khalil, M., Espinosa-Oviedo, J. A., & Zechinelli-Martini, J.-L. (2022). GREENHOME: a household energy consumption and CO_2 footprint metering environment. *ACM Transactions on Internet Technology*, 22(3), 72. http://dx.doi.org/10.1145/3505264.

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, 127042. http://dx.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 & Industrial Engineering*, 144, 106487. http://dx.doi.org/10.1016/j.cie.2020.106487.

Wu, J., & Wu, T. (2012). Sustainability indicators and indices: an overview. In C. N. Madu & C. Kuei (Eds.), *Handbook of sustainable management* (pp. 65-86). London: Imperial College Press. http://dx.doi.org/10.1142/9789814354820_0004.

Yang, Y., Chen, B., & Yang, Z. (2020). An algorithm for ordinal classification based on pairwise comparison. *Journal of Classification*, 37(1), 158-179. http://dx.doi.org/10.1007/s00357-019-9311-4.

Zhai, C., & Massung, S. (2016). *Text data management and analysis: a practical introduction to information retrieval and text mining*. New York: Association for Computing Machinery/Morgan & Claypool Publishers/University of Waterloo.

Zhou, P., Zhou, P., Yüksel, S., Dinçer, H., & Uluer, G. S. (2019). Balanced scorecard-based evaluation of sustainable energy investment projects with IT2 fuzzy hybrid decision making approach. *Energies*, 13(1), 82. http://dx.doi.org/10.3390/en13010082.

Authors contribution

Ayrton Benedito Gaia do Couto worked on the conceptualization, theoretical-methodological approach, theoretical review, collection of studies, writing and final review of the manuscript. Luis Alberto Duncan Rangel worked on the writing and final revision of the manuscript.