IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 January 2023, accepted 16 February 2023, date of publication 28 February 2023, date of current version 3 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3250262

RESEARCH ARTICLE

Method to Produce More Reasonable Candidate Solutions With Explanations in Intelligent Decision Support Systems

FLÁVIO ROSENDO DA SILVA OLIVEIRA[1,2] AND FERNANDO BUARQUE DE LIMA NETO[1], (Senior Member, IEEE)
[1]Polytechnic School of Engineering, University of Pernambuco (UPE), Recife, Pernambuco 50100-010, Brazil
[2]Federal Institute of Education, Science and Technology, Paulista, Pernambuco 53441-601, Brazil

Corresponding author: Flávio Rosendo Da Silva Oliveira (frso@ecomp.poli.br)

This work was supported by the University of Pernambuco (UPE) and the Improvement Coordination of Higher Education Personnel—Brazil (CAPES) under Grant 001.

ABSTRACT The integration of Artificial Intelligence techniques into Decision Support Systems yields effective solutions to decision problems, especially when complex scenarios are at hand. However, the use of intelligent black-box models can hinder the decision support system's potential to be fully adopted because opaque processes raise suspicions and doubts among careful decision makers. Moreover, appropriate and comprehensible explanations may foster trustworthiness and allow for reasonable adjustments or even corrections. This work proposed an approach that incorporates three reasonability aspects into Decision Systems: feasibility, rationality, and plausibility. Thus, by providing decision makers with reasonable candidate solutions for a complex problem, they are expected to perform their tasks more effectively (i.e. decide with more efficiency as well as efficacy). The new approach is accompanied by two proofs of concept in the health and public security areas. Comparative results using random and rational approaches, including the simulation of distinct user profiles, are presented. The proposed approach achieved superior metrics with regard to feasibility and plausibility, suggesting that this proposition can be applied to real-world applications.

INDEX TERMS Decision support systems, machine learning, explainable artificial intelligence, reasonability.

I. INTRODUCTION

Artificial Intelligence (AI) models have been applied in many areas, such as manufacturing [1], B2B enterprises [2], digital forensics [3], transportation planning [4], and health sciences [5], [6]. Despite the immense potential gains generated by the hybridization of AI and Information Systems (IS), primarily the augmentation of user cognitive capabilities [7], some relevant challenges may also arise.

Among the chief limitations, when intelligent opaque-box models are employed in Decision Support Systems (DSS) to handle high-impact problems [8], for example, decisions involving human lives or high financial value, the lack of transparency is a central concern [9]. On one hand, there is a growing trend regarding regulatory measures [10] to ensure system transparency among different purposes and contexts of use [11]. However, it is difficult to accept that a Decision Maker (DM) might properly trust a DSS based on opaque model inferences [12].

The eXplainable Artificial Intelligence (XAI) area [13] has emerged in recent years with the aim of increasing the transparency of opaque-box models. Among other aims, XAI techniques are geared towards mitigating user doubts about the inner functioning of models. However, higher aspects concerning explanations [14], such as comprehensibility and proper treatment of human factors, are still mostly absent [15] in the vast majority of XAI studies.

The associate editor coordinating the review of this manuscript and approving it for publication was Turgay Celik.

Since most XAI explanations seem to leave the DM out of explanation generation [16], it would not be uncommon to face situations when the generated explanation is not fully comprehended (by the DM). Even when the explanation is comprehended, the proposed solution obtained from the intelligent models might not seem plausible or appropriate. The main concern here is that inappropriate pairs of solution-explanations may hinder the proper use of DSS. Our hypothesis is that it is possible to obtain appropriate and user-centered solutions (i.e. qualitative) with adequate explanations from Intelligent DSS without excessively compromising the objective metrics of decisions.

The key objective of this study is to address the existing gap in the XAI literature, approximating system explanations to the DM, while allowing DM preferences to be considered in the DSS solution generation. So, this research article proposes a method to generate more reasonable solutions accompanied by explanations within the scope of an Intelligent DSS. The reasonability of such envisioned solutions derives from the combination of aspects related to: (i) feasibility– how much the solution considers problem constraints; (ii) rationality – how well the proposed solution is aligned with the decision utility; and (iii) plausibility – how much this solution seems to be appropriate, with respect to DM preferences and/or mental model of problem resolution. In short, the proposed approach comprises the characterization of what a reasonable decision is, how to produce these decisions, and how to obtain explanations about them.

The remainder of this article is structured as follows: Section II contains background comments and concepts, comprising DSS, main XAI concepts, and work related to user-centered decisions and explanations. Section III contains the proposed method, encompassing an approach for reasonable decision-making and an explanation of how it is computationally supported. Section IV describes the relevant details of two proofs of concept regarding a model-based DSS [17] and a compound DSS, each of which is created in response to real-world problems in health and public security areas, respectively. Both included comments regarding the experimental results. Finally, in Section V, the conclusion, discussion, and future work are presented.

II. BACKGROUND

This section introduces relevant concepts used as a theoretical foundation for this study. Subsection II-A deals specifically with DSS, whereas subsection II-B deals with the XAI field and some connections with DSS. Finally, subsection II-C presents the main influencing concepts and approaches.

A. DECISION SUPPORT SYSTEMS

Semi-structured problems [18] are characterized by having many options to be analyzed, typically within a short period of time, and by having a relatively high impact on the implemented decision. This category of problems used to be solved by combining the expertise of a DM with the processing and analytical capabilities of computers. The Information System used in many of these tasks is called Decision Support Systems, as information and suggestions/hints are provided to the DM.

A classical architecture of DSS, such as that proposed by Sprague and Watson [19], comprises: (i) a Database Manager, to deal with available raw or processed data; (ii) a Decision Model Manager (DMM), aimed at dealing with available techniques or methods to solve problems; and finally, (iii) a User Interface Manager, responsible for mediating the inputs and outputs from and to DMs. This architecture was later improved by Watson [20] to encompass new technologies, such as Artificial Intelligence and Data Lakes, among other available options to improve DSSs. This work contribution is focused on boosting the Database Manager to create intelligent models mediated by a special type of DMM, which is detailed in Section III.

Aqel et al. [8] synthesized a broad categorization of DSS by mode of assistance, orientation, user relationship, scope of use, focus area, type, and frequency of decision-making, showing how diverse and consolidated DSS are. According to Hasan et al. [21], DSS is summarized in three categories when observing the focus area: model focus, data focus, and knowledge focus. By observing this categorization, it is possible to infer that each type of DSS might have different explanatory needs, especially those that employ black-box models. In Section IV, two proofs of concept are presented to illustrate the applicability of this work proposal to address different types of DSS.

B. EXPLAINABLE ARTIFICIAL INTELLIGENCE

Although XAI is an area with a growing volume of contributions, it is possible to identify in the literature [22] at least two major approaches to explain opaque intelligent models: intrinsic and post-hoc. The post-hoc approach aims to increase the transparency of an intelligent opaque-box model trained to minimize error and is not concerned with the explanation capability. The intrinsic approach, on the other hand, consists of designing natively explainable models, addressing different needs and ways of providing explanations.

Arrieta et al. [23] categorized black-box model explanation methods into six main types. These explanation methods are: (i) attribute impact over model inference, (ii) relevance of attributes in model inference, (iii) supporting examples and counterexamples, (iv) explainable texts, (v) simplification of complex models, and (vi) explanation by data visualization. Although these explanation methods are mostly focused on single models, in the scope of this work, they were used as part of higher-level explanations, as detailed in Section III.

Even with the wide range of approaches and types of explanations found in the literature, there is a growing concern regarding how much XAI focus is put into explaining and how individual decisions are reached by solely considering the black-box inference. Liao et al. [15] explained that there are few shared practices regarding the design of user-friendly XAI applications and that the suitability of explanations depends on user-specific questions for the application. Carvalho et al. [24] also suggested that the explainability of AI systems must be considered in the context of the explanation needs, problem domain, and user types. Shneider and Handali [25] addressed the need to bring together explanations and receptors of explanations, proposing a conceptualization of explanation personalization. This work proposal is based on the premise that good solutions suggested by a DSS must consider both objective and subjective measures. In addition, these solutions must be accompanied by explanations suitable for the specific DM operating the DSS. Dikmen and Burns [26] explored a human-centered approach to XAI that integrates domain knowledge. It was shown that, especially for less experienced users, when domain knowledge was available, there was less reliance on AI, especially when it was incorrect.

To conclude this brief account, Coussement and Benoit [27] introduced the concept of an Interpretable DSS. The Interpretable DSS is proposed as a combination of interpretable data science and improved decision making. The five properties of Interpretable DSS are: (i) performance, (ii) scalability, (iii) comprehensibility, (iv) justifiability, and (v) actionability. This work is aligned with the proposal of an Interpretable DSS as one possible approach to build such systems.

C. MAIN INFLUENCING CONCEPTS AND APPROACHES

Complementing DSS functionality with explanations is a relevant research topic with a broad spectrum of applications, ranging from agriculture [28] to medical systems [29] and information security [30]. As each application scenario has specific explanation requirements and processing strategies, the works closest to our proposal are presented next.

As shown in Figure 1, Gunning and Aha [31] presented an ontology comprising a given user, receiving an explanation, and the XAI process and actions that might lead to the appropriate use of an XAI system. The main concepts contained in this ontology which influenced this work are the 'Test of Understanding', 'Test of Performance' and 'Test of Satisfaction'. In the scope of this work, a reasonable decision must: (i) pass the 'Test of Performance', being a good solution with regard objective perspective (e.g. decision utility and problem restrictions), and (ii) must pass the 'Test of Comprehension', as well as the 'Test of Satisfaction'. This means that being comprehensible and considered adequate according to subjective DM preferences about solutions and/or methods for problem resolution. The characterization of reasonable decisions that is going to be detailed in subsection III-A, is related to these three tests.

Other studies have addressed explainability in the context of DSS. The work of Buron Brarda et al. [32] proposed an approach to supply argument-based multicriteria DSS with conditional preferences and explainable answers. In Brarda's work, graphs are simultaneously the solution and its explanation and are generated by a rule-based reasoning process. This contrasts with this work, as graphs are used to perform inferences and assessments about these inferences, which are collected to create an auxiliary structure used to explain the decision. Thus, in this proposal, there is a clear separation between the solution and its explanation, allowing them to be applied to more types of DSS, even those that do not employ graphs or rules, in the inference engine.

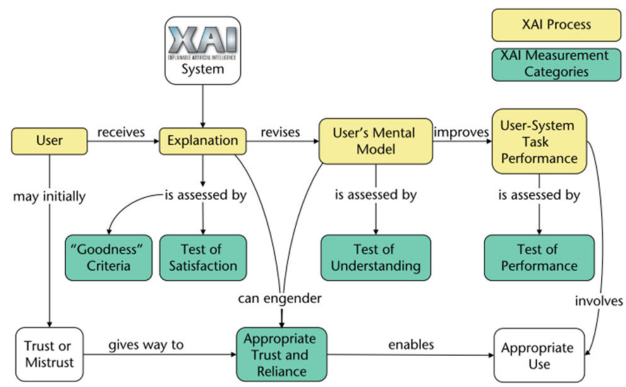


FIGURE 1. Ontology of how explanations might lead to appropriate DSS use. Source: [31].

The work of Dazeley et al. [16] states that explanations might be used in different levels of abstraction, aiming to provide human-aligned conversational explanations and, thus, trying to overcome some current XAI limitations. This work can also be considered an instantiation of Dazeley's work, which is primarily theoretical in the sense that the structure used to explain solutions has multiple levels. These multiple levels and aspects can be used to derive graphical or textual explanations to address multiple levels of decision explanation dialogues.

III. REASONABLE DECISION - CONCEPTS AND PROPOSED APPROACH

The proposed reasonable decision-making method involves the generation of reasonable solutions accompanied by auxiliary data structures used to explain each of these solutions. The following subsections discuss and present: (subsection III-A) conceptual characterization of reasonable solutions; (subsection III-B) overview of the proposed reasonable decision method; (subsection III-C) the generation mechanism of reasonable solutions and (subsection III-D) mechanisms for obtaining reasonable decision structures and explanations.

A. CONCEPTUAL CHARACTERIZATION OF REASONABLE SOLUTIONS

Reasonability, as proposed in this study, is a property that can be attributed to a given solution for a decision problem. It is characterized by three aspects: (i) feasibility, (ii) rationality, and (iii) plausibility, which together are deemed to ascribe much-needed qualitative value to final users.

The feasibility aspect is related to how much a given solution considers available resources and/or how much it follows the limits imposed by the problem itself or inherent context restrictions. The rationality aspect is related to the direct connection of a given solution when observing its capability of solving the problem as posed with respect to any appropriate metric or indicator that can objectively assess it. Finally, the plausibility aspect is related to how much a decision explanation can be comprehended by the specific DM as well as how much a given solution is aligned with the DM preferences and/or mental model of problem resolution. These three aspects were partially drawn from the work of Gunning and Aha [31] regarding proper use and trust in XAI systems, as discussed in subsection II-C.

In Figure 2, a Venn diagram is provided so that various degrees of reasonability (RS), as proposed here, are represented as intersections of these three concepts. Some adjacent interpretations are: (R+P) When a solution is Rational and Plausible but not Feasible, it will likely seem to have adequate utility (R) and will make sense to the DM (P), but it will not be doable, possibly due to not respecting problem restrictions or constraints (not F); (F+R) When a solution is Feasible and Rational but not Plausible, it will likely seem to be doable (F) and having adequate utility (R) but the DM will probably not comprehend its explanation and/or trust it enough to select it (not P); and, (F+P) when a solution is Feasible and Plausible but not Rational, it will likely seem a good pick that makes sense to the DM (P) and could be implemented (F) but lacks quantitative evidence for adequately solving the problem (not R).

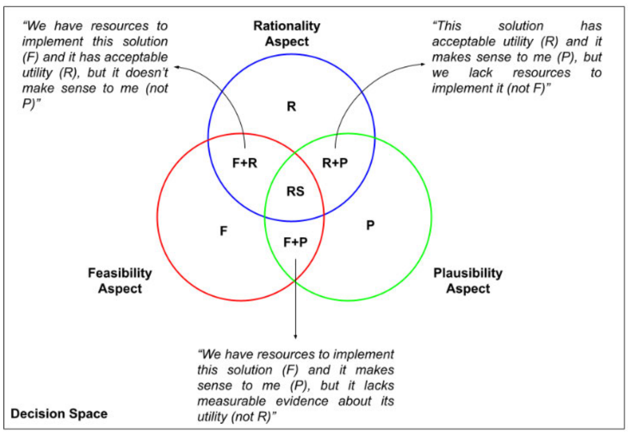


FIGURE 2. Venn diagram composed of the three constituent aspects of proposed reasonability (RS), and examples of real-world interpretations that could emerge given by intersections.

Here, it is proposed that a reasonable decision-making process must consider all three aspects in a balanced manner. Subsection III-B provides details on how a reasonable decision-making approach can be implemented computationally.

B. OVERVIEW OF THE COMPUTATION OF REASONABLE DECISION-MAKING APPROACH

To balance the three conceptual reasonability aspects, it is proposed that the process of finding adequate reasonable solutions be conducted as a constrained multi-objective optimization process [33]. In the context of the proposed approach, this optimization process aims to find non-dominated solutions when considering the feasibility, rationality, and plausibility aspects (all as independent optimization objectives) while dealing with any hard constraints imposed by the decision problem. Despite many similarities between search and optimization processes, the choice for the latter is motivated by the following reasons: (i) the evaluation of each aspect can be performed by a continuous number instead of crisp pertinency (e.g. is rational or is not rational), and (ii) this allows rankings of solutions with finer granularity instead of simply attributing solutions as being contained/not in each of the three reasonability proposed aspects. Figure 3 presents an overview of the proposed computations for a reasonable decision-making approach.

As can be seen in the three steps in Figure 3, after the collection of any relevant Decision Inputs, the Internal DSS Processing takes place to ultimately produce satisfactory (Reasonable) decisions. The three conceived steps are: Step-1, the Inference Graph (IG) extracts the Decision Inputs, producing multiple and potentially diverse candidate reasonable solutions to the problem; Step-2: each of these candidate solutions can be used to assemble a Reasonable Decision Structure (RDS) that might be useful for two purposes: (i) calculating each of the three aspect scores (i.e. feasibility, rationality, and plausibility scores) and (ii) being the raw material to generate explanations about the decision process; Step-3: a reasonability Pareto containing non-dominated Reasonable Decision Structures is produced, with respect to each aspect score, which is pre-selected to be presented for DM's evaluation and selection for implementation.

Because the approach here subscribes to the human-in-the-loop, a DM is asked to inspect each RDS contained in the pareto. The number is not likely to be high (owing to the performed optimization) and is accompanied by explanations. Both are helpful for selecting the most appropriate for actual implementation. If an appropriate solution is deemed satisfactory, the decision process ends. Otherwise, another processing cycle is initiated using the DM feedback for future improvement.

It is worth mentioning that the proposed decision approach is inspired by Simon's Bounded Rationality Decision Model (SBRDM) [34], which contributes to the Design Phase and aids in the Selection Phase. IG processing generates diverse candidate solutions to be instantiated further in the SBRDM Design Phase. The RDS Pareto construction, based on aspect scores, followed by DM evaluation, aided by explanations, is an improvement to the SBRDM's Choice Phase.

Subsections III-C and III-D delves into (i) the generation of reasonable solutions and (ii) evaluations of RDS.

C. GENERATION OF REASONABLE SOLUTIONS

To generate reasonable solution alternatives, two conceptual elements of this proposal must be instantiated: IG and RDS architectures. IG defines the inner DSS inference alternatives that are potentially composed of different layers and internal instances. Both are instrumental in generating reasonable solution alternatives, which are appropriate for each type of DSS and the problem being addressed. The RDS architecture detailed in the next subsection, defines the data and/or facts that must be collected during the processing of the IG so that all three reasonability aspects can be measured and used to construct the reasonability pareto.

The following steps are proposed to setup the IG and RDS architecture:

1) Define what is the DSS most relevant task and specify which processing stages must be performed during an inference. For example, for a DSS whose main task is to prioritize entities to be inspected based on predictions about these entities, the following processing stages should be conducted: (i) collecting historical data and training predictor models, (ii) predicting the relevant events for the next day, and (iii) grouping these predictions into a priority set of entities.
2) Define for the specific problem what must be measured about a solution for each of the reasonability aspects suggested in this approach, namely, feasibility, rationality, and plausibility.
3) Define DSS explanation requirements. In addition, define which data and/or facts must be stored and how to obtain low-level explanations about each processing step.
4) Define how to quantify each explanation or fact and how they might be mapped into Aspect Scores. For complex problems with multiple low-level facts and/or explanations, it is valuable to use intermediate Key Performance Indicators (KPI), which are later combined into Aspect Scores.
5) Define the IG, considering different approaches to accomplish each processing step and how the RDS information will be collected.
6) Instantiate IG and RDS in distinct pipelines, receiving the Decision Inputs, and yield the assembled RDS to pareto formation and pre-selection, as described in Figure 3.

An abstract view of IG processing and the generation of Reasonable Decision Structures is shown in Figure 4.

The practical result of processing an Inference Graph is the generation of different and possibly diverse solutions. An Inference Engine instance might be any intelligent technique used to generate a part or complete solution. Such techniques can include, for example, a trained Artificial Neural Network [35], a parameterized Particle Swarm Optimization [36] instance, or even a combination of techniques. This is when the problem resolution requires it, as in the case of a compound DSS. It is worth emphasizing that no mathematical formulas were supplied here regarding the training of each Inference Engine instance, because this aspect is dependent on the problem being solved.

In Figure 4, after the input data are transformed in the first layer, their transformed forms are used as inputs for an Inference Engine instance. Inference Engines can also be dynamically created when appropriate for cases in which a previous instance is not adequate to be reused.

The processing result of each Inference Engine instance is then stored as part of the Reasonable Decision Solution set as a candidate solution to the problem.

This work proposes the use of a graph instead of fixed optimized pipelines for three main reasons: (i) the possibility of exploring different paths in the graph to generate diversity in the pool of candidate solutions, (ii) to maximize the number of points where relevant data or facts can be selected to generate explanations, and (iii) to allow the evolution of this graph, taking profit from the feedback of the DM. Regarding the latter, it is worth mentioning that over time, an IG might be optimized by omitting connections that tend to produce low-quality solutions as means to foster system scalability even in problems with high computational training cost, multiple layers in the graph, or a large number of solutions to be evaluated. IG evolution is also expected to occur in cases when the pool of candidate solutions is not sufficiently diverse, compromising the quality of the decision alternative generation. This kind of diversity-increasing evolution is also triggered by DM feedback.

D. INSPECTING AND EVALUATING REASONABLE DECISION STRUCTURES

After IG processing is performed, according to Figure 4, all evidence necessary to explain each RDS will be collected so that each solution alternative can be explained later, if needed. The RDS was assembled in a bottom-up manner, as shown in Figure 5. To interpret the explanation of a given RDS, the DM might want to inspect it top-down, first verifying the aspect scores, if needed, drilling down to the KPI level, and finally, drill down to the facts level.

As stated in the previous subsection, during the setup of the RDS architecture, the quantification of facts/explanations in KPIs and the aggregation of KPIs into aspect scores is problem dependent. Therefore, it must be defined as ad hoc. The idea of assembling a multilevel structure is aligned with the work of Dazeley et al. [16]. The existence of multiple levels in the RDS might imply explanations with different abstraction levels and allow for a broad overview at the top level and a low granularity view at the facts level.

It is worth mentioning that the facts level is where explanations generated by other stablished XAI techniques such as LIME [37] will be stored.

Figure 6 depicts an example of the inspection flow of a RDS, which is part of a hypothetical Pareto result of the proposed approach. The aspect scores level allows for a possible simple way to perceive how pre-selected solutions

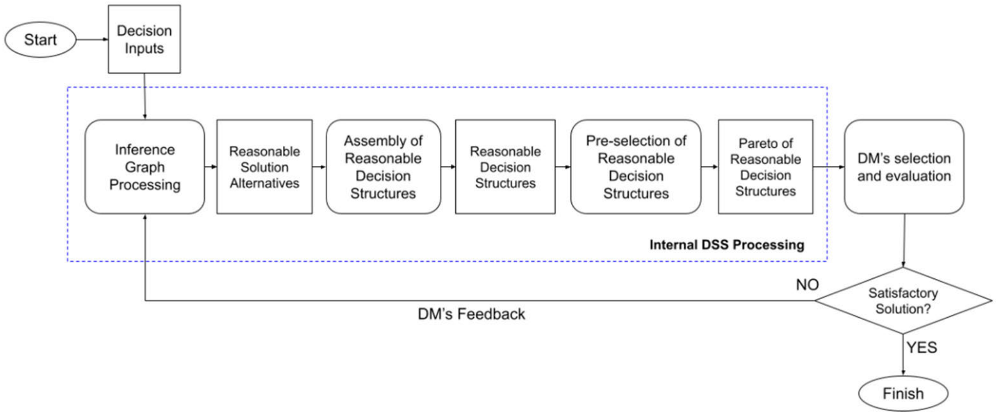


FIGURE 3. Overview of reasonable decision-making approach. Rectangles represent information used as input and/or produced in each step. Rounded rectangles represent system or human processing activities.

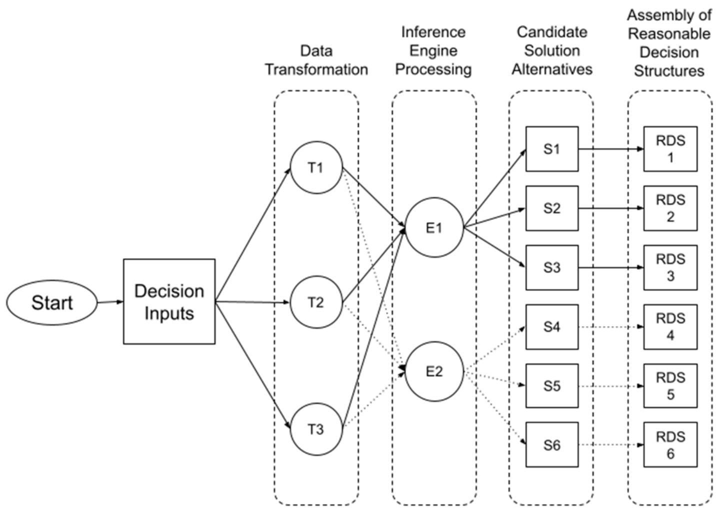


FIGURE 4. Example of processing over an abstract Inference Graph, leading to Reasonable Decision Structures. Solid arrows are related to Inference Engine 1 while dashed arrows are related to Inference Engine 2. Circles represent trained instances of data transformers or machine learning models. Squares represent candidate solutions or reasonable decision structures.

relate to each other. The inspection of the KPI level provides a contextualized overview of the inspected solution, possibly allowing a better comprehension of the properties of each specific decision with respect to each reasonability aspect. The fact level can also be inspected in cases where a lower-level view must be checked to ensure that the DM can trust a decision and/or will be able to adequately use it. For a graphical representation of the possibilities offered by the proposed approach, please refer to Figure 6, items 1–4.

It is worth emphasizing that the RDS is a data structure that might be inspected directly, for example, by AI experts improving the DSS or experienced DMs. However, the RDS may also be used to generate derived graphical or textual explanations geared toward simplifying its inspection and/or aiming at less experienced DMs. The next section describes two proofs of concept employing the proposed approach.

IV. EXPERIMENTS AND RESULTS

In this section, further emphasis is given to the application of the approach in model-based and compound DSS. The former is rather frequent in the literature, while the latter has been less studied by the XAI community. The two proofs of concept presented show not only the proposal's applicability but also its features for covering the three aspects of reasonability. All experiments were coded using the Python version 3.8 programming language and Scikit Learn API, version 1.0 [38], and on purpose, used a regular personal computer.

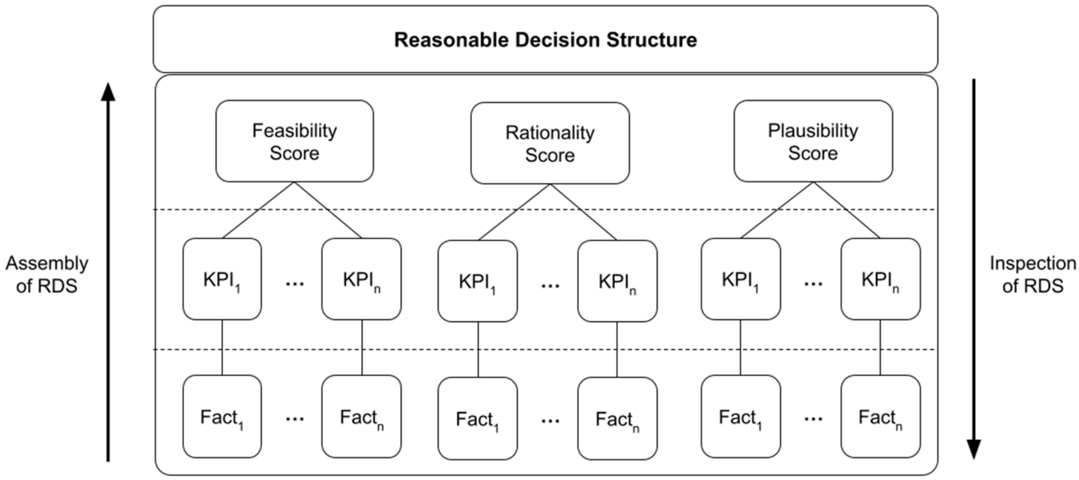


FIGURE 5. Example of an abstract Reasonable Decision Structure. From the bottom-up, each low-level explanation and/or fact is collected to assemble the RDS. From the top-down, the different levels of RDS and explanation paths can be inspected and interpreted to comprehend each reasonability aspect score.

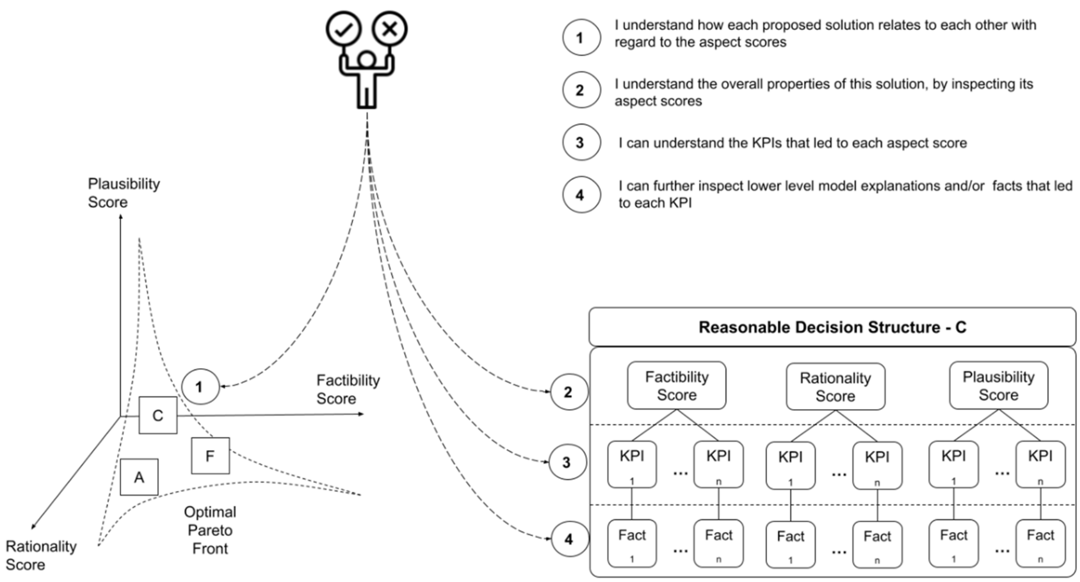


FIGURE 6. Example of DM interacting with pareto and RDS structure. Items 1 to 4 show different levels of inspection available to a DM when using the proposed approach.

A. PROOF OF CONCEPT 1 – MODEL-BASED DSS

A medical classification problem was selected as the first proof of concept. For this hypothetical scenario, a given medic intends to confirm the diagnosis of a patient with a heart disease. In a broader sense, this proof of concept is generalizable to the class of model-based DSS with the capability of automatic feature selection.

This problem is suitable for application of an AI powered DSS because it allows a data driven approach as well as it can also take profit of medic's expertise. In addition, the system was formulated to deal with time and budget restrictions, making it more realistic. Incorporating the benefits of the proposed approach allows for the delivery of reasonable decisions with comprehensible and justifiable explanations.

To implement the reasonable and explainable DSS used in this proof of concept, the following steps were conducted, as proposed in subsection III-C

1. The main objective of DSS is to confirm the diagnosis of a patient with possible heart disease. The main processing tasks consisted of a feature selection task followed by a classification task. The classification task consisted of classifying the patient as sick or not sick, considering the attributes selected for use. Information about the attributes used with cost and type considered during the experiments is shown in Table 1. The cost and type assigned to each

TABLE 1. Attributes used in proof of concept 1, along with cost and type.

Attribute	Cost	Type
age	0	t1
sex	0	t1
cp	1	t1
trestbps	1	t2
chol	1	t2
fbs	2	t2
restecg	2	t3
thalach	2	t3
exang	3	t3
oldpeak	3	t4
slope	3	t4
ca	2	t4
thal	2	t5

attribute are hypothetical and are used only to enrich the employed dataset for this proof of concept.

2. For this problem, the following characteristics were selected as relevant and presented as part of each reasonability aspect:
   a. With respect to the feasibility aspect, the diagnosis must be performed under a maximum amount of time and under a maximum cost, respecting possible constraints over both dimensions.
   b. With respect to the rationality aspect, the diagnosis must be performed with models containing the highest overall accuracy and with the highest accuracy among a certain number of most similar previous cases, combining global and specific performance.
   c. With respect to the plausibility aspect, the diagnostic must employ the maximum number of DM preferred attributes (i.e. exams) and the maximum number of DM-preferred types of attributes (i.e. categories of exams), allowing it to explore different strategies comprised in the problem-solving models.
3. The DSS is meant to provide explanations about complying with cost and time limits while considering: the selected exams, model accuracy, and DM preferences. For this matter, the following facts or low-level explanations were stored:
   a. List of used attributes and cost.
   b. List of used attributes – it was considered that each used attribute required the use of one time unit.
   c. List of patterns used in model testing, highlighting whether correct or incorrect classifications are received.
   d. List of most similar test patterns, highlighting if received correct or incorrect classification.
   e. List of preferred attributes used in the model.
   f. List of preferred types of attributes used in the model.
4. For this proof of concept, each reasonability aspect contained two KPIs as intermediate RDS levels. Each aspect score was quantified as the arithmetic mean of the corresponding aspect KPIs. The aspect KPIs were calculated as follows: The considered Rationality KPIs are R-KPI1 and R-KPI2, where S is the number of training samples, SS is the number of similar training samples, SC and SSC are the number of training samples correctly classified, and the number of training similar samples correctly classified, respectively. The Plausibility KPIs are P-KPI1 and P-KPI2 where PAS is the number of preferred attributes used for inference, PA is the number of DM preferred attributes, PTS is the number of preferred types of attributes used for inference, and PT is the number of DM preferred types of attributes. The considered Feasibility KPIs are F-KPI1 and F-KPI2 where TR is the ratio of the attributes used by the maximum number of attributes and CR is the ratio of the cost used over the maximum cost allowed. The abstract RDS is shown in Figure 7.

$$R-KPI1 = 100 * \frac{SC}{S} \tag{1}$$

$$R-KPI2 = 100 * \frac{SSC}{SS} \tag{2}$$

$$P-KPI1 = \min(100, 100 * \frac{PAS}{PA}) \tag{3}$$

$$P-KPI2 = \min(100, 100 * \frac{PTS}{PT}) \tag{4}$$

$$F-KPI1 = \begin{cases} 0, & for\ TR > 1.5 \\ 100, & for\ TR \leq 0.50 \\ 100 - 100 * (TR - 0.5), & c.c. \end{cases} \tag{5}$$

$$F-KPI2 = \begin{cases} 0, & for\ CR > 1.5 \\ 100, & for\ CR \leq 0.50 \\ 100 - 100 * (CR - 0.5), & c.c. \end{cases} \tag{6}$$

5. Considering the DSS processing task contained in Step 1, abstract IG was defined as the graph contained in Figure 8. For this implementation, in accordance with Step 1, IG contains two layers. The feature selection and classification layers contain instances obtained by an optimization meta-heuristic, the goal of which is to maximize the test accuracy. A sample of the feature selection masks and classifiers is presented in Table 2.
6. The abstracts RDS and IG were implemented and integrated into a pipeline, as shown in Figure 3.

The data used to train the IG classifier instances were obtained from the UCI Heart Disease dataset [39]. After removal of duplicates and patterns containing missing or null values, this dataset contained 297 patterns and 13 attributes and was adjusted to be a binary classification signaling the presence or absence of heart disease. Figure 9 depicts the pseudocode of the simulations performed using this proof of concept.

Two DM profiles were simulated considering the maximum amounts of time and cost and different preferences regarding attributes and types of attributes. The following user profiles were evaluated in the simulations: User profile 1

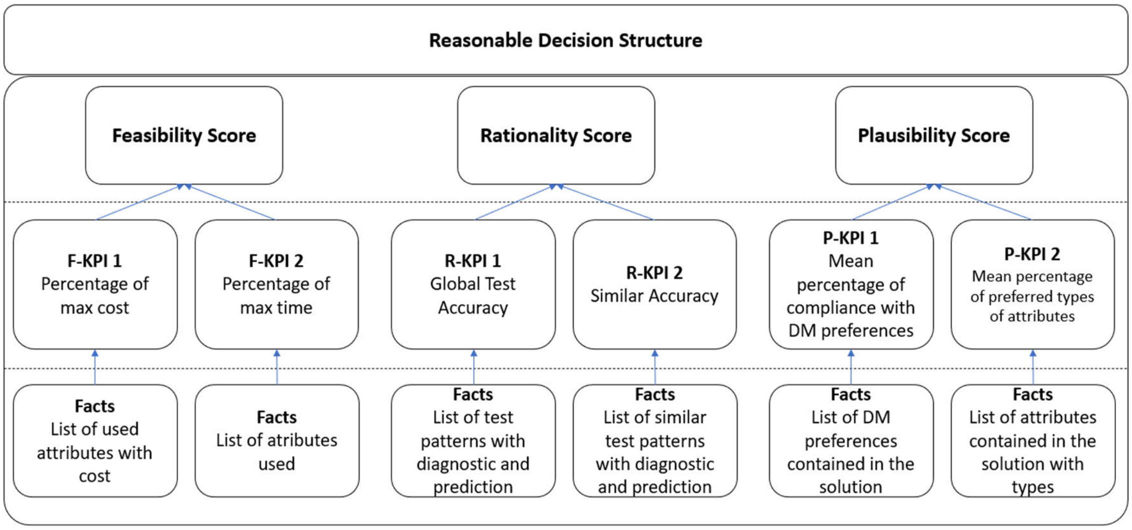


FIGURE 7. Abstract RDS build considering the model based DSS explanatory requirements.

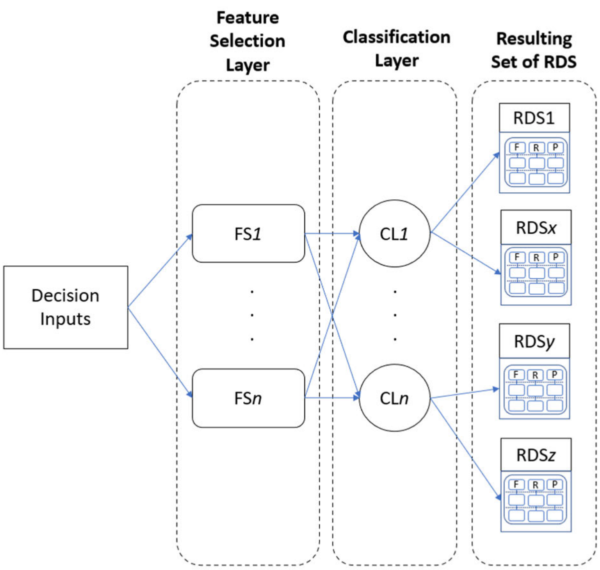


FIGURE 8. Abstract view of IG composed of layers for feature selection and classification as required by the model based DSS.

TABLE 2. Excerpt of feature selection and classification instances.

Path ID	Feature Selection Masks	Model Name	Test Accuracy (%)
0	0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1	MLP Classifier	85.00%
4	1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 1	Logistic Regression	91.66%
17	0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1	Logistic Regression	93.33%
19	0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1	Logistic Regression	91.66%

preferred to use some of the cheapest attributes, emulating a medic with restricted access to diagnostic resources. User profile 2 preferred using some of the attributes with a higher impact on the classification, emulating a more experienced medic.

- User Profile 1:
  - Preferred attributes: ‘age’, ‘sex’, ‘cp’, ‘trestbps’, ‘chol’;
  - Preferred types of exams: t1 and t4.
- User Profile 2:
  - Preferred attributes: ‘sex’, ‘restecg’, ‘thalach’, ‘slope’, ‘ca’;
  - Preferred types of exams: t2 and t3.

For this proof of concept, the following items will be analyzed: (i) how was the general behavior regarding each aspect score, and (ii) how could this approach comprising reasonability and explanations aid in this kind of decision?

The Proposed Approach was compared with two other DSS approaches to evaluate the general behavior of each aspect score. The first, which is referred to as the Random Approach, was performed by randomly selecting a model contained in the Model Database, not considering the aspect scores. This approach was used as a sanity check to evaluate whether an extremely simple strategy could adequately solve this problem. The second, mentioned as the Rational Approach, was performed by selecting predictions with a higher Rationality Score. A Rational Approach was included to evaluate whether a strict decision utility strategy could adequately solve this problem. It can be seen in Table 3 that the results of the experiments run over the first proof of concept with all three approaches.

It is possible to observe that for both user profiles simulated, the Random Approach achieved inferior results when compared with other approaches, except for the Plausibility of User Profile 2. On the other hand, as expected, the Rational Approach achieved the highest results for the Rationality

TABLE 3. Aspect scores (mean and standard deviation) for proof of concept 1, concerning each user profile with different approaches. Bold values are the highest among all three approaches, for each aspect score.

User profile	Aspect Score	Random Approach	Rational Approach	Proposed Approach
User profile 1 (i.e. least resources)	Rationality	88.4887 +/- 8.5046	**94.2232 +/- 4.2874**	92.6083 +/- 5.9304
	Feasibility	82.0621 +/- 23.3876	93.6332 +/- 15.1146	**97.8488 +/- 2.6279**
	Plausibility	80.5141 +/- 6.7104	81.8588 +/- 5.6155	**84.0621 +/- 4.5549**
User profile 2 (i.e. experienced medic)	Rationality	87.0904 +/- 9.6665	**94.2232 +/- 4.2874**	93.5640 +/- 4.6611
	Feasibility	83.7462 +/- 21.6811	93.6332 +/- 15.1146	**94.2742 +/- 5.7600**
	Plausibility	65.0508 +/- 7.4971	61.5989 +/- 6.1903	**66.2994 +/- 7.2607**

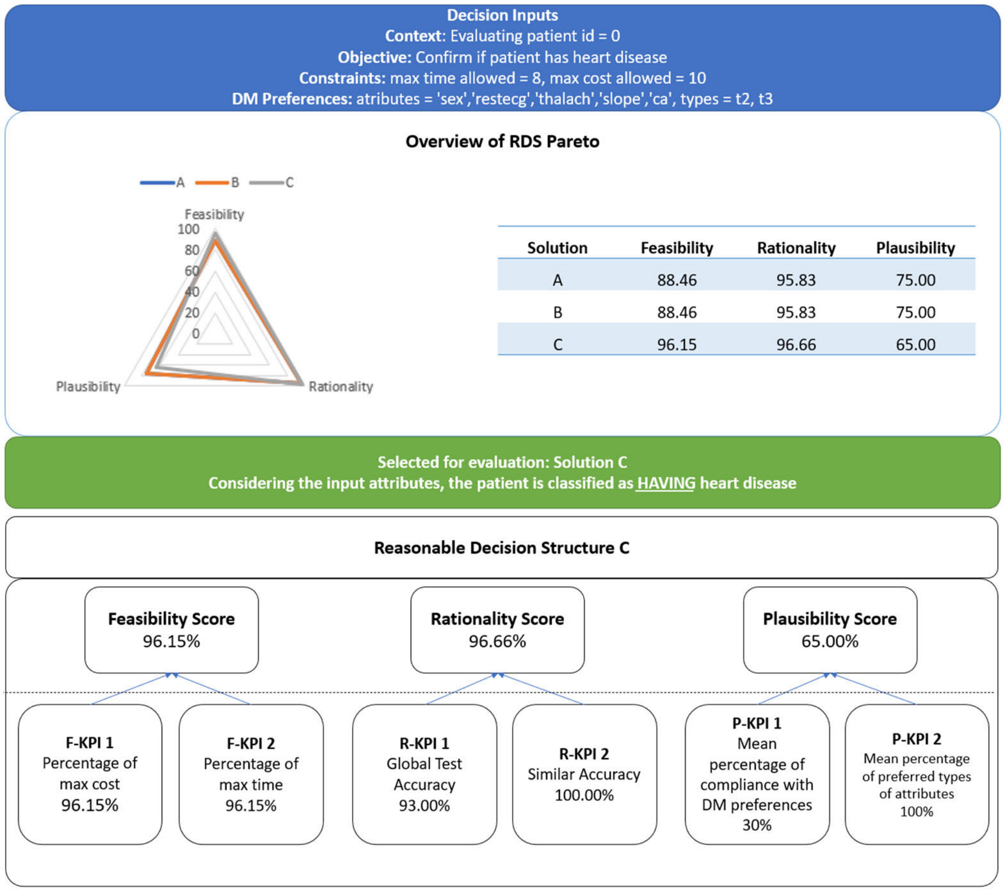


FIGURE 9. Overview of usage scenario, comprising (upper part) the three selected RDS to be inspected and (lower part) two levels of RDS created according to solution C.

```
IG = Inference Graph
UP = User Profile
DI = Decision Inputs
for feature_selector in IG.feature_selectors_list:
    selected_features = feature_selector.select(DI)
    for classifier in IG.classifiers_list:
        solution = classifier.predict(selected_features)
        rds = create_rds(solution,UP)
        rds_set.append(rds)
rds_pareto = extract_reasonability_pareto(rds_set)
```

FIGURE 10. Pseudocode of simulations performed over Proof of Concept 1.

Aspect, being above 94.00% for both users. Finally, the Proposed Approach achieved the best results for Feasibility and Plausibility, and second-best values for rationality. These results are strongly related to the assembly of the Database Model, which optimizes features to obtain the best accuracy. In this specific case, the selected features were more aligned with the preferences of User Profile 1.

To evaluate how this approach, comprising reasonability and explanations, could aid in this kind of decision, the description of one case is provided next, based on Figure 10. The decision inputs for this scenario are displayed at the top of the figure. The three selected solutions to be inspected by the DM are presented below. As shown in the radar chart and table, solutions A and B are equal with respect to all three aspects. Solution C is better when considering Rationality and Feasibility, but inferior when considering plausibility.

Considering that Solution C was selected to be inspected, follows a closer look at its corresponding RDS – it is worth mentioning that patient ID 0 had a classification of having a heart disease.

The DM could ask, 'Why is the Feasibility Score 96.15%?' and, in this case, both Feasibility KPIs had the same value. The answer to the question 'Why F-KPI1 is 96.15%?' is that the maximum cost allowed is 10, and this solution requires a total of 10 cost units. If the DM inspected the Rationality KPIs, it could check the global and similar accuracy of the model, and if required, it could check what patterns were correctly and incorrectly classified during model evaluation. Finally, when evaluating the Plausibility Score, the lowest value among all three solutions, it can be seen that Plausibility KPI 1 is only 30%. The answer to the question 'Why P-KPI1 is 30%?' would be a list of the preferred attributes used by the model – only 'sex' and 'ca' from the list of 5 preferred attributes. If the user decides not to use this solution, because User Profile 2 emulates a more experienced user, he/she would avoid an error because Patient ID 0 is in fact not sick.

B. PROOF OF CONCEPT 2 – COMPOUND DSS

For the second proof of concept, a real-world problem and a dataset were selected. This problem is faced by the Brazilian Federal Highway Police and is related to the creation of optimized patrolling routes. As will be presented later, the proposed solution concerns compound DSS. Nevertheless, this proof of concept is generalizable to the class of DSS composed of classifiers, followed by combinatorial optimizers.

This problem requires the use of an AI-powered DSS because it deals with data that change over time, and whose solution is not trivial. Since each police precinct must patrol a large length of roads, it is not easy to make sense of all historical data and details to come up with an optimized route for the whole day, and not all officers create patrolling routes in the same way. Some lean more on their past experiences and knowledge of road interest points, while others present a more analytical posture, using more intensively historical accident information. It would be desirable to use historical data primarily as well as to profit from the DM's tacit knowledge.

To implement the reasonable and explainable DSS used in this proof of concept, the following steps were conducted, as proposed in subsection III.C:

1. The main task of this DSS is to suggest reasonable patrolling routes for a given operational unit, in this case, a local precinct. The main processing task can be decomposed into a classification task, followed by a combinatorial optimization task. The classification task consists of predicting the occurrence of road accidents per road segment of 10 km, considering every three-hour time frame of a given day. The options for tree-hour windows and 10 km segments were selected after exploratory studies and were in accordance with the current PRF practices. After the predictions for each km and each time frame, combinatorial optimization takes place, selecting reasonable routes for the latter inspection of the police officer acting as the DM.
2. For this problem, the following characteristics were selected as relevant and presented as part of each reasonability aspect:
   a. For feasibility aspect: respect the maximum number of kilometers required by the route.
   b. For rationality aspect: maximize the number of probable accidents covered according to the prediction.
   c. For the plausibility aspect: the route must contain the maximum number of preferred KM segments proposed by the DM, allowing it to incorporate tacit knowledge.
3. The DSS is meant to provide explanations about respecting the maximum KM limit, the selected KM to be patrolled, and how much the proposed route is aligned with the preferences of the DM police officer. For this matter, the following facts or low-level explanations must be stored:
   a. List of km to be patrolled.
   b. List of km selected or not to be patrolled with predicted number of accidents.
   c. List of km matching DM preferences.
4. For this proof of concept, each reasonability aspect contained only one KPI as an intermediate RDS level. Each aspect score was quantified directly from the corresponding aspect KPI, as explained below. The F-KPI, R-KPI, and P-KPI are related to feasibility, rationality, and plausibility aspects, respectively. In Equation 7, TKM is the total km in the route and MKM is the maximum allowed km. In Equation 8, PAR is the total number of predicted accidents in the route, and TAP is the total number of predicted accidents. Finally, in Equation 9, NPR is the number of DM preferred km in the route, whereas NP is the number of DM preferred km. Equation 10 describes the calculation of the Feasibility Score, while Rationality and Plausibility scores are equal to their respective KPIs. The abstract RDS is shown in Figure 11.B.

$$F-KPI = 100 * \frac{TKM}{MKM} \tag{7}$$

$$R-KPI = 100 * \frac{PAR}{TAP} \tag{8}$$

$$P-KPI = 100 * \frac{NPR}{NP} \tag{9}$$

$$F-score = \begin{cases} 0, & for\ F-KPI > 150 \\ & or\ F-KPI < 50 \\ 100, & for\ F-KPI \geq 50 \\ & or\ F-KPI \leq 100 \\ 200-F-KPI, & c.c. \end{cases} \tag{10}$$

5. Considering the description of the DSS processing task contained in step 1, the abstract IG was defined as the graph contained in Figure 11.A. For this specific implementation, in accordance with step 1, the IG contains two layers: one for the classification task and the other

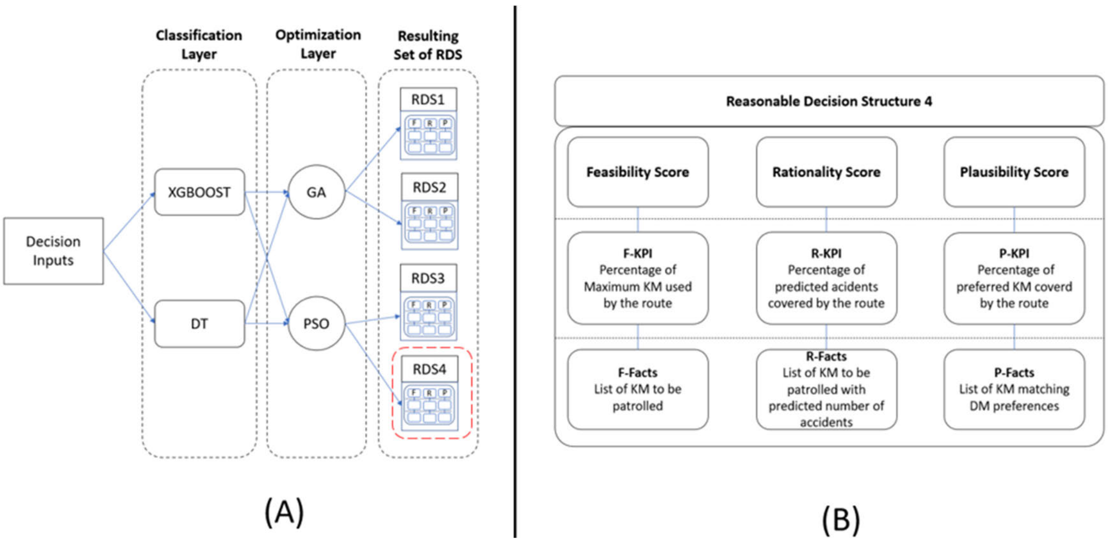


FIGURE 11. **(A) IG composed of layers for classification and for combinatorial optimization. (B) Abstract RDS build considering the DSS explanatory requirements, matching the highlighted RDS in (A).**

for the combinatorial optimization task. The classification layer contains optimized instances of an XGBoost [40] and a Decision Tree [41]. The optimization layer contains instances of a parameterized Genetic Algorithm [35] and combinatorial PSO [42].

6. The abstracts RDS and IG were implemented and integrated into a pipeline, as shown in Figure 3.

The original dataset contained data on 354,192 accidents collected from 2017 to 2021. After exploratory analysis of all available data, the following information was selected for use with classifiers: date of prediction, time of prediction, and km range to predict. For this experiment, data on BR101 from km 0 to 211 were selected because of its wide range, encompassing rural and city stretches, and having diverse landscape characteristics and road conditions. The number of samples containing accidents was 395, and that with no accidents was 1504. We used SMOTE [43] to deal with class imbalance, synthetically generating 1109 more samples containing accidents to better train classifiers, for a total of 3008 samples used in the experiments.

As described in step 1, the first part of the inference graph comprises a classification task. Different techniques were evaluated to select model instances to add to the Classification Layer. Logistic Regression, K Nearest Neighbors, DT, Artificial Neural Networks, Support Vector machines, Random Forests and XGBoost were evaluated and compared. The best-performing instances of classifiers were selected, in this case XGBoost and DT were obtained after optimizing the hyperparameters using a grid search over the training set. For this experiment, 80% of the data were used as the training set, and 20% were used as the test set to evaluate the generalization performance of the classifiers. A whole week was used as simulated decision problems, evaluating individually from Monday to Sunday.

```
IG = Inference Graph
UP = User Profile
DI = Decision Inputs
for classifier in IG.classifiers_list:
    for km_range in DI.km_range_list:
        km_predictions.append( [km_range, classifier.predict(km_range,x)] )
    for optimizer in IG.optimizers_list:
        solution = optimizer.optimize(DI,km_predictions,UP)
        rds = create_rds(solution,UP)
        rds_set.append(rds)
rds_pareto = extract_reasonability_pareto(rds_set)
```

FIGURE 12. **Pseudocode of simulations performed over Proof of Concept 2.**

Figure 12 depicts the pseudocode of the simulations performed over this proof of concept.

The following three user profiles were evaluated in the simulations, each of which explored different human strategies for generating patrolling routes:

- User Profile 1: Focused preferred KM in areas with the most past accidents. The preferred km were the ranges from to 50-59, 60-69 e 70-79;
- User Profile 2: Spread the preferred KM to areas with intermediate historical values. The preferred KM were in the ranges 30-39, 70-79 e 180-189;
- User Profile 3: Did not consider the number of past accidents and focused on the second part of BR-101 (i.e., above KM 100), forcing more coverage of the patrols. The preferred km were the ranges 100-109, 160-169, 200-209.

For this proof of concept, the following items will be analyzed: (i) What was the general behavior regarding each aspect score? and (ii) how does this approach comprising reasonability and explanations aid in this kind of decision?

The Proposed Approach was compared with two other DSS approaches to evaluate the general behavior of each aspect score. The first, referred to as the Random Approach,

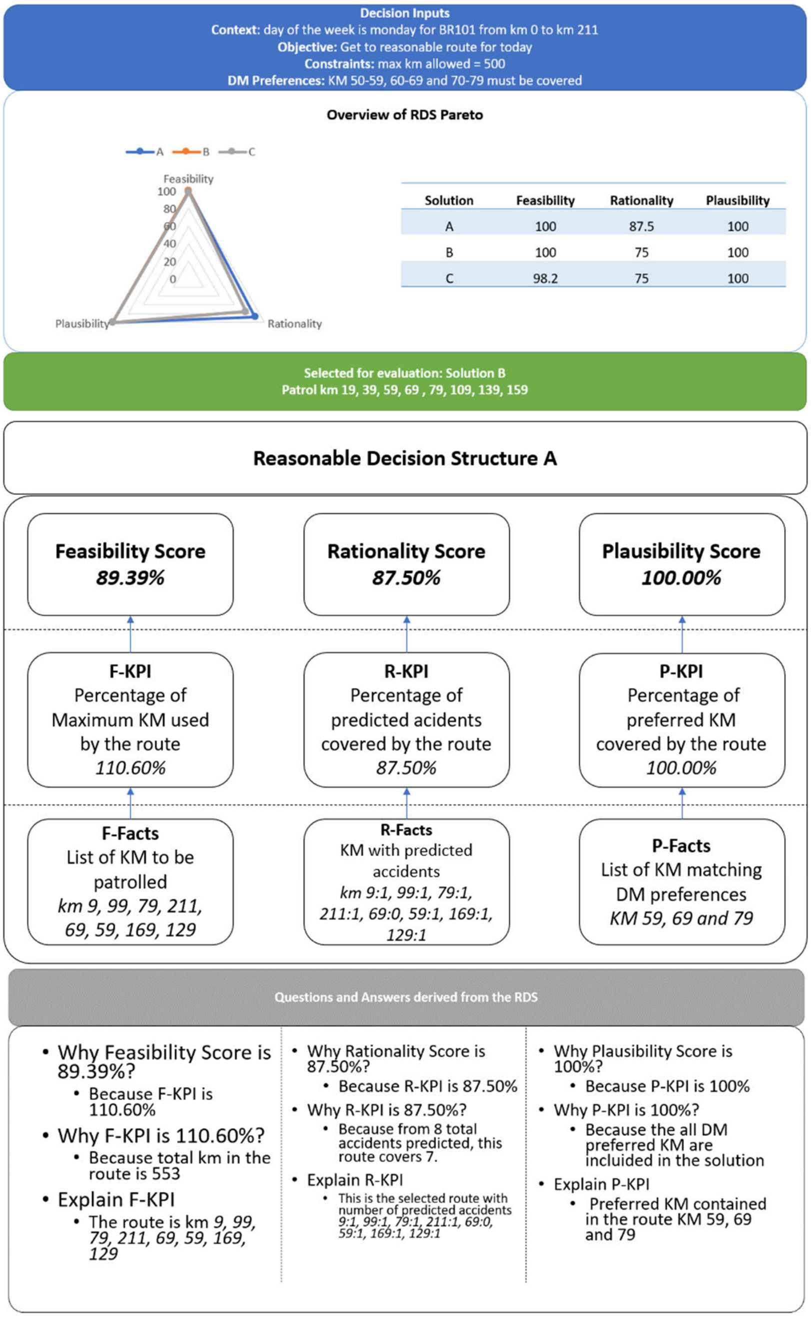


FIGURE 13. (A) Example of decision screen, containing all RDS selected for DM inspection. (B) RDS for solution A. (C) Questions and Answers extracted from RDS of solution A.

was performed by randomly selecting the accident predictions produced by XGBoost or DT and randomly selecting KM to patrol. The Random Approach was included as a dummy to evaluate whether an extremely simple strategy could adequately solve the problem. The second, mentioned as the Rational Approach, was performed by selecting predictions with higher accuracy, and after employing a greedy selection mechanism, striving to cover the maximum number of predicted accidents, despite other characteristics, such as respecting the maximum number of KM or preferred KM to inspect. A Rational Approach was included to evaluate whether a strict decision utility strategy could adequately solve this problem. It can be seen in Table 4 that the results of the experiments run over the first proof of concept with all three approaches.

It can be observed that for all three simulated user profiles, the Random Approach achieved results for each aspect, inferior to those of the Proposed Approach. In addition, in most cases, the aspect results were inferior to those of the Rational Approach, suggesting that the problem could not be solved

TABLE 4. Aspect scores (mean and standard deviation) for proof of concept 2, concerning each user profile with different approaches. Bold values are the highest among all three approaches, for each aspect score.

User profile	Aspect Score	Random Approach	Rational Approach	Proposed Approach
User profile 1	Rationality	53.6990 +/- 15.3867	**99.1071 +/- 2.1870**	74.6599 +/- 8.5838
	Feasibility	67.7286 +/- 19.3257	33.7286 +/- 43.6543	**100.0000 +/- 0.0000**
	Plausibility	33.3333 +/- 25.1976	83.3333 +/- 23.5702	**98.4127 +/- 3.8881**
User profile 2	Rationality	50.2551 +/- 13.8462	**99.1071 +/- 2.1870**	80.0170 +/- 9.8932
	Feasibility	75.0714 +/- 24.2164	33.7286 +/- 43.6543	**98.6000 +/- 1.8714**
	Plausibility	45.2381 +/- 19.3429	59.5238 +/- 12.1405	**96.8254 +/- 5.0195**
User Profile 3	Rationality	50.1276 +/- 13.6032	**99.1071 +/- 2.1870**	74.1497 +/- 7.2961
	Feasibility	65.7429 +/- 21.6609	33.7286 +/- 43.6543	**93.7333 +/- 3.8621**
	Plausibility	11.9048 +/- 11.6642	7.1429 +/- 8.2479	**95.2381 +/- 5.4986**

by chance or by a simple and not optimized heuristic. On the other hand, as expected, the Rational Approach achieved the highest results for the Rationality Aspect, above 99.00%. Finally, the Proposed Approach achieved the second-best results for rationality, between 74 and 80% among all user profiles. However, it achieved the best results for Feasibility and Plausibility aspects for all user profiles, being above 93%.

To evaluate how this approach, comprising reasonability and explanations, could aid in this kind of decision, the description of one case is provided next, based on Figure 13. The decision inputs for this scenario are displayed at the top of the figure. All three selected solutions to be inspected by the DM are presented, with close values for all three aspects, as can be seen in the radar chart and table. At this high level of abstraction, it is possible to perceive differences among all three solutions. Considering that Solution B was selected for inspection, it follows a closer look at its corresponding RDS. The multiple levels contained in the RDS allow for the inspection and understanding of the characteristics of the solution, going to the lowest fact level only if needed. The hierarchical RDS structure allows for the possible saving of cognitive resources, going to the lowest levels only if needed. In addition, it allows the identification of possible problems with a given solution regarding a specific aspect or level. For example, to understand why the Plausibility Score is 100%, it is possible to check its KPI level and see that all preferences were fulfilled. When in doubt, the Facts level display these preferences contained in the solution. Finally, there is a textual view of the RDS that can be generated using templating and presenting all questions and answers that can be extracted from this RDS. It is worth mentioning that a textual view of the RDS could contribute to a more human-centered inspection, allowing a broader audience to interact and profit from DSS explanations.

V. CONCLUSION

This work has proposed a method to obtain reasonable solutions with explanations by means of a decision approach that encompasses how to generate solutions that are rational, feasible, and plausible.

Moreover, this approach details how to obtain comprehensible and justifiable explanations. Two proofs of concept were presented regarding a medical problem, instantiated as a model-based DSS, and a public security problem, instantiated as a compound-oriented DSS.

Both proofs of concept explored how the approach can be used and their simulation results, suggesting that they could be employed even in real-world problems. Some of the KPIs employed in each proof of concept could be reused as evaluation indexes in other studies, considering their practical value. When compared to a purely rational approach, the reasonable approach put forward here delivered solutions that are not as good in terms of decision utility but are much more comprehensive in tackling problem restrictions and user preferences. In the first proof of concept, as shown in Table 3, for both simulated user profiles, the proposed approach achieved results up to 4% better for feasibility and 5% better for plausibility, while being worse by at most 2% for rationality score. In the second proof of concept, as shown in Table 4, the proposed approach achieved results that were more than 60% better for feasibility and at least 10% better for plausibility for all the simulated profiles. On the other hand, the results for rationality were 25% lower in the worst case. Even with the models optimized for accuracy, the results suggest that the solutions selected for DM inspection were adequate for each simulated user profile.

Despite the proofs of concept dealt with common types of Intelligent DSS, signaling the applicability of the approach, more experiments and instantiations are required to properly explore the boundaries of its application. Another concern may arise about scalability, because in real applications, the number of alternatives to be evaluated tends to be much larger than that used in the proof of concepts. Thus, new mechanisms should be integrated into this approach to address such challenges. One highlight of this approach is its flexibility in using various intelligent algorithms, such as optimizers, regressors, classifiers, or data transformers in the respective IG layers. Another concern might be the need for more results regarding the use of decision structures, such as those proposed here, and the need for more applications of the method. Because the method application and use of the decision structure are specific to each decision problem, two proofs of concept were provided as applications of the method, and instantiated decision structures were presented. Nevertheless, more studies are needed to derive the best practices for method and decision structure employment.

Future works are in the direction of investigating mechanisms to allow the dynamic optimization of IG connections, tackling more challenging personalization or dynamic scenarios, and developing new pre-hoc explainability algorithms to be incorporated in the proposal. Further studies on the use of the proposed decision structures will be conducted, along with more application scenarios. In addition to these investigation paths, the relation of how to better assemble the IG in terms of instance diversity is also to be investigated. Another improvement to be made on the road is to investigate more evaluation indexes than those used as KPI in both proofs of concept.

ACKNOWLEDGMENT

The authors would like to thank UPE and IFPE for their support.

REFERENCES

[1] Z. Nie, S. Jung, L. B. Kara, and K. S. Whitefoot, "Optimization of part consolidation for minimum production costs and time using additive manufacturing," *J. Mech. Des.*, vol. 142, no. 7, pp. 1–16, Jul. 2020, doi: 10.1115/1.4045106.

[2] A. Farrokhi, F. Shirazi, N. Hajli, and M. Tajvidi, "Using artificial intelligence to detect crisis related to events: Decision making in B2B by artificial intelligence," *Ind. Marketing Manage.*, vol. 91, pp. 257–273, Nov. 2020, doi: 10.1016/J.INDMARMAN.2020.09.015.

[3] A. A. Solanke, "Explainable digital forensics AI: Towards mitigating distrust in AI-based digital forensics analysis using interpretable models," *Forensic Sci. Int., Digit. Invest.*, vol. 42, Jul. 2022, Art. no. 301403, doi: 10.1016/j.fsidi.2022.301403.

[4] M. Amini, A. Bagheri, and D. Delen, "Discovering injury severity risk factors in automobile crashes: A hybrid explainable AI framework for decision support," *Rel. Eng. Syst. Saf.*, vol. 226, Oct. 2022, Art. no. 108720, doi: 10.1016/j.ress.2022.108720.

[5] F. Shaikh, J. Dehmeshki, S. Bisdas, D. Roettger-Dupont, O. Kubassova, M. Aziz, and O. Awan, "Artificial intelligence-based clinical decision support systems using advanced medical imaging and radiomics," *Current Problems Diagnostic Radiol.*, vol. 50, no. 2, pp. 262–267, Mar. 2021, doi: 10.1067/J.CPRADIOL.2020.05.006.

[6] T. H. Vo, N. T. K. Nguyen, Q. H. Kha, and N. Q. K. Le, "On the road to explainable AI in drug-drug interactions prediction: A systematic review," *Comput. Struct. Biotechnol. J.*, vol. 20, pp. 2112–2123, Apr. 2022, doi: 10.1016/j.csbj.2022.04.021.

[7] M. H. Jarrahi, "Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making," *Bus. Horizons*, vol. 61, no. 4, pp. 577–586, Jul. 2018, doi: 10.1016/j.bushor.2018.03.007.

[8] M. J. Aqel, O. A. Nakshabandi, and A. Adeniyi, "Decision support systems classification in industry," *Periodicals Eng. Natural Sci.*, vol. 7, no. 2, pp. 774–785, 2019, doi: 10.21533/pen.v7i2.550.

[9] J. Zerilli, U. Bhatt, and A. Weller, "How transparency modulates trust in artificial intelligence," *Patterns*, vol. 3, no. 4, Apr. 2022, Art. no. 100455, doi: 10.1016/j.patter.2022.100455.

[10] K. Stuurman and E. Lachaud, "Regulating AI. A label to complete the proposed act on artificial intelligence," *Comput. Law Secur. Rev.*, vol. 44, Apr. 2022, Art. no. 105657, doi: 10.1016/j.clsr.2022.105657.

[11] H. Felzmann, E. F. Villaronga, C. Lutz, and A. Tamò-Larrieux, "Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns," *Big Data Soc.*, vol. 6, no. 1, pp. 1–14, 2019, doi: 10.1177/2053951719860542.

[12] H. de Bruijn, M. Warnier, and M. Janssen, "The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making," *Government Inf. Quart.*, vol. 39, no. 2, Apr. 2022, Art. no. 101666, doi: 10.1016/j.giq.2021.101666.

[13] E. M. Kenny, C. Ford, M. Quinn, and M. T. Keane, "Explaining black-box classifiers using post-hoc explanations-by-example: The effect of explanations and error-rates in XAI user studies," *Artif. Intell.*, vol. 294, May 2021, Art. no. 103459, doi: 10.1016/j.artint.2021.103459.

[14] M. Langer and R. N. Landers, "The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers," *Comput. Hum. Behav.*, vol. 123, Oct. 2021, Art. no. 106878, doi: 10.1016/J.CHB.2021.106878.

[15] Q. V. Liao, D. Gruen, and S. Miller, "Questioning the AI: Informing design practices for explainable AI user experiences," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2020, pp. 1–15, doi: 10.1145/3313831.3376590.

[16] R. Dazeley, P. Vamplew, C. Foale, C. Young, S. Aryal, and F. Cruz, "Levels of explainable artificial intelligence for human-aligned conversational explanations," *Artif. Intell.*, vol. 299, Oct. 2021, Art. no. 103525, doi: 10.1016/j.artint.2021.103525.

[17] L. Weber, S. Lapuschkin, A. Binder, and W. Samek, "Beyond explaining: Opportunities and challenges of XAI-based model improvement," *Inf. Fusion*, vol. 92, pp. 154–176, Apr. 2023, doi: 10.1016/j.inffus.2022.11.013.

[18] F. N. Ford, "Decision support systems and expert systems: A comparison," *Inf. Manage.*, vol. 8, no. 1, pp. 21–26, 1985, doi: 10.1016/0378-7206(85)90066-7.

[19] R. H. Sprague and H. J. Watson, "A decision support system for banks," Omega, vol. 4, no. 6, pp. 657–671, 1976, doi: 10.1016/0305-0483(76)90093-1.

[20] H. J. Watson, "Revisiting ralph Sprague's framework for developing decision support systems," *Commun. Assoc. Inf. Syst.*, vol. 42, no. 1, pp. 363–385, 2018, doi: 10.17705/1CAIS.04213.

[21] M. S. Hasan, Z. Ebrahim, W. H. Wan Mahmood, and M. N. A. Rahman, "Decision support system classification and its application in manufacturing sector: A review," *Jurnal Teknologi*, vol. 79, no. 1, pp. 153–163, Dec. 2016, doi: 10.11113/jt.v79.7689.

[22] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018, doi: 10.1109/ACCESS.2018.2870052.

[23] A. B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020, doi: 10.1016/J.INFFUS.2019.12.012.

[24] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics," *Electronics*, vol. 8, no. 8, pp. 1–34, 2019, doi: 10.3390/electronics8080832.

[25] J. Schneider and J. P. Handali, "Personalized explanation for machine learning: A conceptualization," in *Proc. 27th Eur. Conf. Inf. Syst. Inf. Syst. Sharing Soc., (ECIS)*, 2019, pp. 1–17.

[26] M. Dikmen and C. Burns, "The effects of domain knowledge on trust in explainable AI and task performance: A case of peer-to-peer lending," *Int. J. Hum.-Comput. Stud.*, vol. 162, Jun. 2022, Art. no. 102792, doi: 10.1016/j.ijhcs.2022.102792.

[27] K. Coussement and D. F. Benoit, "Interpretable data science for decision making," *Decis. Support Syst.*, vol. 150, Nov. 2021, Art. no. 113664, doi: 10.1016/j.dss.2021.113664.

[28] E. M. Kenny, E. Ruelle, A. Geoghegan, L. Shalloo, M. O'Leary, M. O'Donovan, and M. T. Keane, "Predicting grass growth for sustainable dairy farming: A CBR system using Bayesian case-exclusion and post-hoc, personalized explanation-by-example (XAI)," in *Case-Based Reasoning Research and Development*, vol. 11680. Cham, Switzerland: Springer, 2019, pp. 172–187, doi: 10.1007/978-3-030-29249-2_12.

[29] A. J. Barda, C. M. Horvat, and H. Hochheiser, "A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare," *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, pp. 1–16, Dec. 2020, doi: 10.1186/s12911-020-01276-x.

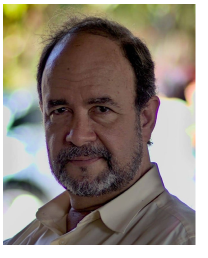
[30] D. Cirqueira, M. Helfert, and M. Bezbradica, "Towards design principles for user-centric explainable AI in fraud detection," in *Proc. Int. Conf. Hum.-Comput. Interact.*, vol. 12797, 2021, pp. 21–40, doi: 10.1007/978-3-030-77772-2_2.

[31] D. Gunning and D. W. Aha, "DARPA's explainable artificial intelligence program," *AI Mag.*, vol. 40, no. 2, pp. 44–58, 2019, doi: 10.1609/aimag.v40i2.2850.

[32] M. E. B. Brarda, L. H. Tamargo, and A. J. García, "An approach to enhance argument-based multi-criteria decision systems with conditional preferences and explainable answers," *Expert Syst. Appl.*, vol. 126, pp. 171–186, Jul. 2019, doi: 10.1016/j.eswa.2019.02.021.

[33] Y. Yang, J. Liu, and S. Tan, "A multi-objective evolutionary algorithm for steady-state constrained multi-objective optimization problems," *Appl. Soft Comput.*, vol. 101, Mar. 2021, Art. no. 107042, doi: 10.1016/j.asoc.2020.107042.
[34] H. A. Simon, "Theories of bounded rationality," in *Decision and Organization*. Amsterdam, The Netherlands: North Holland, 1972, ch. 8, pp. 161–176.
[35] L. H. S. de Menezes, L. L. Carneiro, I. M. de Carvalho Tavares, P. H. Santos, T. P. das Chagas, A. A. Mendes, E. G. P. da Silva, M. Franco, and J. R. de Oliveira, "Artificial neural network hybridized with a genetic algorithm for optimization of lipase production from penicillium roqueforti ATCC 10110 in solid-state fermentation," *Biocatalysis Agricult. Biotechnol.*, vol. 31, Jan. 2021, Art. no. 101885, doi: 10.1016/J.BCAB.2020.101885.
[36] J. Lu, J. Zhang, and J. Sheng, "Enhanced multi-swarm cooperative particle swarm optimizer," *Swarm Evol. Comput.*, vol. 69, Mar. 2022, Art. no. 100989, doi: 10.1016/J.SWEVO.2021.100989.
[37] X. Li, H. Xiong, X. Li, X. Zhang, J. Liu, H. Jiang, Z. Chen, and D. Dou, "G-LIME: Statistical learning for local interpretations of deep neural networks using global priors," *Artif. Intell.*, vol. 314, Jan. 2023, Art. no. 103823, doi: 10.1016/j.artint.2022.103823.
[38] F. Pedregosa, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Jan. 2011.
[39] D. Dua and C. Graff. (2017). *UCI Machine Learning Repository*. [Online]. Available: http://archive.ics.uci.edu/ml
[40] J. Li, X. An, Q. Li, C. Wang, H. Yu, X. Zhou, and Y.-A. Geng, "Application of XGBoost algorithm in the optimization of pollutant concentration," *Atmos. Res.*, vol. 276, Oct. 2022, Art. no. 106238, doi: 10.1016/j.atmosres.2022.106238.
[41] E. Dolotov and N. Zolotykh, "Evolutionary algorithms for constructing an ensemble of decision trees," in *Proc. Int. Conf. Anal. Images, Social Netw. Texts*, W. M. P. van der Aalst, V. Batagelj, D. I. Ignatov, M. Khachay, V. Kuskova, A. Kutuzov, S. O. Kuznetsov, I. A. Lomazova, N. Loukachevitch, A. Napoli, P. M. Pardalos, M. Pelillo, A. V Savchenko, and E. Tutubalina, Eds. Cham, Switzerland: Springer, 2020, pp. 9–15.
[42] K. Ihara, S. Kato, T. Nakaya, T. Ogi, and H. Masuda, "Application of PSO-based constrained combinatorial optimization to segment assignment in shield tunneling," in *Agents and Artificial Intelligence*, J. van den Herik, A. P. Rocha, and L. Steels, Eds. Cham, Switzerland: Springer, 2019, pp. 166–182.
[43] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Dec. 2002, doi: 10.1613/jair.953.

FLÁVIO ROSENDO DA SILVA OLIVEIRA received the B.Sc. and M.Sc. degrees in computer engineering from the University of Pernambuco, in 2007 and 2009, respectively, where he is currently pursuing the Ph.D. degree. He is also an Assistant Professor with the Federal Institute of Education, Science and Technology, Paulista, Pernambuco. He has authored several research papers published in well reputed peer-reviewed conferences. His research interests include information systems, machine learning, deep learning, and explainable artificial intelligence.



FERNANDO BUARQUE DE LIMA NETO (Senior Member, IEEE) received the Diploma degree in artificial neural networks from the Imperial College London, in 2002, and the Ph.D. degree in artificial intelligence from the University of London. He is currently an Associate Professor with the University of Pernambuco, Brazil, where he leads the Computational Intelligence Research Group. His current research interests include computational intelligence (metaheuristics evolutionary, social, and hybrid), modeling/simulation of stochastic and complexes problems, and intelligent decision support systems (including aspects of computational semiotics). He is also a Humboldt Fellow. He is also a member of the Computational Engineering Postgraduate Program.