Insights into the Potential of Fuzzy Systems for Medical AI Interpretability

Hafsaa Ouifak[1][a] and Ali Idri[1,2][b]
[1]Faculty of Medical Sciences, Mohammed VI Polytechnic University, Ben Guerir, Morocco
[2]Software Project Management Research Team, ENSIAS, Mohammed V University, Rabat, Morocco

Keywords: Explainable AI, Interpretability, Black-Box, Machine Learning, Fuzzy Logic, Neuro-Fuzzy, Medicine.

Abstract: Machine Learning (ML) solutions have demonstrated significant improvements across various domains. However, the complete integration of ML solutions into critical fields such as medicine is facing one main challenge: interpretability. This study conducts a systematic mapping to investigate primary research focused on the application of fuzzy logic (FL) in enhancing the interpretability of ML black-box models in medical contexts. The mapping covers the period from 1994 to January 2024, resulting in 67 relevant publications from multiple digital libraries. The findings indicate that 60% of selected studies proposed new FL-based interpretability techniques, while 40% of them evaluated existing techniques. Breast cancer emerged as the most frequently studied disease using FL interpretability methods. Additionally, TSK neuro-fuzzy systems were identified as the most employed systems for enhancing interpretability. Future research should aim to address existing limitations, including the challenge of maintaining interpretability in ensemble methods

1 INTRODUCTION

With the emergence of social networks and the digital transformation of most of the aspects of our lives, data has become abundant (Yang et al., 2017). Based on this data, Machine Learning (ML) techniques can provide decision-makers with future insights and help them make informed decisions. ML techniques are now being used in various fields given engineering (Thai, 2022), industry (Bendaouia et al., 2024), medicine (Zizaan and Idri, 2023), etc.

ML techniques can be divided into two classes: white-box and black-box models. White-box models, like decision trees or linear classifiers, are transparent and easily interpretable, allowing for straightforward explanations of the knowledge they learn. On the other hand, black-box models, such as Support Vector Machines (SVMs), Random Forests, and Artificial Neural Networks (ANNs) (Loyola-Gonzalez, 2019), are not interpretable.

With the popularity of Deep Learning (DL), black box techniques have been extensively and successfully used: the more data these techniques are fed, the better their performance capabilities (Alom et al., 2019).

Despite their effectiveness, black box techniques lack an acceptable performance-interpretability tradeoff, and this represents a major obstacle to their acceptance in several domains where the cost of an error is very high and intolerable (Alom et al., 2019). For example, in the medical context, a “wrong” decision is likely to cost the life of a patient. Thus, interpretability in medicine can be used to argue the diagnosis or treatments given and makes the ML technique used trustworthy to physicians and patients.

Interpretability refers to how well humans can comprehend the reasons behind a decision made by a model (Christoph, 2020). The evaluation and assessment of interpretability techniques are challenging and sometimes left to subjectivity as it has no common interpretability measure.

A common technique to make black box techniques interpretable is to use fuzzy logic (FL). Works attempting to use FL to interpret ML black box models do so in two ways: 1) fuzzy rule extraction (Markowska-Kaczmar and Trelak, 2003), where FL is used to extract fuzzy rules explaining the behavior of the model; fuzzy rules are composed of linguistic variables that are more comprehensible to humans

[a] https://orcid.org/0000-0002-4611-6987
[b] https://orcid.org/0000-0002-4586-4158

Ouifak, H. and Idri, A.
Insights into the Potential of Fuzzy Systems for Medical AI Interpretability.
DOI: 10.5220/0013072900003838

In *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2024) - Volume 1: KDIR*, pages 525-532
ISBN: 978-989-758-716-0; ISSN: 2184-3228

(Zadeh, 1974). And 2) neuro-fuzzy systems which are used to add the interpretability aspect to ANNs while maintaining their learning and performance capabilities (de Campos Souza, 2020).

To the best of our knowledge, no Systematic Mapping Study (SMS) dealing with the use of FL in ML black box models' interpretability has been carried out for medical applications. However, there are some works related to this topic. For instance, Souza (de Campos Souza, 2020) reviewed the theory behind hybrid models, i.e., the models based on FL and ANNs, and concluded that such models present a certain degree of interpretability while maintaining a high level of performance. Similarly, Das and al. (Das et al., 2020) reviewed the improvements FL can bring to DNNs and the real-life applications of such models. Other recent studies have reviewed fuzzy interpretability to highlight its emerging trend and the promises of this field (Padrón-Tristán et al., 2021).

This study presents an SMS of the use of FL in ML interpretability for medical applications. We conducted a search on six digital libraries: IEEE Xplore, ScienceDirect, PubMed, ACM Digital Library, Wiley, and Google Scholar. The search was conducted in the period between 1994 and January 2014 and has identified 67 primary studies. The selected studies were analyzed according to four Mapping Questions (MQs):

- Publication channels and years of publications (MQ1).
- Type of presented contribution (MQ2).
- Identifying the studied medical diseases (MQ3).
- Discovering the FL categories and systems used the most by the selected papers (MQ4).

The structure of this paper is as follows: Section 2 provides an introduction to ML interpretability and FL. Section 3 outlines the research methodology used to carry out this SMS. Section 4 details the findings from the mapping study. Lastly, the conclusions are discussed in Section 5.

2 BACKGROUND

This section presents an overview of the concepts and techniques that will be referred to in this study.

2.1 Interpretability

Interpretability techniques (i.e., post-hoc or post-modeling interpretability techniques) are used to explain the behavior of certain ML models that are not intrinsically interpretable (i.e., black box) (Barredo Arrieta et al., 2020). These techniques can be classified based on their applicability and their scope. In terms of applicability, post-hoc interpretability techniques can be divided into two main groups: 1) model-agnostic methods which can be applied to any ML model (Barredo Arrieta et al., 2020). These methods work without accessing the model's internal architecture and are applied after the training (e.g. Fuzzy rule extraction (Markowska-Kaczmar and Trelak, 2003)). 2) Model-specific methods (Barredo Arrieta et al., 2020), on the other hand, rely on the internal structure of a particular model and can only explain that model (Carvalho et al., 2019) (e.g., feature relevance, visualization).

Another type of interpretability techniques classification can be done using the scope of the explanations they generate. 1) Global interpretability techniques which try to explain the whole behavior of a model; and 2) Local interpretability techniques which are only concerned with explaining the process that led the model to a particular decision (Doshi-Velez and Kim, 2017). Examples of global interpretability techniques are permuted feature importance (Fisher et al., 2018) and global surrogates (Christoph, 2020). Local interpretable model-agnostic explanations (LIME) (Barredo Arrieta et al., 2020) and SHapley Additive exPlanations (SHAP) (Lundberg et al., 2017) are two of the popular local interpretability techniques). Moreover, methods that combine a white-box and a black-box to achieve a tradeoff between performance and interpretability are referred to as hybrid architectures (e.g., neuro-fuzzy systems (Ouifak and Idri, 2023a)).

2.2 Fuzzy Inference Systems

Fuzzy inference systems (FIS) use a set of fuzzy rules to map inputs to outputs (Jang, 1993). There are two primary types of FIS: Mamdani and Takagi-Sugeno-Kang (TSK). The difference between these types occurs in the consequent part of their fuzzy rules (Zhang et al., 2020).

Mamdani FIS (Mamdani and Assilian, 1975): Developed by Mamdani for controlling a steam engine and boiler system, the Mamdani FIS follows four steps: 1) Fuzzifying the inputs, 2) Evaluating the rules (inference), 3) Aggregating the results of the rules, and 4) Defuzzifying the output. This type of FIS is often used in Linguistic Fuzzy Modeling (LFM) because of its interpretable and intuitive rule bases. For example, in a system with one input and one output, a Mamdani fuzzy rule might be structured as:

$$If\ x\ is\ A\ Then\ y\ is\ B \quad (1)$$

where x and y are linguistic variables, A and B are fuzzy sets.

TSK (Takagi-Sugeno-Kang) FIS (Sugeno and Kang, 1988): This type of FIS was introduced by Takagi, Sugeno, and Kang. It also uses fuzzy rules but differs in that the consequent part is a mathematical function of the input variables rather than a fuzzy set. For example, in a system with two inputs, a TSK fuzzy rule might be structured as:

$$If\ x\ is\ A\ and\ y\ is\ B\ then\ z\ is\ f(x,y) \qquad (2)$$

where x and y are linguistic variables, A and B are fuzzy sets, and $f(x,y)$ is a linear function.

3 METHODOLOGY

Kitchenham and Charters (Kitchenham and Charters, 2007) proposed a mapping and review process consisting of six steps as shown in Figure 1. The present mapping study follows their process.

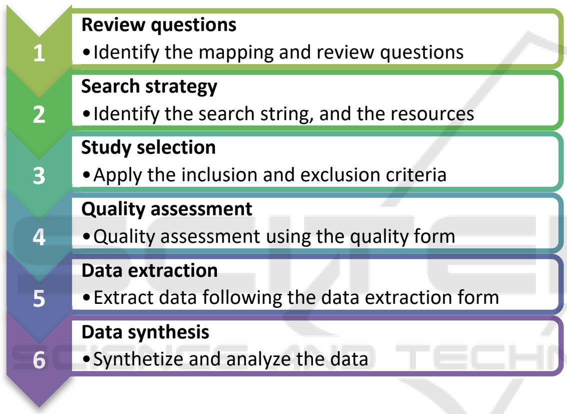

Figure 1: Mapping methodology steps (Kitchenham and Charters, 2007).

3.1 Mapping Questions

The purpose of this SMS is to select and organize research works focused on using fuzzy systems to interpret ML models for medical applications. The proposed MQs for this study are outlined in Table 1.

Table 1: Mapping questions of the study.

ID	Question	Motivation
MQ1	What are the publication channels and years of publications?	To determine if there is a dedicated publication channel and to identify the number of articles discussing the use of FL in enhancing the interpretability of ML black box models for medicine over the years
MQ2	What are the types of contributions presented in the literature?	To identify the different types of studies dealing with the use of FL for ML black box models' interpretability
MQ3	What are the most studied diseases?	To find out the diseases and the medical applications that were mostly studied using the fuzzy systems to make ML decisions interpretable
MQ4	What are is the type of fuzzy systems most evaluated?	To discover the FL technique category claimed to have a better chance of enhancing the interpretability of ML black box models

3.2 Search Strategy

To address the suggested MQs, we initially created a search string and then selected six digital libraries: IEEE Xplore, ScienceDirect, ACM Digital Library, PubMed, Wiley, and Google Scholar. These libraries were frequently used in previous reviews in the field of medicine (Ouifak and Idri, 2023b; Zizaan and Idri, 2023).

3.2.1 Search String

To ensure comprehensive coverage, the search string included key terms related to the study questions along with their synonyms. Synonyms were connected using the OR Boolean operator, while the main terms were linked with the AND Boolean operator. The full search string was constructed as follows:

("black box" OR "neural networks" OR "support vector machine" OR "random forest" OR "ensemble") AND (fuzz*) AND (interpretab* OR explainab* OR “rule extraction” AND (medic* OR health*).

3.2.2 Search Process

The search process of the present SMS was based on titles, abstracts, and keywords of the primary retreived studies indexed by the six digital libraries.

3.3 Study Selection

At this point, the searches carried out returned a set of candidate studies. To further filter the candidate studies, we used a set of ICs and ECs, described in Table 2, and evaluated each one of the candidate papers based on the titles and abstracts. In case no final decision can be made based on the abstract and/or title, the full paper was reviewed.

3.4 Quality Assessment

The quality assessment (QA) phase is used to further filter high-quality papers and limit the selection. To do this, we created a questionnaire with six questions aimed at evaluating the quality of the relevant papers, as shown in Table 3.

Table 2: Inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
Paper proposing/improving a new/existing FL-based ML interpretability technique for a medical application	Papers not written in English
Paper providing an overview of FL-based ML interpretability techniques	Unavailability of the full-text
Paper evaluating/comparing FL-based ML interpretability techniques of ML black box models	Paper using FL for any purpose other than increasing the interpretability of ML black box models
	Paper attempting to improve the interpretability of ML black box models without the use of FL

Table 3: Quality assessment form.

	Question	Possible answers
QA1	Is the FL-based ML interpretability method presented in detail?	“Yes”, “Partially” or “No”
QA2	Does the study evaluate the performance of the proposed FL-based ML interpretability technique?	“Yes”, “Partially” or “No”
QA3	Was the assessment done quantitatively or qualitatively?	“Quantitatively” or “Qualitatively”
QA4	Does the study compare the proposed technique with other techniques?	“Yes” or “No”
QA5	Does the study discuss the benefits and limitations of the proposed technique?	“Yes”, “Partially” or “No”
QA6	Is the Journal/Conference recognized?	**Conferences**: Core A: +1.5 Core B: +1 Core C: +0.5 Not ranked: +0 **Journals** : Q1 : +2 Q2 : +1.5 Q3 or Q4: +1 Not ranked: +0

3.5 Data Extraction

A data extraction form was utilized for each selected paper to answer the MQs. The extraction process was divided into two phases: initially, the first author reviewed the full texts of the studies to collect relevant data, followed by a verification step where the co-author ensured the accuracy of the extracted information.

3.6 Data Synthesis

During the data synthesis stage, the extracted data is consolidated and reported for each MQ. To simplify this process, we used the vote-counting method, and narrative synthesis to interpret the results. Then, visualization tools such as bar and pie charts, created using MS Excel were used for a better presentation.

3.7 Threats to Validity

Highlighting the study's limitations is as important as presenting its findings, enhancing reliability. Some main threats to validity in this study can be:

Study selection bias: A search string using the search string may miss some studies due to the broad scope. To address this, we set minimum criteria in the QA for objective decisions and included three possible answers to minimize disagreement (“Yes”, “Partially” and “No”).

To ensure accuracy during the data extraction phase, the results were reviewed consecutively by two authors.

4 MAPPING RESULTS

This section gives a summary of the selected articles, addresses the MQs listed in Table 1, and discusses the results of the synthesis.

4.1 Selection Process

The searches across the six selected digital libraries returned a total of 2,561 potential articles. By applying IC/EC and performing a quality assessment, we identified the papers relevant to our SMS, resulting in 67 pertinent studies, as depicted in Figure 2.

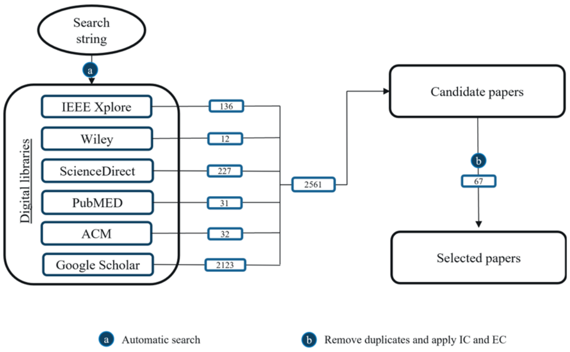

Figure 2: Papers selection steps.

4.2 MQ1: Publication Channels and Years

The 67 selected studies were distributed across journals and conferences, as depicted in Figure 3. Specifically, 67% of these papers were published in journals, and 33% in conference proceedings.

The selected papers were published in the journals IEEE Transactions on Fuzzy Systems, Expert Systems with Applications, and Applied Soft Computing, each featuring six publications. The International Conference on Fuzzy Systems (FUZZ-IEEE) was the most common conference, appearing three times among the selected papers, whereas other conferences were cited only once or twice.

The bar chart in Figure 3 shows the distribution of papers published each year from 1999 to 2023. There are several years with low numbers of publications, mostly between 2 to 4 papers, 1999 (4 papers), 2005 (3 papers), and 2006 (3 papers). A significant increase is observed starting in 2020, with 6 papers, followed by 14 papers in 2021, and peaking at 15 papers in 2022. In 2023, the number of publications decreased to 4.

The observed increase in studies focusing on the interpretability of ML black-box models using FL in 2022 may be related to the increased interest in transparency and trustworthiness in ML models. The necessity for explainable AI (XAI) has become particularly pressing in critical domains such as medicine (Chaddad et al., 2023). Consequently, researchers have been exploring various XAI approaches, with fuzzy systems being one notable avenue of investigation.

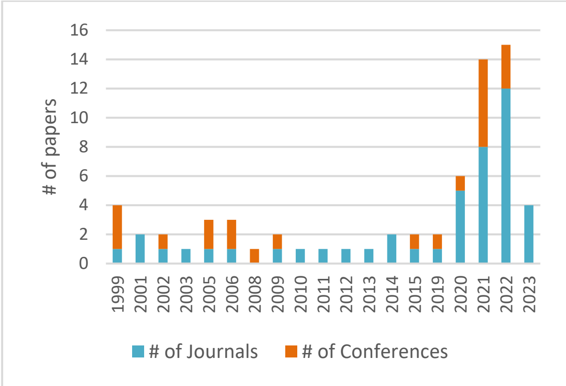

Figure 3: Distribution of the qualified studies per year and channels.

The decrease in the number of papers in 2023 can be attributed to several challenges, such as the complexity involved in training neuro-fuzzy systems for high-dimensional datasets (Ouifak and Idri, 2023a). As the rule bases expand, the rules themselves can become lengthy and difficult to interpret (Ouifak and Idri, 2023b, 2023a). Another factor may be the transparency these models offer when dealing with tabular data, where linguistic rules are more easily understood. However, many ML applications in medicine are related to medical imagery, where this clarity is less apparent. Additionally, it remains unclear to many medical professionals how FL can be integrated into their daily work. For instance, during diagnosis, patients often describe symptoms with some degree of ambiguity (e.g., 'a not strong pain,' 'a medium pain,' 'a little bit of pain'). These degrees of truth should be considered by doctors, but managing numerous symptoms with varying degrees of truth can be very complicated. A system capable of handling such fuzziness would be effective in these cases. Furthermore, there is a limited number of high quality open-source medical datasets, whether tabular or image-based, available for research (Chrimes and Kim, 2022). The lack of open data in this field can also pose a significant barrier to the evaluation of new techniques.

Contributions to FL and related systems are still evolving, but there is a need to showcase more practical applications and simplified models across different domains to maximize the potential and fully leverage the benefits of this research area.

4.3 MQ2: Type of Contributions

As shown in Figure 4, two types of contribution are identified: Solution Proposal (SP), and Evaluation Research (ER).

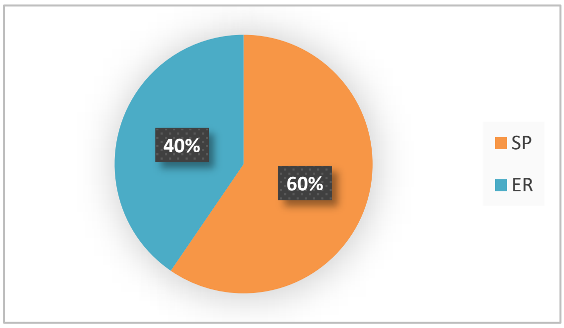

Figure 4: Type of contribution in the selected studies.

As illustrated in Figure 5, ER and SP are more prevalent compared to other types of contributions such as reviews or opinions. This indicates a significant interest in proposing and evaluating new FL-based interpretability techniques for medicine. Moreover, the prevalence of SP over evaluating existing FL techniques indicates that the field is still immature and requires further development. It's important to note that even when papers introduce a new approach, they still conduct evaluations using at least one dataset.

4.4 MQ3: Studied Diseases

The chart in Figure 5 displays the number of papers addressing different diseases. The distribution indicates a significant research focus on breast cancer and diabetes compared to other diseases. Breast cancer has the highest representation with 18 papers, followed by diabetes with 15 papers, and heart disease with 13 papers. Liver cancer and hepatitis each have 5 papers, while sleep disorder and mammography are addressed in 4 papers each. EEG signals related to bipolar disorder are discussed in 3 papers. Hypothyroid, mental health disorders, and bipolar disorder each have 2 papers. Additionally, there are 2 papers focusing on hepatobiliary disorders, Wisconsin, and Parkinson's.

Breast cancer is a significant health issue and is the leading cause of death among women worldwide (Zerouaoui and Idri, 2021). It has become a major focus in the field of ML for diagnosis, prognosis, and treatment. The importance of this topic and the availability of open-source data have contributed to its prominence in research, explaining why it is frequently studied in the selected papers.

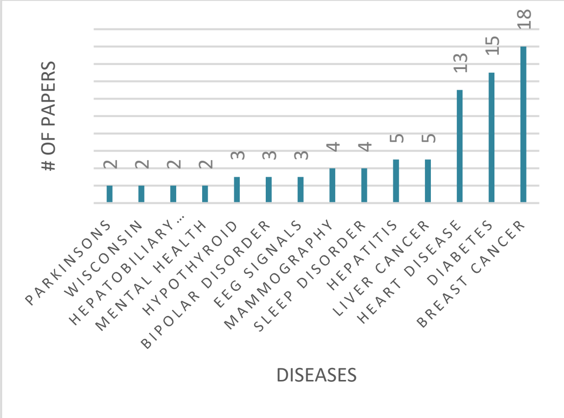

Figure 5: Most Studied Diseases.

4.5 MQ4: Types of FL Techniques

The selected studies have mainly either trained: (1) an FL-based ML model to leverage the interpretability features of FL (e.g. neuro-fuzzy systems for cancer diagnosis (Nguyen et al., 2022) or association rules for medical diagnosis based on medical records (Fernandez-Basso et al., 2022)), or (2) an ML model and then extracted FL rules from it to explain its decisions (e.g. rule extraction from SVM on lung cancer (Fung et al., 2005) or liver cancer (Chaves et al., 2005)). 14 of the selected studies used TSK fuzzy systems (e.g. (Shen et al., 2020; Zhou et al., 2021)), 9 of them specified the Mamdani category fuzzy system (e.g. (Ahmed et al., 2021; Liu et al., 2006)), while others didn't specify. Also, 36 of the papers mentioned using type-1 fuzzy systems.

32 papers used neuro-fuzzy systems and fuzzy linguistic rules (Nguyen et al., 2022) for a performance-interpretability tradeoff, while others used other techniques like the visualization (Sabol et al., 2019).

The research community has tended towards the use of the neuro-fuzzy framework. This can be explained by the fact that neuro-fuzzy networks combine both the powerful performance capabilities of ANNs and the interpretability that FL provides (Ouifak and Idri, 2023a). For example, (Nguyen et al., 2022) used the adaptive neuro-fuzzy system (ANFIS) (Jang, 1993), which is a popular model used across domains (Ouifak and Idri, 2023b). They combine fuzzy inference in a hierarchical architecture with attention to select the important rules to interpret the results of medical diagnosis. Others also used neuro-fuzzy systems for different tasks and diseases like sleep disorders (Juang et al., 2021), heart diseases (Bahani et al., 2021), and ovarian cancer (Tan et al., 2005) and showed the potential of FL system in interpreting ML rules, especially in the form of rules (Bahani et al., 2021; Chaves et al., 2005; Fung et al., 2005; Nguyen et al., 2022; Ouifak and Idri, 2023a).

5 CONCLUSION

This paper aimed to perform an SMS on the use of FL in the interpretability of ML black boxes in medicine. First, using a search string, a search was conducted in six different digital libraries. Second, a study selection process was performed, it started with identifying the papers within the scope of our SMS, and then the quality scores were computed to get only relevant papers. The study selection and quality assessment phases returned 67 relevant papers which were used to answer the MQs of this study. The main findings of each MQ are summarized below:

- MQ1. The data extracted to answer this MQ revealed that the interest in using FL to tackle the black box ML models is a hot research topic that is attracting attention once more. This was especially the case in 2022 with 15 papers. Moreover, two publication avenues were identified: journals and conferences.
- MQ2. Evaluation Research and Solution Proposal were the two main types of contributions made by the selected papers. Most of the selected papers conducted experiments and compared existing or new FL-based ML interpretability techniques.
- MQ3. Breast cancer and diabetes diseases were the most studied using FL techniques for ML interpretability.
- MQ4. Neuro-fuzzy systems specifically type-1 TSK systems are the most evaluated and studied to generate ML explanations.

Future work aims to delve deeper into neuro-fuzzy systems, which show great promise despite some limitations. One key issue is the loss of interpretability when using ensembles. To address this, we plan to develop a single rule base model that effectively represents the ensemble and maintains interpretability.

REFERENCES

Ahmed, U., Lin, J.C.W., Srivastava, G., 2021. Fuzzy Explainable Attention-based Deep Active Learning on Mental-Health Data. IEEE International Conference on Fuzzy Systems 2021-July. https://doi.org/10.1109/FUZZ45933.2021.9494423

Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Van Essen, B.C., Awwal, A.A.S., Asari, V.K., 2019. A state-of-the-art survey on deep learning theory and architectures. Electronics (Switzerland) 8, 292. https://doi.org/10.3390/electronics8030292

Bahani, K., Moujabbir, M., Ramdani, M., 2021. An accurate fuzzy rule-based classification systems for heart disease diagnosis. Sci Afr 14, e01019. https://doi.org/10.1016/J.SCIAF.2021.E01019

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 58, 82–115. https://doi.org/10.1016/J.INFFUS.2019.12.012

Bendaouia, A., Abdelwahed, E.H., Qassimi, S., Boussetta, A., Benzakour, I., Benhayoun, A., Amar, O., Bourzeix, F., Baïna, K., Cherkaoui, M., Hasidi, O., 2024. Hybrid features extraction for the online mineral grades determination in the flotation froth using Deep Learning. Eng Appl Artif Intell 129, 107680. https://doi.org/10.1016/J.ENGAPPAI.2023.107680

Carvalho, D. V., Pereira, E.M., Cardoso, J.S., 2019. Machine Learning Interpretability: A Survey on Methods and Metrics. Electronics 2019, Vol. 8, Page 832 8, 832. https://doi.org/10.3390/ELECTRONICS8080832

Chaddad, A., Lu, Q., Li, J., Katib, Y., Kateb, R., Tanougast, C., Bouridane, A., Abdulkadir, A., 2023. Explainable, Domain-Adaptive, and Federated Artificial Intelligence in Medicine. IEEE/CAA Journal of Automatica Sinica 10, 859–876. https://doi.org/10.1109/JAS.2023.123123

Chaves, A.D.C.F., Vellasco, M.M.B.R., Tanscheit, R., 2005. Fuzzy rule extraction from support vector machines. Proceedings - HIS 2005: Fifth International Conference on Hybrid Intelligent Systems 2005, 335–340. https://doi.org/10.1109/ICHIS.2005.51

Chrimes, D., Kim, C., 2022. Review of Publically Available Health Big Data Sets. 2022 IEEE International Conference on Big Data (Big Data) 6625–6627. https://doi.org/10.1109/BIGDATA55660.2022.10020258

Christoph, M., 2020. Interpretable Machine Learning A Guide for Making Black Box Models Explainable., Book.

Das, R., Sen, S., Maulik, U., 2020. A Survey on Fuzzy Deep Neural Networks. ACM Computing Surveys (CSUR) 53. https://doi.org/10.1145/3369798

de Campos Souza, P.V., 2020. Fuzzy neural networks and neuro-fuzzy networks: A review the main techniques and applications used in the literature. Appl Soft Comput 92, 106275. https://doi.org/10.1016/J.ASOC.2020.106275

Doshi-Velez, F., Kim, B., 2017. Towards A Rigorous Science of Interpretable Machine Learning.

Fernandez-Basso, C., Gutiérrez-Batista, K., Morcillo-Jiménez, R., Vila, M.A., Martin-Bautista, M.J., 2022. A

fuzzy-based medical system for pattern mining in a distributed environment: Application to diagnostic and co-morbidity. Appl Soft Comput 122, 108870. https://doi.org/10.1016/J.ASOC.2022.108870
Fisher, A., Rudin, C., Dominici, F., 2018. Model Class Reliance: Variable Importance Measures for any Machine Learning Model Class, from the.
Fung, G., Sandilya, S., Bharat Rao, R., 2005. Rule extraction from linear support vector machines. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 32–40. https://doi.org/10.1145/1081870.1081878
Jang, J.S.R., 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Trans Syst Man Cybern 23, 665–685. https://doi.org/10.1109/21.256541
Juang, C.F., Wen, C.Y., Chang, K.M., Chen, Y.H., Wu, M.F., Huang, W.C., 2021. Explainable fuzzy neural network with easy-to-obtain physiological features for screening obstructive sleep apnea-hypopnea syndrome. Sleep Med 85, 280–290. https://doi.org/10.1016/J.SLEEP.2021.07.012
Kitchenham, B., Charters, S., 2007. Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, School of Computer Science and Mathematics, Keele University.
Liu, F., Ng, G.S., Quek, C., Loh, T.F., 2006. Artificial ventilation modeling using neuro-fuzzy hybrid system. IEEE International Conference on Neural Networks - Conference Proceedings 2859–2864. https://doi.org/10.1109/IJCNN.2006.247215
Loyola-Gonzalez, O., 2019. Black-box vs. White-Box: Understanding their advantages and weaknesses from a practical point of view. IEEE Access 7, 154096–154113. https://doi.org/10.1109/ACCESS.2019.2949286
Lundberg, S.M., Allen, P.G., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions. Adv Neural Inf Process Syst 30.
Mamdani, E.H., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Int J Man Mach Stud 7, 1–13. https://doi.org/10.1016/S0020-7373(75)80002-2
Markowska-Kaczmar, U., Trelak, W., 2003. Extraction of Fuzzy Rules from Trained Neural Network Using Evolutionary Algorithm *. European Symposium on Artificial Neural Networks.
Nguyen, T.L., Kavuri, S., Park, S.Y., Lee, M., 2022. Attentive Hierarchical ANFIS with interpretability for cancer diagnostic. Expert Syst Appl 201, 117099. https://doi.org/10.1016/J.ESWA.2022.117099
Ouifak, H., Idri, A., 2023a. On the performance and interpretability of Mamdani and Takagi-Sugeno-Kang based neuro-fuzzy systems for medical diagnosis. Sci Afr e01610. https://doi.org/10.1016/J.SCIAF.2023.E01610
Ouifak, H., Idri, A., 2023b. Application of neuro-fuzzy ensembles across domains: A systematic review of the two last decades (2000–2022). Eng Appl Artif Intell 124, 106582. https://doi.org/10.1016/J.ENGAPPAI.2023.106582
Padrón-Tristán, J.F., Cruz-Reyes, L., Espín-Andrade, R.A., Llorente-Peralta, C.E., 2021. A Brief Review of Performance and Interpretability in Fuzzy Inference Systems. Studies in Computational Intelligence 966, 237–266. https://doi.org/10.1007/978-3-030-71115-3_11/TABLES/6
Sabol, P., Sincak, P., Ogawa, K., Hartono, P., 2019. Explainable Classifier Supporting Decision-making for Breast Cancer Diagnosis from Histopathological Images. Proceedings of the International Joint Conference on Neural Networks 2019-July. https://doi.org/10.1109/IJCNN.2019.8852070
Shen, T., Wang, J., Gou, C., Wang, F.Y., 2020. Hierarchical Fused Model with Deep Learning and Type-2 Fuzzy Learning for Breast Cancer Diagnosis. IEEE Transactions on Fuzzy Systems 28, 3204–3218. https://doi.org/10.1109/TFUZZ.2020.3013681
Sugeno, M., Kang, G.T., 1988. Structure identification of fuzzy model. Fuzzy Sets Syst 28, 15–33. https://doi.org/10.1016/0165-0114(88)90113-3
Tan, T.Z., Quek, C., Ng, G.S., 2005. Ovarian cancer diagnosis by hippocampus and neocortex-inspired learning memory structures. Neural Netw 18, 818–825. https://doi.org/10.1016/J.NEUNET.2005.06.027
Thai, H.T., 2022. Machine learning for structural engineering: A state-of-the-art review. Structures 38, 448–491. https://doi.org/10.1016/J.ISTRUC.2022.02.003
Yang, C., Huang, Q., Li, Z., Liu, K., Hu, F., 2017. Big Data and cloud computing: innovation opportunities and challenges. Int J Digit Earth 10, 13–53. https://doi.org/10.1080/17538947.2016.1239771
Zadeh, L.A., 1974. The Concept of a Linguistic Variable and its Application to Approximate Reasoning. Learning Systems and Intelligent Robots 1–10. https://doi.org/10.1007/978-1-4684-2106-4_1
Zerouaoui, H., Idri, A., 2021. Reviewing Machine Learning and Image Processing Based Decision-Making Systems for Breast Cancer Imaging. J Med Syst 45, 1–20. https://doi.org/10.1007/S10916-020-01689-1/FIGURES/19
Zhang, S., Sakulyeva, T.N., Pitukhin, E.A., Doguchaeva, S.M., Zhang, S., Sakulyeva, T.N., Pitukhin, E.A., Doguchaeva, S.M., 2020. Neuro-Fuzzy and Soft Computing - A Computational Approach to Learning and Artificial Intelligence. International Review of Automatic Control (IREACO) 13, 191–199. https://doi.org/10.15866/IREACO.V13I4.19179
Zhou, T., Zhou, Y., Gao, S., 2021. Quantitative-integration-based TSK fuzzy classification through improving the consistency of multi-hierarchical structure. Appl Soft Comput 106, 107350. https://doi.org/10.1016/J.ASOC.2021.107350
Zizaan, A., Idri, A., 2023. Machine learning based Breast Cancer screening: trends, challenges, and opportunities. Comput Methods Biomech Biomed Eng Imaging Vis 11, 976–996. https://doi.org/10.1080/21681163.2023.2172615