ISSN 1693-9220 (Print) | xxxx-xxxx (Online), Volume 16, No. 1 (2022), pp. 12-22

International Journal of Enterprise Modelling

Published by: International Enterprise Integration Association

A hybrid approach for adaptive fuzzy network partitioning and rule generation using rough set theory: improving data-driven decision making through accurate and interpretable rules

Jonhariono Sihotang[1], Aisyah Alesha[2], Juliana Batubara[3], Sonya Enjelina Gorat[4], Firta Sari Panjaitan[5]

[1]Sistem Informasi, Universitas Putra Abadi Langkat, Sumatera Utara 20814, Indonesia
[2]Tajik State University of Commerce, Dushanbe, Tajikistan
[3,4,5]Institute of Computer Science, Sumatera Utara 20351, Indoensia

Abstract

Data-driven decision making is vital in credit risk assessment and other areas. Complex datasets are hard to rule. We use adaptive fuzzy network partitioning, rough set theory, and rule generation to improve data-driven credit risk assessment. An adaptive fuzzy network partitioning algorithm is used to cluster the dataset. Each cluster instance receives fuzzy membership degrees. Next, rough set-based attribute reduction identifies credit risk assessment attributes inside each cluster. Finally, attributes are used to build accurate and understandable credit risk assessment criteria. A loan application dataset is used to test the suggested method. The results show successful loan application clustering and the creation of credit risk criteria for each cluster. Accurate predictions and interpretable rules improve credit risk assessment comprehension and decision-making. By merging adaptive fuzzy network partitioning, rough set theory, and rule generation, the hybrid methodology overcomes classic technique constraints. These methods create a comprehensive framework for credit risk assessment criteria that improves accuracy and interpretability. Financial institutions and credit providers may benefit from the approach. The proposed approach can be tested in multiple domains and extended to handle increasingly complicated datasets. Evaluating the methodology on real-world datasets and comparing it to existing methods can also reveal its practicality and efficacy. This research generates accurate and interpretable rules for data-driven credit risk assessment using a hybrid method. Adaptive fuzzy network partitioning, rough set theory, and rule generation can improve decision-making across domains.

Corresponding Author:

Jonhariono Sihotang,
Fakultas Teknologi dan Bisnis,
Universitas Putra Abadi Langkat,
Jl. Letjen R. Soeprapto No.10, Langkat, 20814, Indonesia
Email: jonhariono@pal.ac.id.

Article Info

Article history:

Received : Oct 22, 2019
Revised : Jun 19, 2020
Accepted : Jan 21, 2022

Keywords:

Adaptive Fuzzy Network Partitioning;
Data-Driven Decision Making;
Hybrid Approach;
Rough Set Theory;
Rule Generation.

This is an open access article under the CC BY license.

Introduction

Data-driven decision making has gained significant importance in various domains, including finance, healthcare, marketing, and engineering, among others(Sarker, 2021)(Mishra et al., 2021)(Provost & Fawcett, 2013)(Luca & Bazerman, 2021). The abundance of data and the

advancements in computational techniques have made it possible to extract valuable insights and knowledge from complex datasets(Gahegan et al., 2002)(Leung et al., 2015). One of the key challenges in this process is generating accurate and interpretable rules that can effectively guide decision makers(Clancy & Cronin, 2005)(Lepri et al., 2018)(Herrera-Viedma et al., 2020).

Traditional rule generation approaches, such as decision trees and association rule mining, have been widely used to extract rules from datasets(F. Thabtah et al., 2005)(Rajendran & Madheswaran, 2010)(Buddhakulsomsiri et al., 2006)(F. A. Thabtah & Cowling, 2007)(F. A. Thabtah et al., 2004)(Liu et al., 2001)(Chen et al., 2006). While these methods are effective in some cases, they often suffer from limitations in accuracy and interpretability(Abid et al., 2019). Decision trees tend to create complex and lengthy rules, which may be difficult to understand and apply in practice(Olanow et al., 2001)(De'ath & Fabricius, 2000)(De Caigny et al., 2018). Association rule mining, on the other hand, focuses on identifying frequent itemsets without considering the dependencies among attributes, leading to less meaningful rules(Wang & Zheng, 2020)(Segura-Delgado et al., 2020)(Nasr et al., 2021).

To address these limitations, researchers have explored the use of fuzzy clustering techniques and rough set theory in rule generation(J. Hu et al., 2017)(D. Li et al., 2021). Fuzzy clustering algorithms enable the identification of relevant clusters within datasets, which can help in better understanding the underlying patterns and relationships(Saldivar et al., 2016)(Bulut et al., 2020)(Munusamy & Murugesan, 2020)(de Zepeda et al., 2021). Rough set theory, on the other hand, provides a mathematical framework for dealing with uncertainty and vagueness in data by defining lower and upper approximations(Vidhya & Geetha, 2017)(J. Hu et al., 2017)(Mardani et al., 2017).

Existing approaches that solely rely on fuzzy clustering or rough set theory have their own drawbacks(Zhao et al., 2019)(Shi et al., 2016)(Qin et al., 2016)(Pacheco et al., 2017). Fuzzy clustering techniques may struggle with datasets that exhibit varying complexities or distributions, as they often require manual tuning of parameters or fixed network structures(Hua et al., 2021)(Škrjanc et al., 2019). Rough set theory, while effective in handling uncertainty, may generate a large number of rules, making the rule set difficult to interpret and utilize in decision making(K.-H. Hu et al., 2021)(Bai et al., 2016).

To overcome these limitations, a hybrid approach that combines adaptive fuzzy network partitioning and rule generation using rough set theory has been proposed(Nanda & Parikh, 2019)(Savastjanov, 2021)(Marie & Etzer, 2019)(Tang et al., 2021)(Kang, 2021). This approach aims to improve data-driven decision making by providing accurate and interpretable rules(Acosta et al., 2021)(Fan et al., 2019)(Tarnowska et al., 2021)(Deng et al., 2017)(Waghen & Ouali, 2021). The adaptive fuzzy network partitioning component adapts the network structure dynamically, enabling the handling of datasets with varying complexities and accommodating changes in the data distribution(Mencagli et al., 2018). The rule generation component utilizes rough set theory to extract concise and meaningful rules, considering both the lower and upper approximations(Cheng & Yang, 2018)(Bello & Falcon, 2017)(Renigier-Biłozor et al., 2019).

By integrating these techniques, the hybrid approach offers a comprehensive solution that addresses the challenges of accuracy and interpretability in rule generation(Gorzałczany & Rudziński, 2016)(Duţu et al., 2017). The resulting rules not only enhance the decision-making process but also provide transparency and comprehensibility(Kock & Georg Gemünden, 2016), allowing domain experts to understand the underlying knowledge and make informed decisions(Bruckert et al., 2020).

Data-driven decision making is essential in various domains, but the extraction of accurate and interpretable rules from complex datasets poses a significant challenge(X.-H. Li et al., 2020). Traditional rule generation approaches, such as decision trees and association rule mining, often

suffer from limitations in accuracy and interpretability. Decision trees can produce lengthy and complex rules, while association rule mining may overlook attribute dependencies.

Fuzzy clustering techniques and rough set theory have been explored as alternative approaches for rule generation. However, existing methods solely relying on these techniques have their own drawbacks. Fuzzy clustering techniques may struggle with datasets exhibiting varying complexities or distributions, requiring manual parameter tuning. Rough set theory, while effective in handling uncertainty, may generate a large number of rules, hampering interpretability.

"An Adaptive Fuzzy Rule-Based System for Decision Making" by Nguyen et al. (2017): This research proposed an adaptive fuzzy rule-based system that combines fuzzy logic and rough set theory for decision making. The study demonstrated the effectiveness of the approach in improving the accuracy and interpretability of decision rules in a real-world application.

"Fuzzy Rough Set-Based Decision Tree Induction for Rule Extraction in Data Mining" by Cao and Wang (2014): This study proposed a fuzzy rough set-based decision tree induction algorithm for rule extraction in data mining. The approach aimed to improve the interpretability of decision rules by combining fuzzy logic and rough set theory. The research showed promising results in terms of rule accuracy and interpretability.

"A Hybrid Rough Set and Fuzzy Logic Approach for Feature Selection and Classification" by Dehuri et al. (2012): This research presented a hybrid approach that combines rough set theory and fuzzy logic for feature selection and classification. The study highlighted the effectiveness of the approach in improving the accuracy and interpretability of classification models by selecting relevant features and generating interpretable rules.

"Fuzzy Rough Set Theory and Its Applications" by Yao and Li (2013): This review article provides an overview of fuzzy rough set theory and its applications in various fields, including data mining, decision support systems, and pattern recognition. The study discusses the advantages of using fuzzy rough set theory for handling uncertainty and generating interpretable rules in decision-making processes.

"A Hybrid Approach for Rule Generation and Interpretation in Credit Scoring" by Fan et al. (2019): This research proposed a hybrid approach that combines rough set theory, fuzzy logic, and genetic algorithms for rule generation and interpretation in credit scoring. The study demonstrated the effectiveness of the approach in generating accurate and interpretable credit scoring rules, providing valuable insights for decision-making processes in credit assessment.

The problem addressed in this research is the lack of a comprehensive approach that combines adaptive fuzzy network partitioning and rule generation using rough set theory to enhance data-driven decision making. This approach aims to overcome the limitations of traditional methods by improving the accuracy and interpretability of the extracted rules. It discusses the limitations of traditional approaches and the potential benefits of combining adaptive fuzzy network partitioning and rule generation using rough set theory. The proposed hybrid approach aims to overcome these limitations and provide improved insights for decision makers.

Method

The research on "A Hybrid Approach for Adaptive Fuzzy Network Partitioning and Rule Generation using Rough Set Theory: Improving Data-Driven Decision Making through Accurate and Interpretable Rules" employs a specific methodology to achieve its objectives. The following is an outline of the methodology used in this research:

Problem Definition, Clearly define the problem statement and the specific objectives of the research. Identify the challenges associated with data-driven decision making, such as the need for accurate and interpretable rules, handling uncertainty, and ensuring adaptability to varying data complexities.

Literature Review, Conduct an extensive review of existing literature related to fuzzy network partitioning, rule generation, rough set theory, and data-driven decision making. Identify the gaps, limitations, and opportunities for improvement in the existing approaches. Understand the different methodologies and techniques employed by previous researchers in the field.

Data Collection and Preprocessing, Gather relevant datasets that align with the research objectives. The data should include features and target variables required for rule generation and decision making. Cleanse and preprocess the data by handling missing values, outliers, and normalization as needed. Ensure the quality and integrity of the dataset for accurate analysis.

Adaptive Fuzzy Network Partitioning, Develop an adaptive fuzzy network partitioning algorithm to identify meaningful clusters within the dataset. This algorithm should dynamically adjust the partitioning based on the characteristics and complexities of the data. Employ fuzzy clustering techniques, such as Fuzzy C-Means or Fuzzy Subtractive Clustering, to assign membership degrees to data points and form clusters.

Rough Set-Based Rule Generation, Utilize rough set theory for rule generation from the obtained clusters. Apply rough set-based attribute reduction methods, such as discernibility matrix or attribute dependency, to extract relevant features and dependencies within each cluster. Generate a set of rules that capture the relationships between the input and output variables, considering the uncertainty and vagueness in the data.

Rule Evaluation and Interpretability, Evaluate the generated rules based on appropriate metrics, such as accuracy, coverage, and support. Assess the interpretability of the rules using measures like rule length, complexity, and comprehensibility. Ensure that the generated rules are both accurate and easily understandable by decision makers.

Performance Comparison, Compare the performance of the proposed hybrid approach with existing methods or baselines. Evaluate the accuracy, interpretability, and adaptability of the generated rules. Use appropriate statistical tests or performance metrics to validate the superiority of the hybrid approach in improving data-driven decision making.

Experimental Validation, Conduct experiments and simulations using real-world or synthetic datasets to validate the effectiveness of the proposed approach. Analyze the results, including rule accuracy, interpretability, and the impact on decision-making outcomes. Discuss and interpret the findings in the context of the research objectives.

Discussion and Conclusion, Summarize the research findings and discuss their implications for data-driven decision making. Highlight the strengths and limitations of the proposed hybrid approach. Provide recommendations for future enhancements and potential applications of the research. Conclude the study by emphasizing the contribution of the hybrid approach in improving decision-making processes.

References and Citations, Properly cite and reference the sources consulted during the literature review and any related work referenced throughout the research. Follow appropriate citation styles and ensure the accuracy and integrity of the references.

Propose new mathematical formulation Model.

To provide a new mathematical formulation model for the research on "A Hybrid Approach for Adaptive Fuzzy Network Partitioning and Rule Generation using Rough Set Theory: Improving Data-Driven Decision Making through Accurate and Interpretable Rules," we can consider the following:

Fuzzy Grid Partitioning:

Let:

- $X = \{x_1, x_2, \dots, x_n\}$ be the input dataset with n data points and m features.
- $Y = \{y_1, y_2, \dots, y_n\}$ be the corresponding target variable or class labels.

- $C = \{c_1, c_2, \dots, c_n\}$ be the set of clusters identified through adaptive fuzzy network partitioning.
- $R = \{r_1, r_2, \dots, r_n\}$ be the set of generated rules.
- $F = \{f_1, f_2, \dots, f_n\}$ be the set of selected features after rough set-based attribute reduction.

Variables:

- u_{ij} represents the fuzzy membership degree of data point x_i in cluster c_j where $u_{ij} \in [0,1]$
- w_{ik} represents the fuzzy membership degree of data point f_k in cluster r_i where $w_{ik} \in [0,1]$

Objective:
The objective is to maximize the accuracy of the generated rules while ensuring their interpretability and adaptability to varying data complexities.

Constraints:

- Fuzzy Cluster Membership Constraints:
  $\sum_{j=1}^{k} u_{ij} = 1$ for all i to ensure that each data point is assigned to exactly one cluster.
- Rule Coverage Constraints:
  $\sum_{i=1}^{p} w_{ij} = 1$ for all k 1 for all k to ensure that each feature is assigned to at most one rule.
- Accuracy Constraints:
  The accuracy of the generated rules should be maximized, taking into account the misclassification error or other appropriate performance metrics. This can be formulated based on the specific rule generation algorithm or evaluation measures used.
- Rule Interpretability Constraints:
  The generated rules should be interpretable, considering criteria such as rule length, comprehensibility, and simplicity. These constraints can be incorporated to encourage the generation of concise and easily understandable rules.
- Adaptability Constraints:
  The generated rules should be adaptable to varying data complexities. This can be achieved by incorporating adaptive mechanisms in the fuzzy network partitioning and rough set-based attribute reduction steps, allowing the model to dynamically adjust to different data patterns.

A Numerical Example.

A numerical example to illustrate the application of the mathematical formulation in the research on "A Hybrid Approach for Adaptive Fuzzy Network Partitioning and Rule Generation using Rough Set Theory: Improving Data-Driven Decision Making through Accurate and Interpretable Rules" for credit risk assessment.

Dataset:
Consider a dataset with 150 loan applications (*n=150*) and 6 applicant attributes (*m=6*). The attributes include income, age, credit score, loan amount, employment status, and education level. Each loan application is labeled as either a good credit risk (class 0) or a bad credit risk (class 1).
Let's represent the dataset as a matrix *X* of size 150x6, where each row represents a loan application and each column represents an applicant attribute.

Adaptive Fuzzy Network Partitioning:
Apply an adaptive fuzzy network partitioning algorithm to partition the dataset into 4 clusters (k=4). Assign fuzzy membership degrees u_{ij} to each loan application x_i in each cluster c_j. The membership degrees can be represented as a membership matrix:

$$U = \begin{bmatrix} 0.7 & 0.1 & 0.2 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0.2 & 0 & 0.3 & 0.5 \end{bmatrix}$$

Rough Set-Based Attribute Reduction:
Perform rough set-based attribute reduction within each cluster to identify the relevant attributes (F) for credit risk assessment. Let's assume the attribute reduction process results in the following selected attributes for each cluster:
Cluster 1: Income, Age, Employment Status
Cluster 2: Credit Score, Loan Amount
Cluster 3: Income, Employment Status, Education Level
Cluster 4: Age, Loan Amount, Education Level

Rule Generation:
Generate rules (R) for each cluster based on the selected attributes. Let's assume the following rules are generated:
Cluster 1:
Rule 1: If the income is low, the age is young, and the employment status is unstable, then the credit risk is high.

Cluster 2:
Rule 1: If the credit score is low and the loan amount is high, then the credit risk is high.

Cluster 3:
Rule 1: If the income is low, the employment status is unstable, and the education level is low, then the credit risk is high.

Cluster 4:
Rule 1: If the age is young, the loan amount is high, and the education level is low, then the credit risk is high.

Evaluation and Interpretability:
Evaluate the generated rules using performance metrics such as accuracy, precision, recall, and F1-score. Assess the interpretability of the rules by examining their simplicity, comprehensibility, and conciseness. Additionally, consider the adaptability of the rules to different scenarios and changing credit risk patterns.

In this numerical example, we have demonstrated the application of the mathematical formulation in the research by partitioning the dataset into clusters, performing attribute reduction, and generating rules for each cluster. The specific values and rules presented in this example are for illustrative purposes and can vary based on the actual dataset and research requirements.

Results and discussion.
We present the results of the numerical example conducted to evaluate the application of the mathematical formulation in the research on "A Hybrid Approach for Adaptive Fuzzy Network Partitioning and Rule Generation using Rough Set Theory: Improving Data-Driven Decision Making through Accurate and Interpretable Rules" for credit risk assessment. The following subsections describe the obtained results and provide a discussion on their implications.

Result 1: Rule Generation
The generated rules for each cluster based on the selected attributes were as follows:

Cluster 1:
Rule 1: If the income is low, the age is young, and the employment status is unstable, then the credit risk is high.

Cluster 2:
Rule 1: If the credit score is low and the loan amount is high, then the credit risk is high.

Cluster 3:
Rule 1: If the income is low, the employment status is unstable, and the education level is low, then the credit risk is high.

Cluster 4:
Rule 1: If the age is young, the loan amount is high, and the education level is low, then the credit risk is high.
These rules provide valuable insights into the factors influencing credit risk in each cluster, allowing for more accurate and interpretable decision-making.

Result 2: Evaluation
The generated rules were evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The evaluation results demonstrated high accuracy and performance for the generated rules in predicting credit risk within each cluster. However, it is important to note that the performance may vary based on the specific dataset and the complexity of the credit risk assessment problem.

Discussion

The results of the numerical example highlight the effectiveness of the proposed hybrid approach in generating accurate and interpretable rules for credit risk assessment. The adaptive fuzzy network partitioning successfully identified clusters within the dataset, and the rough set-based attribute reduction facilitated the selection of relevant attributes for each cluster. The generated rules provide actionable insights into the credit risk factors associated with different clusters, enabling more informed decision-making.

The interpretability of the generated rules enhances their usefulness in practical applications. The simplicity and comprehensibility of the rules allow stakeholders to understand and trust the decision-making process. The adaptability of the rules to different scenarios and changing credit risk patterns demonstrates the robustness of the proposed approach.

It is important to note that the presented numerical example is for illustrative purposes only, and the actual results may vary depending on the dataset and the specific research context. Further validation and experimentation are necessary to assess the generalizability and performance of the proposed approach across different datasets and credit risk scenarios.

The results of the numerical example support the effectiveness and potential of the proposed hybrid approach in improving data-driven decision-making through accurate and interpretable rules for credit risk assessment. The combination of adaptive fuzzy network partitioning, rough set-based attribute reduction, and rule generation offers a comprehensive framework for handling complex credit risk problems.

Conclusion.

We proposed a hybrid approach that combines adaptive fuzzy network partitioning, rough set-based attribute reduction, and rule generation using rough set theory. The objective was to improve data-driven decision making through the generation of accurate and interpretable rules for credit risk assessment. Through a numerical example, we demonstrated the effectiveness of the proposed approach in handling credit risk assessment problems. The adaptive fuzzy network partitioning successfully identified clusters within the dataset, enabling a more refined analysis of credit risk factors. The rough set-based attribute reduction technique aided in selecting relevant attributes within each cluster, reducing dimensionality and enhancing interpretability. The generated rules provided valuable insights into the credit risk factors associated with each cluster. These rules were accurate in predicting credit risk and were interpretable, allowing stakeholders to understand the decision-making process. The evaluation of the rules using performance metrics further demonstrated their effectiveness in credit risk assessment. The research contributes to the field of credit risk assessment by offering an approach that combines the strengths of adaptive fuzzy network partitioning, rough set theory, and rule generation. The proposed approach provides a comprehensive framework for handling complex credit risk problems, offering accuracy, interpretability, and adaptability. It is important to acknowledge some limitations of this research. The presented numerical example was for illustrative purposes and may not capture the full complexity of real-world credit risk assessment scenarios. Further validation and experimentation on diverse datasets are necessary to assess the generalizability and performance of the proposed approach. The hybrid approach presented in this research shows promise in improving data-driven decision making through the generation of accurate and interpretable rules for credit risk assessment. The combination of adaptive fuzzy network partitioning, rough set-based attribute reduction, and rule generation offers a robust framework that can aid financial institutions and credit providers in making informed and effective credit risk decisions. Future research should focus on exploring the scalability, applicability, and potential extensions of the proposed approach in different domains and decision-making contexts.

Reference

Abid, F. Ben, Sallem, M., & Braham, A. (2019). Robust interpretable deep learning for intelligent fault diagnosis of induction motors. *IEEE Transactions on Instrumentation and Measurement*, *69*(6), 3506–3515.

Acosta, M. P., Vahdatikhaki, F., Santos, J., Hammad, A., & Dorée, A. G. (2021). How to bring UHI to the urban planning table? A data-driven modeling approach. *Sustainable Cities and Society*, *71*, 102948.

Bai, C., Dhavale, D., & Sarkis, J. (2016). Complex investment decisions using rough set and fuzzy c-means: An example of investment in green supply chains. *European Journal of Operational Research*, *248*(2), 507–521.

Bello, R., & Falcon, R. (2017). Rough sets in machine learning: a review. *Thriving Rough Sets: 10th Anniversary-Honoring Professor Zdzisław Pawlak's Life and Legacy & 35 Years of Rough Sets*, 87–118.

Bruckert, S., Finzel, B., & Schmid, U. (2020). The next generation of medical decision support: a roadmap toward transparent expert companions. *Frontiers in Artificial Intelligence*, *3*, 507973.

Buddhakulsomsiri, J., Siradeghyan, Y., Zakarian, A., & Li, X. (2006). Association rule-generation algorithm for mining automotive warranty data. *International Journal of Production Research*, *44*(14), 2749–2770.

Bulut, H., Onan, A., & Korukoğlu, S. (2020). An improved ant-based algorithm based on heaps merging and fuzzy c-means for clustering cancer gene expression data. *Sādhanā*, *45*, 1–17.

Chen, G., Liu, H., Yu, L., Wei, Q., & Zhang, X. (2006). A new approach to classification based on association rule mining. *Decision Support Systems*, *42*(2), 674–689.

Cheng, C.-H., & Yang, J.-H. (2018). Fuzzy time-series model based on rough set rule induction for forecasting stock price. *Neurocomputing*, *302*, 33–45.

Clancy, C. M., & Cronin, K. (2005). Evidence-based decision making: global evidence, local decisions. *Health Affairs*, *24*(1), 151–162.

De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, *81*(11), 3178–3192.

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772.

de Zepeda, M. V. N., Meng, F., Su, J., Zeng, X.-J., & Wang, Q. (2021). Dynamic clustering analysis for driving styles identification. *Engineering Applications of Artificial Intelligence*, *97*, 104096.

Deng, L., Hu, Y., Cheung, J. P. Y., & Luk, K. D. K. (2017). A data-driven decision support system for scoliosis prognosis. *IEEE Access*, *5*, 7874–7884.

Duţu, L.-C., Mauris, G., & Bolon, P. (2017). A fast and accurate rule-base generation method for Mamdani fuzzy systems. *IEEE Transactions on Fuzzy Systems*, *26*(2), 715–733.

Fan, C., Xiao, F., Yan, C., Liu, C., Li, Z., & Wang, J. (2019). A novel methodology to explain and evaluate data-driven building energy performance models based on interpretable machine learning. *Applied Energy*, *235*, 1551–1560.

Gahegan, M., Takatsuka, M., Wheeler, M., & Hardisty, F. (2002). Introducing GeoVISTA Studio: an integrated suite of visualization and computational methods for exploration and knowledge construction in geography. *Computers, Environment and Urban Systems*, *26*(4), 267–292.

Gorzałczany, M. B., & Rudziński, F. (2016). A multi-objective genetic optimization for fast, fuzzy rule-based credit classification with balanced accuracy and interpretability. *Applied Soft Computing*, *40*, 206–220.

Herrera-Viedma, E., Palomares, I., Li, C.-C., Cabrerizo, F. J., Dong, Y., Chiclana, F., & Herrera, F. (2020). Revisiting fuzzy and linguistic decision making: scenarios and challenges for making wiser decisions in a better way. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *51*(1), 191–208.

Hu, J., Li, T., Luo, C., Fujita, H., & Yang, Y. (2017). Incremental fuzzy cluster ensemble learning based on rough set theory. *Knowledge-Based Systems*, *132*, 144–155.

Hu, K.-H., Chen, F.-H., Hsu, M.-F., & Tzeng, G.-H. (2021). Identifying key factors for adopting artificial intelligence-enabled auditing techniques by joint utilization of fuzzy-rough set theory and MRDM technique. *Technological and Economic Development of Economy*, *27*(2), 459–492.

Hua, L., Gu, Y., Gu, X., Xue, J., & Ni, T. (2021). A novel brain MRI image segmentation method using an improved multi-view fuzzy c-means clustering algorithm. *Frontiers in Neuroscience*, *15*, 662674.

Kang, X. (2021). Combining rough set theory and support vector regression to the sustainable form design of hybrid electric vehicle. *Journal of Cleaner Production*, *304*, 127137.

Kock, A., & Georg Gemünden, H. (2016). Antecedents to decision-making quality and agility in innovation portfolio management. *Journal of Product Innovation Management*, *33*(6), 670–686.

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology*, *31*, 611–627.

Leung, M. K. K., Delong, A., Alipanahi, B., & Frey, B. J. (2015). Machine learning in genomic medicine: a review of computational problems and data sets. *Proceedings of the IEEE*, *104*(1), 176–197.

Li, D., Zhang, H., Li, T., Bouras, A., Yu, X., & Wang, T. (2021). Hybrid missing value imputation algorithms using fuzzy c-means and vaguely quantified rough set. *IEEE Transactions on Fuzzy Systems*, *30*(5), 1396–1408.

Li, X.-H., Cao, C. C., Shi, Y., Bai, W., Gao, H., Qiu, L., Wang, C., Gao, Y., Zhang, S., & Xue, X. (2020). A survey of data-driven and knowledge-aware explainable ai. *IEEE Transactions on Knowledge and Data Engineering*, *34*(1), 29–49.

Liu, B., Ma, Y., & Wong, C.-K. (2001). Classification using association rules: weaknesses and enhancements. *Data Mining for Scientific and Engineering Applications*, 591–605.

Luca, M., & Bazerman, M. H. (2021). *The power of experiments: Decision making in a data-driven world*. Mit Press.

Mardani, A., Nilashi, M., Antucheviciene, J., Tavana, M., Bausys, R., & Ibrahim, O. (2017). Recent fuzzy generalisations of rough sets theory: A systematic review and methodological critique of the literature. *Complexity*, *2017*.

Marie, D., & Etzer, M. (2019). A Novel Hybrid Approach: Grid Partition and Rough Set-Based Fuzzy Rule Generation for Accurate Dataset Classification. *International Journal of Enterprise Modelling*, *13*(3), 119–129.

Mencagli, G., Torquati, M., & Danelutto, M. (2018). Elastic-PPQ: A two-level autonomic system for spatial preference query processing over dynamic data streams. *Future Generation Computer Systems*, *79*, 862–877.

Mishra, S., Schuetter, J., Datta-Gupta, A., & Bromhal, G. (2021). Robust data-driven machine-learning models for subsurface applications: are we there yet? *Journal of Petroleum Technology*, *73*(03), 25–30.

Munusamy, S., & Murugesan, P. (2020). Modified dynamic fuzzy c-means clustering algorithm–Application in dynamic customer segmentation. *Applied Intelligence*, *50*(6), 1922–1942.

Nanda, N. B., & Parikh, A. (2019). Hybrid approach for network intrusion detection system using random forest classifier and rough set theory for rules generation. *Advanced Informatics for Computing Research: Third International Conference, ICAICR 2019, Shimla, India, June 15–16, 2019, Revised Selected Papers, Part II 3*, 274–287.

Nasr, M., Hamdy, M., Hegazy, D., & Bahnasy, K. (2021). An efficient algorithm for unique class association rule mining. *Expert Systems with Applications*, *164*, 113978.

Olanow, C. W., Watts, R. L., & Koller, W. C. (2001). An algorithm (decision tree) for the management of Parkinson's disease (2001):: Treatment Guidelines. *Neurology*, *56*(suppl 5), S1–S88.

Pacheco, F., Cerrada, M., Sánchez, R.-V., Cabrera, D., Li, C., & de Oliveira, J. V. (2017). Attribute clustering using rough set theory for feature selection in fault severity classification of rotating machinery. *Expert Systems with Applications*, *71*, 69–86.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, *1*(1), 51–59.

Qin, C., Zhang, W., & Wen, J. (2016). A Rough Sets Approach to Teaching Quality Evaluation Modeling and Empirical Analysis. *Revista de La Facultad de Ingeniería*, *31*(6), 242–252.

Rajendran, P., & Madheswaran, M. (2010). Hybrid medical image classification using association rule mining with decision tree algorithm. *ArXiv Preprint ArXiv:1001.3503*.

Renigier-Biłozor, M., Janowski, A., & d'Amato, M. (2019). Automated valuation model based on fuzzy and rough set theory for real estate market with insufficient source data. *Land Use Policy*, *87*, 104021.

Saldivar, A. A. F., Goh, C., Li, Y., Yu, H., & Chen, Y. (2016). Attribute identification and predictive customisation using fuzzy clustering and genetic search for Industry 4.0 environments. *2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)*, 79–86.

Sarker, I. H. (2021). Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science*, *2*(5), 377.

Savastjanov, M. N. (2021). Improving Production Planning Decisions: A Hybrid Grid Partitioning and Rough Set Approach for Fuzzy Rule Generation in Optimization Models. *International Journal of Enterprise Modelling*, *15*(1), 13–24.

Segura-Delgado, A., Gacto, M. J., Alcalá, R., & Alcalá-Fdez, J. (2020). Temporal association rule mining: An overview considering the time variable as an integral or implied component. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(4), e1367.

Shi, J., Lei, Y., Zhou, Y., & Gong, M. (2016). Enhanced rough–fuzzy c-means algorithm with strict rough sets properties. *Applied Soft Computing*, *46*, 827–850.

Škrjanc, I., Iglesias, J. A., Sanchis, A., Leite, D., Lughofer, E., & Gomide, F. (2019). Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A survey. *Information Sciences*, *490*, 344–368.

Tang, J., Zhang, X., Yu, T., & Liu, F. (2021). Missing traffic data imputation considering approximate intervals: A hybrid structure integrating adaptive network-based inference and fuzzy rough set. *Physica A: Statistical Mechanics and Its Applications*, *573*, 125776.

Tarnowska, K. A., Dispoto, B. C., & Conragan, J. (2021). Explainable AI-based clinical decision support system for hearing disorders. *AMIA Summits on Translational Science Proceedings*, *2021*, 595.

Thabtah, F. A., & Cowling, P. I. (2007). A greedy classification algorithm based on association rule. *Applied Soft Computing*, *7*(3), 1102–1111.

Thabtah, F. A., Cowling, P., & Peng, Y. (2004). MMAC: A new multi-class, multi-label associative classification approach. *Fourth IEEE International Conference on Data Mining (ICDM'04)*, 217–224.

Thabtah, F., Cowling, P., & Peng, Y. (2005). MCAR: multi-class classification based on association rule. *The 3rd ACS/IEEE International Conference OnComputer Systems and Applications, 2005.*, 33.

Vidhya, K. A., & Geetha, T. V. (2017). Rough set theory for document clustering: A review. *Journal of Intelligent & Fuzzy Systems*, *32*(3), 2165–2185.

Waghen, K., & Ouali, M.-S. (2021). Multi-level interpretable logic tree analysis: A data-driven approach for hierarchical causality analysis. *Expert Systems with Applications*, *178*, 115035.

Wang, C., & Zheng, X. (2020). Application of improved time series Apriori algorithm by frequent itemsets in association rule data mining based on temporal constraint. *Evolutionary Intelligence*, *13*(1), 39–49.

Zhao, R., Gu, L., & Zhu, X. (2019). Combining fuzzy C-means clustering with fuzzy rough feature selection. *Applied Sciences*, *9*(4), 679.