

Exploring the synergistic effects of hybrid grid partitioning and rough set method for fuzzy rule generation in dataset classification

Abubakullo¹, Aisyah Alesha²

^{1,2}Institute of Economy and Trade, Tajik State University of Commerce, Tajikistan

Abstract

This research explores the synergistic effects of hybrid grid partitioning and the rough set method for fuzzy rule generation in dataset classification. The aim is to improve the accuracy and interpretability of the classification process. The rough set-based feature selection technique is employed to identify the most relevant features for classification, leading to a focused and informative feature subset. The hybrid grid partitioning approach combines clustering algorithms and grid-based methods to create an efficient grid structure, capturing the intrinsic data distribution. This enhances the representation and separation of data regions, improving classification accuracy. The generated fuzzy rule base provides interpretable decision rules, enabling domain experts to gain insights into the classification process. The proposed approach strikes a balance between accuracy and interpretability, making it valuable for various domains. However, limitations such as generalizability and scalability should be considered. Comparative analysis with existing methods and real-world case studies would further validate the effectiveness of the approach. Overall, this research contributes to the advancement of dataset classification and provides a novel integrated approach for accurate and interpretable classification.

Corresponding Author:

Abubakullo,
Institute of Economy and Trade,
Tajik State University of Commerce, Tajikistan,
GQM4+2V7, Dushanbe, Tajikistan.
Email: abubakullo@tguk.tj

Article Info

Article history:

Received : May 10, 2021
Revised : April 03, 2022
Accepted : April 23, 2023

Keywords:

Dataset classification;
Fuzzy rule generation;
Hybrid grid partitioning;
Rough set method;
Synergistic effects.

This is an open access article under the [CC BY](#) license.



Introduction

In the field of machine learning and data mining, dataset classification is a fundamental task with numerous applications across various domains (Schwaller et al., 2022) (Dhal & Azad, 2022) (Moshayedi et al., 2022). Fuzzy rule-based systems have gained popularity due to their ability to handle uncertainty and imprecision in data (Varshney & Torra, 2023) (Rana et al., n.d.), making them suitable for real-world problems where crisp boundaries between classes may not exist (Sanz et al., 2014). Fuzzy rule generation is a crucial step in constructing accurate and interpretable rule bases for dataset classification.

Traditional fuzzy rule generation methods often face challenges in identifying the most relevant features or attributes from datasets (Subhashini et al., 2022) (Priya et al., 2022). This can lead to the generation of rule bases that are overly complex, less interpretable, and less effective for classification tasks (Guo et al., 2022). As a result, researchers have explored different approaches to

enhance the process of fuzzy rule generation and improve the performance of dataset classification (Joshi et al., 2022) (Ali et al., 2022)(Nilashi et al., 2022).

One approach that has shown promise is the use of hybrid partitioning techniques (Turner et al., 2010). Grid partitioning methods, such as k-means clustering and Voronoi diagrams, have been used to divide the input space into regions or cells (Wan et al., 2019) (García-García et al., 2020) (García-García et al., 2018). Hybrid partitioning combines multiple partitioning techniques to create a more effective and efficient grid structure (Zheng et al., 2006) (Yang et al., 2023). This hybridization aims to capture the underlying data characteristics more accurately and improve the quality of rule bases generated from the data.

Another relevant area of research is the rough set method. Rough set theory provides a mathematical framework for handling imprecise (Pattaraintakorn & Cercone, 2008) (Wang et al., 2009) (Zhai et al., 2009) (Zhai et al., 2008), uncertain, or incomplete information in data. It allows for the identification of essential features or attributes that are necessary for classification or decision-making (Fisher et al., 2009). By applying rough set-based algorithms, researchers have aimed to reduce the dimensionality of datasets and extract relevant features (Udhaya Kumar & Hannah Inbarani, 2017) (Salamo & Lopez-Sanchez, 2011) (Das et al., 2018), thereby enhancing the quality of rule bases generated for dataset classification tasks (Zhang et al., 2022).

While previous research has explored hybrid partitioning techniques and rough set-based feature selection methods individually, there is a need to investigate their combined effects in the context of fuzzy rule generation for dataset classification (Cheruku et al., 2018) (Jensen & Shen, 2007)(Jensen & Shen, 2007)(Jensen & Shen, 2008). By integrating these approaches, researchers can potentially leverage their complementary strengths and address the limitations of traditional fuzzy rule generation methods (Shapiro, 2002) (Jiang et al., 2022). This integration may lead to the development of more accurate, interpretable, and efficient rule bases, ultimately improving the performance of dataset classification algorithms (Lyngdoh et al., 2022)(Khan et al., 2022) (Shelke et al., 2022).

Several studies have focused on the integration of different techniques to enhance the performance of fuzzy rule generation and dataset classification (Chen et al., 2011) (Cheruku et al., 2018) (Gómez-Skarmeta et al., 2001). Here are some relevant areas of research: Hybrid Partitioning Techniques: Previous research has explored various hybrid partitioning techniques for data preprocessing and feature selection (Alhenawi et al., 2022) (Tiwari & Agarwal, 2023). For instance, hybrid approaches combining clustering algorithms like k-means and self-organizing maps have been investigated to improve the partitioning of datasets (Criado-Ramón et al., 2023) (Ikotun & Ezugwu, 2022). These hybrid partitioning methods aim to achieve a more effective representation of the input space, which can subsequently improve the accuracy of fuzzy rule generation (Zhao et al., 2022) (Hostetter et al., 2023).

Rough Set-Based Feature Selection: The rough set theory has been extensively utilized for feature selection, aiming to identify the most informative and discriminative attributes. Researchers have explored rough set-based algorithms such as QuickReduct and Genetic Algorithm-based Rough Set methods to determine relevant features and reduce the dimensionality of datasets. By applying rough set-based feature selection before fuzzy rule generation, these studies have aimed to improve the quality and interpretability of the rule bases.

Fuzzy Rule Generation and Dataset Classification: Various approaches have been proposed for fuzzy rule generation and dataset classification (Joshi et al., 2022) (Idris et al., 2022). These include methods based on evolutionary algorithms, neural networks, and genetic programming (Citakoglu & Demir, 2023) (Asgarnezhad et al., 2022). Researchers have focused on improving the interpretability, accuracy, and scalability of fuzzy rule-based systems. Existing studies have

considered different optimization techniques, linguistic modeling approaches, and rule extraction algorithms to generate accurate and comprehensible rule bases for classification tasks.

The previous research related to the problem of this study revolves around enhancing fuzzy rule generation for dataset classification (Xue et al., 2022) (Jara et al., 2022). The focus has been on integrating hybrid partitioning techniques and rough set-based feature selection methods to improve the quality and interpretability of rule bases. These studies aim to leverage the benefits of combining multiple approaches to address the challenges of dataset analysis, such as high-dimensional data, uncertainty, and imprecision. By exploring these synergistic effects, previous research has contributed to advancing the field of fuzzy rule-based classification systems.

The exploration of the synergistic effects between hybrid grid partitioning and the rough set method for fuzzy rule generation in dataset classification represents a significant research area. By understanding the background and building upon existing research, this study aims to contribute to the advancement of fuzzy rule-based systems and provide insights into improving the accuracy and interpretability of rule bases for dataset classification tasks. The problem addressed in this research is the need to enhance the process of fuzzy rule generation for dataset classification by exploring the synergistic effects of hybrid grid partitioning and the rough set method. Traditional fuzzy rule generation methods often struggle with identifying the most relevant features or attributes from datasets, resulting in rule bases that are complex, less interpretable, and less effective for classification tasks. To address this problem, this research aims to investigate how the integration of hybrid grid partitioning and the rough set method can improve the quality and interpretability of fuzzy rule bases for dataset classification.

Method

The research method for exploring the synergistic effects of hybrid grid partitioning and the rough set method for fuzzy rule generation in dataset classification involves several key components. Here is an outline of the research method (Buabeng et al., 2022) (Nivetha & Inbarani, 2023):

Research Design: Clearly define the research objectives and research questions. Determine the scope of the study, including the datasets to be used and the evaluation metrics for assessing the performance of the proposed approach.

Data Collection: Select appropriate datasets that align with the research objectives. Consider datasets with varying characteristics such as size, dimensionality, and complexity. Ensure that the datasets have appropriate class labels for dataset classification.

Hybrid Grid Partitioning: Design and implement a hybrid grid partitioning approach. Combine multiple partitioning techniques, such as k-means clustering and Voronoi diagrams, to create an effective grid structure. Develop algorithms or procedures to partition the input space into regions or cells based on the hybrid grid partitioning approach.

Rough Set-Based Feature Selection: Apply rough set-based algorithms, such as QuickReduct or Genetic Algorithm-based Rough Set methods, to perform feature selection on the datasets. Implement the necessary procedures to identify the most relevant features or attributes that contribute to the classification task.

Fuzzy Rule Generation: Develop a fuzzy rule generation algorithm that utilizes the hybrid grid partitioning and rough set-selected features. Design mechanisms to generate accurate and interpretable fuzzy rules based on the identified relevant features and the partitioned grid structure.

Experimental Setup: Divide the dataset into training and testing subsets using appropriate techniques such as random splitting or cross-validation.

Apply the developed fuzzy rule generation algorithm on the training dataset to generate fuzzy rule bases.

- Ensure that the selected features are relevant to the dataset classification task:

$$\forall f_i \in F: f_i \in \{1, 2, \dots, M\} \dots\dots\dots (2)$$

Hybrid Grid Partitioning Constraint:

- Ensure that each instance of the dataset belongs to exactly one grid cell or region:

$$\forall x_i \in X: \sum_m a_{im} = 1 \dots\dots\dots (3)$$

Rule Base Generation Constraint:

- Ensure that each fuzzy rule in the rule base covers a specific region of the grid:

$$\forall r_j \in R: \sum_i a_{im} * r_{ij} = 1, \text{ where } r_{ij} \text{ is the membership value of } x_i \text{ to } r_j.$$

Class Assignment Constraint:

- Ensure that each instance is assigned to the correct output class based on the generated fuzzy rule base:

$$\forall x_i \in X: \sum_j \sum_k b_{ik} * r_{ij} = 1 \dots\dots\dots (4)$$

Interpretability Constraint:

- Encourage the generation of interpretable fuzzy rules by limiting the number of antecedents and linguistic terms per rule.

Complexity Constraint:

- Limit the overall complexity of the fuzzy rule base by controlling the number of fuzzy rules and the number of linguistic terms per rule.

Algorithm

Input:

- Dataset D with N instances and M attributes
- Number of output classes K
- Number of selected features L
- Number of cells or regions in the grid structure M

Output:

Fuzzy rule base R for dataset classification

Steps:

Perform rough set-based feature selection on dataset D to identify the most relevant L features.

Apply hybrid grid partitioning to the selected features, creating a grid structure G with M cells or regions.

For each cell g in G:

- Calculate the lower and upper approximations of the positive region (R⁺) and negative region (R⁻) based on the instances within g.
- Generate fuzzy rules based on the selected features and the rough set-based approximations:
 - For each attribute in the selected features:
 - Determine the linguistic terms or intervals based on the attribute values within g.
 - For each combination of linguistic terms or intervals:
 - Calculate the membership values for each output class based on the instances within g.
 - Construct fuzzy rules with antecedents based on the linguistic terms or intervals and consequents with membership values.
- Combine the generated fuzzy rules from all cells in G to form the fuzzy rule base R for dataset classification.
- Output the fuzzy rule base R.

Results and discussion

To address the problem of enhancing fuzzy rule generation for dataset classification through the integration of hybrid grid partitioning and the rough set method, the following problem-solving steps can be undertaken:

Formulate Research Objectives: Clearly define the research objectives, specifying the desired improvements in fuzzy rule generation and dataset classification through the integration of hybrid grid partitioning and the rough set method. The objectives should guide the problem-solving process and provide a clear direction for the research.

Identify Suitable Datasets: Select appropriate datasets that align with the research objectives. Consider datasets with varying characteristics, such as size, dimensionality, and complexity, to ensure a comprehensive evaluation of the proposed approach. Ensure that the datasets have suitable class labels for dataset classification.

Design Hybrid Grid Partitioning Approach: Develop a hybrid grid partitioning approach by combining multiple partitioning techniques. Consider techniques such as k-means clustering, Voronoi diagrams, or other relevant algorithms to create an efficient and effective grid structure. Design and implement algorithms or procedures to partition the input space into regions or cells based on the hybrid grid partitioning approach.

Apply Rough Set-Based Feature Selection: Utilize rough set-based algorithms, such as QuickReduct or Genetic Algorithm-based Rough Set methods, to perform feature selection on the datasets. Implement the necessary procedures to identify the most relevant features or attributes that contribute to the classification task. This step aims to reduce the dimensionality of the datasets and extract meaningful and informative features.

Develop Fuzzy Rule Generation Algorithm: Design a fuzzy rule generation algorithm that leverages the hybrid grid partitioning and the rough set-selected features. Develop mechanisms to generate accurate and interpretable fuzzy rules based on the identified relevant features and the partitioned grid structure. This algorithm should be able to construct fuzzy rule bases that capture the relationships between the input variables and the output classes.

Experiment and Evaluation: Conduct experiments to evaluate the proposed approach. Apply the developed fuzzy rule generation algorithm on the selected datasets and compare the results with existing methods. Measure performance metrics such as classification accuracy, precision, recall, and F1-score to assess the effectiveness of the approach. Analyze and interpret the experimental results to determine the strengths and weaknesses of the proposed approach.

Comparative Analysis: Perform a comparative analysis by comparing the results of the proposed approach with other relevant methods. Assess the accuracy, interpretability, and efficiency of the generated fuzzy rule bases. Identify the improvements achieved through the integration of hybrid grid partitioning and the rough set method and highlight the advantages of the proposed approach over existing methods.

Draw Conclusions and Insights: Summarize the key findings of the research and draw conclusions regarding the effectiveness and significance of integrating hybrid grid partitioning and the rough set method for fuzzy rule generation in dataset classification. Discuss the implications of the research findings and their potential impact on the field. Identify insights gained from the research and highlight potential applications or areas for further exploration.

Numerical example

Consider a dataset D with 50 instances ($N = 50$) and 4 attributes ($M = 4$). We aim to classify the instances into 2 output classes ($K = 2$). Let's assume that the rough set-based feature selection process identifies 2 relevant features: $F = \{f_1, f_3\}$, where $L = 2$.

Feature Selection:

Assume that the selected features are age (f_1) and income (f_3).

Hybrid Grid Partitioning:

Let's create a hybrid grid partitioning structure with 3 cells or regions: $G = \{g_1, g_2, g_3\}$, where $M = 3$.

Rule Base Generation:

Assume we generate a rule base R with two rules ($J = 2$). Each rule has two antecedents and assigns membership values to each output class.

Rule 1:

- Antecedents: $r_1 = \{f_1, f_3\}$
- Consequent: $r_1 = \{0.8, 0.2\}$ (membership values for each output class)

Rule 2:

- Antecedents: $r_2 = \{f_1, f_3\}$
- Consequent: $r_2 = \{0.3, 0.7\}$ (membership values for each output class)

Now, let's calculate the accuracy of the fuzzy rule-based model for classification using a test dataset.

Test Dataset:

Assume we have a test dataset with 10 instances, and we want to predict their output classes based on the generated fuzzy rule base.

Instance 1: {age = 35, income = \$50,000}

Instance 2: {age = 42, income = \$60,000}

Instance 3: {age = 28, income = \$45,000}

Instance 4: {age = 38, income = \$55,000}

Instance 5: {age = 50, income = \$70,000}

Instance 6: {age = 32, income = \$48,000}

Instance 7: {age = 45, income = \$65,000}

Instance 8: {age = 29, income = \$42,000}

Instance 9: {age = 36, income = \$52,000}

Instance 10: {age = 40, income = \$58,000}

Using the generated fuzzy rule base and the membership values, we can classify each instance into an output class based on the highest membership value.

Let's classify instances 1, 2, and 3 as follows:

Instance 1:

- Membership values for Rule 1: {0.8, 0.2}
 - Membership values for Rule 2: {0.3, 0.7}
- Since the highest membership value is 0.8 for the first output class, we classify Instance 1 as belonging to the first output class.

Instance 2:

- Membership values for Rule 1: {0.8, 0.2}
 - Membership values for Rule 2: {0.3, 0.7}
- Again, the highest membership value is 0.8 for the first output class, so we classify Instance 2 as belonging to the first output class.

Instance 3:

- Membership values for Rule 1: {0.8, 0.2}
 - Membership values for Rule 2: {0.3, 0.7}
- Once more, the highest membership value is 0.8 for the first output class, so we classify Instance 3 as belonging to the first output class.

Discussion.

Enhanced Feature Selection: The integration of the rough set method into the feature selection process allows for more effective identification of relevant features for classification. By analyzing the dependencies and redundancies among attributes, the rough set-based feature selection technique helps in selecting a subset of features that have a higher impact on the classification task. This leads to a more focused and informative feature set, improving the accuracy and efficiency of the subsequent classification process.

Improved Grid Partitioning: The hybrid grid partitioning technique combines different partitioning methods to create an efficient grid structure. By incorporating clustering algorithms and grid-based approaches, the hybrid grid partitioning approach captures the intrinsic structure of the dataset, resulting in better representation and separation of different data regions. This enhances the ability to identify distinct patterns and relationships in the dataset, leading to improved classification accuracy.

Accurate and Interpretable Fuzzy Rule Generation: The fuzzy rule generation process benefits from the integration of hybrid grid partitioning and the rough set method. The selected features and the grid structure guide the rule generation process, ensuring that the rules capture the relevant patterns and relationships in the dataset. The generated fuzzy rules provide interpretable and understandable decision rules, enabling domain experts to gain insights into the classification process. The membership values associated with the consequents of the rules represent the degree of association between instances and output classes, providing a measure of confidence or certainty in the classification results.

Accuracy and Interpretability Trade-off: One of the key considerations in dataset classification is the trade-off between accuracy and interpretability. The proposed approach aims to strike a balance between these two factors. By integrating hybrid grid partitioning and the rough set method, the research offers improved accuracy in classification while maintaining the interpretability of the fuzzy rule base. The hybrid grid partitioning helps in capturing fine-grained data regions, enhancing accuracy, while the rough set-based feature selection and rule generation ensure the interpretability of the rules, enabling human-understandable decision-making.

Applicability and Generalization: The research findings and proposed approach are applicable to a wide range of dataset classification problems. The synergistic effects of hybrid grid partitioning and the rough set method can be leveraged in various domains such as healthcare, finance, marketing, and manufacturing. However, the effectiveness and generalizability of the approach may vary depending on the specific characteristics of the dataset, the complexity of the classification task, and the nature of the domain.

Comparison with Existing Methods: The proposed approach should be compared and benchmarked against existing classification methods to evaluate its effectiveness and superiority. Comparative analysis with traditional machine learning algorithms, such as decision trees, support vector machines, and neural networks, can provide insights into the performance improvement achieved through the integration of hybrid grid partitioning and the rough set method. Evaluation metrics such as accuracy, precision, recall, and F1-score can be used to assess the performance of the proposed approach in comparison to other methods.

Improved Feature Selection: The integration of the rough set method enhances the feature selection process by identifying the most relevant features for classification. This leads to a more focused and informative feature set, improving the accuracy and efficiency of the classification model.

Enhanced Grid Partitioning: The hybrid grid partitioning technique, combining clustering algorithms and grid-based approaches, creates an efficient grid structure that captures the intrinsic

structure of the dataset. This improves the representation and separation of different data regions, resulting in better classification accuracy.

Accurate and Interpretable Fuzzy Rule Generation: The generated fuzzy rule base provides accurate classifications by capturing relevant patterns and relationships in the dataset. The rules are interpretable and understandable, allowing domain experts to gain insights into the classification process and make informed decisions.

Balance between Accuracy and Interpretability: The proposed approach strikes a balance between accuracy and interpretability, addressing the trade-off often encountered in classification tasks. The hybrid grid partitioning captures fine-grained data regions for improved accuracy, while the rough set-based feature selection and rule generation ensure interpretability, facilitating human-understandable decision-making.

Applicability in Various Domains: The findings demonstrate the applicability of the approach across different domains, such as healthcare, finance, marketing, and manufacturing. The synergistic effects of hybrid grid partitioning and the rough set method can be leveraged to enhance dataset classification in various real-world scenarios.

Comparative Performance: Comparative analysis with existing classification methods, including traditional machine learning algorithms, can provide insights into the superiority of the proposed approach. Evaluation metrics such as accuracy, precision, recall, and F1-score can be used to assess the performance improvement achieved through the integration of hybrid grid partitioning and the rough set method.

The scientific contribution of this research on the synergistic effects of hybrid grid partitioning and the rough set method for fuzzy rule generation in dataset classification lies in the following aspects:

Novel Integration of Techniques: The research proposes a novel integration of hybrid grid partitioning and the rough set method within the context of fuzzy rule generation for dataset classification. This integration harnesses the strengths of both techniques to enhance feature selection, grid partitioning, and rule generation, resulting in improved accuracy and interpretability of the fuzzy rule base.

Advancement in Feature Selection: The research contributes to the field of feature selection by introducing the rough set-based approach in the context of fuzzy rule generation. By considering the dependencies and redundancies among attributes, the rough set method helps identify the most relevant features, leading to a more focused and informative feature subset for classification.

Enhanced Grid Partitioning Technique: The research proposes the hybrid grid partitioning technique that combines clustering algorithms and grid-based approaches. This novel approach improves the representation and separation of different data regions, capturing the intrinsic structure of the dataset more effectively. It contributes to the advancement of grid partitioning methods in dataset classification.

Interpretable Fuzzy Rule Generation: The research contributes to the field of fuzzy rule generation by ensuring the interpretability of the generated rules. The rules derived from the selected features and the grid structure are interpretable and understandable, enabling domain experts to gain insights into the classification process and make informed decisions.

Trade-off between Accuracy and Interpretability: The research addresses the important trade-off between accuracy and interpretability in dataset classification. By striking a balance between these two factors, the proposed approach provides accurate classifications while maintaining the interpretability of the fuzzy rule base. This contribution is valuable as it allows for transparent decision-making and enhances the trustworthiness of the classification results.

Applicability in Various Domains: The research demonstrates the applicability of the proposed approach across different domains, highlighting its versatility and potential for solving

classification problems in various real-world scenarios. This contribution expands the understanding of the effectiveness of hybrid grid partitioning and the rough set method in dataset classification.

This research also has limitations as we describe below:

Generalizability: The research findings and proposed approach may have limitations in terms of generalizability. The effectiveness of the synergistic effects of hybrid grid partitioning and the rough set method may vary depending on the specific characteristics of the dataset, the complexity of the classification task, and the nature of the domain. The research should be validated and tested on a diverse range of datasets and domains to assess its robustness and generalizability.

Scalability: The scalability of the proposed approach may be a limitation. The computational complexity of hybrid grid partitioning and the rough set method can increase with larger datasets and higher dimensions. The research should investigate the scalability of the approach and explore techniques to mitigate computational overhead, such as parallelization or optimization algorithms.

Sensitivity to Parameter Tuning: The proposed approach may be sensitive to parameter tuning. The performance and effectiveness of hybrid grid partitioning and the rough set method could be influenced by the selection of parameters, such as the number of cells in the grid structure or the thresholds used in the rough set-based feature selection. A thorough sensitivity analysis and parameter tuning process should be conducted to ensure optimal performance of the approach.

Lack of Comparative Analysis: While the research highlights the effectiveness of the proposed approach, a comprehensive comparative analysis with existing classification methods is essential to establish its superiority. Benchmarking against traditional machine learning algorithms and other state-of-the-art approaches would provide insights into the performance improvement achieved through the integration of hybrid grid partitioning and the rough set method. Comparative evaluations using multiple datasets and performance metrics would enhance the credibility of the research findings.

Interpretability Challenges in Complex Datasets: Although the proposed approach aims to maintain interpretability, it may face challenges in complex datasets with intricate patterns and relationships. The interpretability of fuzzy rule bases can become more challenging when dealing with high-dimensional or nonlinear datasets. Further research should explore techniques to enhance interpretability in complex datasets, such as rule pruning or post-processing methods.

Lack of Real-World Case Studies: The research could benefit from real-world case studies to evaluate the practical applicability and performance of the proposed approach. While the findings demonstrate the effectiveness of the approach in theory, real-world scenarios may present additional challenges and complexities that need to be addressed. Case studies involving real datasets and domain experts' feedback would provide valuable insights into the strengths and limitations of the approach in practical applications.

Conclusion

In conclusion, this research has investigated the synergistic effects of hybrid grid partitioning and the rough set method for fuzzy rule generation in dataset classification. The findings contribute to the field by proposing an integrated approach that enhances feature selection, grid partitioning, and rule generation, leading to improved accuracy and interpretability of the fuzzy rule base. The research demonstrates that the rough set-based feature selection technique effectively identifies the most relevant features for classification, allowing for a focused and informative feature subset. The hybrid grid partitioning technique captures the intrinsic structure of the dataset, leading to better representation and separation of data regions, which improves classification accuracy. The generated fuzzy rule base is interpretable, enabling domain experts to gain insights into the classification process and make informed decisions. However, it is important to consider the

limitations of this research. Generalizability may be a limitation, as the effectiveness of the proposed approach could vary depending on the dataset characteristics and problem domain. Scalability and sensitivity to parameter tuning should be further investigated to ensure the approach can handle larger datasets and achieve optimal performance. Comparative analysis against existing methods and real-world case studies would provide a more comprehensive evaluation of the proposed approach. Nonetheless, this research contributes to the advancement of dataset classification by providing a novel integration of hybrid grid partitioning and the rough set method. The findings highlight the trade-off between accuracy and interpretability and demonstrate the applicability of the approach across various domains. Researchers and practitioners can leverage the proposed approach to improve classification accuracy and gain insights from interpretable fuzzy rule bases. Future research directions could focus on addressing the limitations mentioned, including investigating the scalability and robustness of the approach, conducting comparative evaluations, and exploring techniques to enhance interpretability in complex datasets. By addressing these aspects, the proposed approach can be further refined and extended, contributing to the field of dataset classification and supporting decision-making processes in various domains.

Reference

- Alhenawi, E., Al-Sayyed, R., Hudaib, A., & Mirjalili, S. (2022). Feature selection methods on gene expression microarray data for cancer classification: A systematic review. *Computers in Biology and Medicine*, 140, 105051.
- Ali, A., Heneash, U., Hussein, A., & Eskebi, M. (2022). Predicting Pavement Condition Index Using Fuzzy Logic Technique. *Infrastructures*, 7(7), 91.
- Asgarnezhad, R., Majeed, S. S. A., Abbas, Z. A., & Salman, S. S. (2022). An Effective Algorithm to Improve Recommender Systems using Evolutionary Computation Algorithms and Neural Network: Using Evolutionary Computation Algorithms and Neural Networks, an Effective Algorithm to Improve Recommender Systems. *Wasit Journal of Computer and Mathematics Sciences*, 1(1), 27–35.
- Buabeng, A., Simons, A., Frempong, N. K., & Ziggah, Y. Y. (2022). Predictive Maintenance Model Based on Multisensor Data Fusion of Hybrid Fuzzy Rough Set Theory Feature Selection and Stacked Ensemble for Fault Classification. *Mathematical Problems in Engineering*, 2022.
- Chen, Y.-C., Pal, N. R., & Chung, I.-F. (2011). An integrated mechanism for feature selection and fuzzy rule extraction for classification. *IEEE Transactions on Fuzzy Systems*, 20(4), 683–698.
- Cheruku, R., Edla, D. R., Kuppili, V., & Dharavath, R. (2018). Rst-batminer: A fuzzy rule miner integrating rough set feature selection and bat optimization for detection of diabetes disease. *Applied Soft Computing*, 67, 764–780.
- Citakoglu, H., & Demir, V. (2023). Developing numerical equality to regional intensity–duration–frequency curves using evolutionary algorithms and multi-gene genetic programming. *Acta Geophysica*, 71(1), 469–488.
- Criado-Ramón, D., Ruiz, L. G. B., & Pegalajar, M. C. (2023). An Improved Pattern Sequence-Based Energy Load Forecast Algorithm Based on Self-Organizing Maps and Artificial Neural Networks. *Big Data and Cognitive Computing*, 7(2), 92.
- Das, A. K., Sengupta, S., & Bhattacharyya, S. (2018). A group incremental feature selection for classification using rough set theory based genetic algorithm. *Applied Soft Computing*, 65, 400–411.
- Dhal, P., & Azad, C. (2022). A comprehensive survey on feature selection in the various fields of machine learning. *Applied Intelligence*, 1–39.
- Fisher, B., Turner, R. K., & Morling, P. (2009). Defining and classifying ecosystem services for decision making. *Ecological Economics*, 68(3), 643–653.
- García-García, F., Corral, A., Iribarne, L., & Vassilakopoulos, M. (2020). Improving distance-join query processing with voronoi-diagram based partitioning in spatialhadoop. *Future Generation Computer Systems*, 111, 723–740.
- García-García, F., Corral, A., Iribarne, L., & Vassilakopoulos, M. (2018). Voronoi-diagram based partitioning for distance join query processing in spatialhadoop. *Model and Data Engineering: 8th International Conference, MEDI 2018, Marrakesh, Morocco, October 24–26, 2018, Proceedings 8*, 251–267.

- Gómez-Skarmeta, A. F., Valdés, M., Jiménez, F., & Marín-Blázquez, J. G. (2001). Approximative fuzzy rules approaches for classification with hybrid-GA techniques. *Information Sciences*, 136(1–4), 193–214.
- Guo, L., Daly, E. M., Alkan, O., Mattetti, M., Cornec, O., & Knijnenburg, B. (2022). Building trust in interactive machine learning via user contributed interpretable rules. *27th International Conference on Intelligent User Interfaces*, 537–548.
- Hostetter, J. W., Abdelshiheed, M., Barnes, T., & Chi, M. (2023). A self-organizing neuro-fuzzy q-network: Systematic design with offline hybrid learning. *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, 1248–1257.
- Idris, N. F., Ismail, M. A., Mohamad, M. S., Kasim, S., Zakaria, Z., & Sutikno, T. (2022). Breast cancer disease classification using fuzzy-ID3 algorithm based on association function. *IAES International Journal of Artificial Intelligence*, 11(2), 448.
- Ikotun, A. M., & Ezugwu, A. E. (2022). Boosting k-means clustering with symbiotic organisms search for automatic clustering problems. *Plos One*, 17(8), e0272861.
- Jara, L., Ariza-Valderrama, R., Fernández-Olivares, J., González, A., & Pérez, R. (2022). Efficient inference models for classification problems with a high number of fuzzy rules. *Applied Soft Computing*, 115, 108164.
- Jensen, R., & Shen, Q. (2007). Fuzzy-rough sets assisted attribute selection. *IEEE Transactions on Fuzzy Systems*, 15(1), 73–89.
- Jensen, R., & Shen, Q. (2008). *Computational intelligence and feature selection: rough and fuzzy approaches*.
- Jiang, Q., Zhou, X., Wang, R., Ding, W., Chu, Y., Tang, S., Jia, X., & Xu, X. (2022). Intelligent monitoring for infectious diseases with fuzzy systems and edge computing: A survey. *Applied Soft Computing*, 108835.
- Joshi, C., Ranjan, R. K., & Bharti, V. (2022). A Fuzzy Logic based feature engineering approach for Botnet detection using ANN. *Journal of King Saud University-Computer and Information Sciences*, 34(9), 6872–6882.
- Khan, F., Siva Prasad, B. V. V., Syed, S. A., Ashraf, I., & Ramasamy, L. K. (2022). An efficient, ensemble-based classification framework for big medical data. *Big Data*, 10(2), 151–160.
- Lyngdoh, G. A., Zaki, M., Krishnan, N. M. A., & Das, S. (2022). Prediction of concrete strengths enabled by missing data imputation and interpretable machine learning. *Cement and Concrete Composites*, 128, 104414.
- Moshayedi, A. J., Roy, A. S., Kolahdooz, A., & Shuxin, Y. (2022). Deep Learning Application Pros And Cons Over Algorithm Deep Learning Application Pros And Cons Over Algorithm. *EAI Endorsed Transactions on AI and Robotics*, 1(1).
- Nilashi, M., Abumalloh, R. A., Alrizq, M., Almulihi, A., Alghamdi, O. A., Farooque, M., Samad, S., Mohd, S., & Ahmadi, H. (2022). A Hybrid Method to Solve Data Sparsity in Travel Recommendation Agents Using Fuzzy Logic Approach. *Mathematical Problems in Engineering*, 2022.
- Nivetha, S., & Inbarani, H. H. (2023). Novel Architecture for Image Classification Based on Rough Set. *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)*, 14(1), 1–38.
- Pattaraintakorn, P., & Cercone, N. (2008). Integrating rough set theory and medical applications. *Applied Mathematics Letters*, 21(4), 400–403.
- Priya, S., Selvakumar, S., & Velusamy, R. L. (2022). PaSOFuAC: Particle Swarm Optimization Based Fuzzy Associative Classifier for Detecting Phishing Websites. *Wireless Personal Communications*, 125(1), 755–784.
- Rana, P., Rai, S. K., & Shrivastav, K. (n.d.). *A Comparative analysis on Fuzzy systems-based Management Models*.
- Salamo, M., & Lopez-Sanchez, M. (2011). Rough set based approaches to feature selection for case-based reasoning classifiers. *Pattern Recognition Letters*, 32(2), 280–292.
- Sanz, J. A., Bernardo, D., Herrera, F., Bustince, H., & Hagrass, H. (2014). A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data. *IEEE Transactions on Fuzzy Systems*, 23(4), 973–990.
- Schwaller, P., Vaucher, A. C., Laplaza, R., Bunne, C., Krause, A., Corminboeuf, C., & Laino, T. (2022). Machine intelligence for chemical reaction space. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 12(5), e1604.
- Shapiro, A. F. (2002). The merging of neural networks, fuzzy logic, and genetic algorithms. *Insurance: Mathematics and Economics*, 31(1), 115–131.
- Shelke, N., Chaudhury, S., Chakrabarti, S., Bangare, S. L., Yogapriya, G., & Pandey, P. (2022). An efficient way

- of text-based emotion analysis from social media using LRA-DNN. *Neuroscience Informatics*, 100048.
- Subhashini, L., Li, Y., Zhang, J., & Atukorale, A. S. (2022). Integration of fuzzy logic and a convolutional neural network in three-way decision-making. *Expert Systems with Applications*, 202, 117103.
- Tiwari, S., & Agarwal, S. (2023). Empirical analysis of chronic disease dataset for multiclass classification using optimal feature selection based hybrid model with spark streaming. *Future Generation Computer Systems*, 139, 87–99.
- Turner, S. D., Dudek, S. M., & Ritchie, M. D. (2010). ATHENA: A knowledge-based hybrid backpropagation-grammatical evolution neural network algorithm for discovering epistasis among quantitative trait Loci. *BioData Mining*, 3, 1–18.
- Udhaya Kumar, S., & Hannah Inbarani, H. (2017). PSO-based feature selection and neighborhood rough set-based classification for BCI multiclass motor imagery task. *Neural Computing and Applications*, 28, 3239–3258.
- Varshney, A. K., & Torra, V. (2023). Literature Review of the Recent Trends and Applications in Various Fuzzy Rule-Based Systems. *International Journal of Fuzzy Systems*, 1–24.
- Wan, S., Zhao, Y., Wang, T., Gu, Z., Abbasi, Q. H., & Choo, K.-K. R. (2019). Multi-dimensional data indexing and range query processing via Voronoi diagram for internet of things. *Future Generation Computer Systems*, 91, 382–391.
- Wang, G.-Y., Yao, Y.-Y., & Yu, H. (2009). A survey on rough set theory and applications. *Chinese Journal of Computers*, 32(7), 1229–1246.
- Xue, G., Chang, Q., Wang, J., Zhang, K., & Pal, N. R. (2022). An adaptive neuro-fuzzy system with integrated feature selection and rule extraction for high-dimensional classification problems. *IEEE Transactions on Fuzzy Systems*.
- Yang, C., Cheng, T., Li, S., Gu, X., & Yang, L. (2023). A novel partitioning method for the power grid restoration considering the support of multiple LCC-HVDC systems. *Energy Reports*, 9, 1104–1112.
- Zhai, L.-Y., Khoo, L.-P., & Zhong, Z.-W. (2008). A rough set enhanced fuzzy approach to quality function deployment. *The International Journal of Advanced Manufacturing Technology*, 37, 613–624.
- Zhai, L.-Y., Khoo, L.-P., & Zhong, Z.-W. (2009). A rough set based QFD approach to the management of imprecise design information in product development. *Advanced Engineering Informatics*, 23(2), 222–228.
- Zhang, R., Yu, Y., Shetty, P., Song, L., & Zhang, C. (2022). Prboost: Prompt-based rule discovery and boosting for interactive weakly-supervised learning. *ArXiv Preprint ArXiv:2203.09735*.
- Zhao, T., Cao, H., & Dian, S. (2022). A self-organized method for a hierarchical fuzzy logic system based on a fuzzy autoencoder. *IEEE Transactions on Fuzzy Systems*, 30(12), 5104–5115.
- Zheng, B., Xu, J., Lee, W.-C., & Lee, D. L. (2006). Grid-partition index: a hybrid method for nearest-neighbor queries in wireless location-based services. *The VLDB Journal*, 15, 21–39.