VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v9i2.5015

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

A SYSTEMATIC LITERATURE REVIEW: RECURSIVE FEATURE ELIMINATION ALGORITHMS

Arif Mudi Priyatno[1*]; Triyanna Widiyaningtyas[2]

Department of Electrical and Informatics Engineering[1,2]
Universitas Negeri Malang[1,2]
https://um.ac.id[1,2]
arif.mudi.2305349@students.um.ac.id[1*], triyannaw.ft@um.ac.id[2]

(*) Corresponding Author

Abstract—*Recursive feature elimination (RFE) is a feature selection algorithm that works by gradually eliminating unimportant features. RFE has become a popular method for feature selection in various machine learning applications, such as classification and prediction. However, there is no systematic literature review (SLR) that discusses recursive feature elimination algorithms. This article conducts a SLR on RFE algorithms. The goal is to provide an overview of the current state of the RFE algorithm. This SLR uses IEEE Xplore, ScienceDirect, Springer, and Scopus (publish and publish) databases from 2018 to 2023. This SLR received 76 relevant papers with 49% standard RFEs, 43% strategy RFEs, and 8% modified RFEs. Research using RFE continues to increase every year, from 2018 to 2023. The feature selection method used simultaneously or for comparison is based on a filter approach, namely Pearson correlation, and an embedded approach, namely random forest. The most widely used machine learning algorithms are support vector machines and random forests, with 19.5% and 16.7%, respectively. Strategy RFE and modified RFE can be referred to as hybrid RFEs. Based on relevant papers, it is found that the RFE strategy is broadly divided into two categories: using RFE after other feature selection methods and using RFE simultaneously with other methods. Modification of the RFE is done by modifying the flow of the RFE. The modification process is divided into two categories: before the process of calculating the smallest weight criteria and after calculating the smallest weight criteria. Calculating the smallest weight criteria in this RFE modification is still a challenge at this time to obtain optimal results.*

Keywords: *hybrid, modified, strategy, recursive feature elimination, systematic literature review .*

Intisari—*Recursive feature elimination (RFE) adalah algoritma seleksi fitur yang bekerja dengan menyingkirkan fitur yang tidak penting secara bertahap. RFE telah menjadi metode yang populer untuk seleksi fitur dalam berbagai aplikasi machine learning, seperti klasifikasi, dan prediksi. Tetapi belum ada systematic literature review yang membahas tentang algoritma recursive feature elimination. Artikel ini melakukan systematic literature review (SLR) pada algoritme recursive feature elimination. Tujuannya adalah untuk memberikan gambaran umum tentang kondisi terkini algoritme RFE. SLR ini menggunakan database IEEE Xplore, ScienceDirect, Springer, dan Scopus (publish and perish) dari tahun 2018 hingga 2023. SLR ini mendapatkan 76 paper relevan dengan 49% RFE standar, 43% RFE strategi, dan 8% modifikasi RFE. Penelitian menggunakan RFE selalu meningkat setiap tahunnya dari 2018 hingga 2023. Metode seleksi fitur yang digunakan bersamaan atau pembanding berdasarkan pendeketan filter yaitu pearson correlation, serta berdasarkan pendekatan embedded yaitu random forest. Machine learning yang paling banyak digunakan yaitu support vector machine dan random forest dengan masing-masing 19,5% dan 16,7%. RFE strategi, dan RFE modifikasi dapat disebut sebagai Hybrid RFE. Berdasarkan paper relvan didapatkan bahwa strategi RFE secara garis besar dibedakan menjadi 2 yaitu penggunaan RFE setelah proses metode seleksi fitur lainnya, dan penggunaan RFE secara bersamaan dengan metode lainnya. Modifkasi RFE dilakukan dengan memodifikasi alur dari RFE. Proses modifikasi terbagi menjadi 2 yaitu sebelum proses perhitungan kriteria bobot terkecil, dan sesudah perhitungan kriteria bobot terkecil. Perhitungan kriteria bobot terkecil pada modifikasi RFE ini masih menjadi tantangan saat ini untuk mendapatkan hasil yang optimal.*

Kata Kunci: *hybrid, modifikasi, strategi, recursive feature elimination, systematic literature review.*

VOL. 9. NO. 2 FEBRUARY 2023
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5015

INTRODUCTION

Feature selection (FS) is a technique used to identify the optimal set of features that effectively captures the essence of the data set. It involves selecting the feature that contributes most to the estimated variable in a specific domain as user-interested [1]. Data development has undergone a significant improvement in recent years, mainly due to the development of data collection methodologies in various disciplines. As a result, this surge in data volumes requires wider utilization of computing resources and increased time requirements for implementing machine learning systems. In different sectors, the data collected often indicates a high level of dimension, so the role of machine learning may just be to select an optimal set of features and remove excess features manually.

The machine learning model applied showed limited learning capabilities due to the presence of inappropriate features in the data set, resulting in low recognition levels and a severe decline in performance. The removal of excessive and outdated features by FS resulted in reduced dimensions and improved vector quality of the attributes produced [2]–[4]. Feature selection (FS) has been used for many applications, such as breast cancer and diabetes classification [5], speech recognition [6], gene prediction [7], gait analysis [8], text mining [9], and others.

Feature selection (FS) is characterized by two fundamental objectives: reducing required features and maximizing classification performance to overcome the challenges posed by the dimension curse. There are three main types of FS strategies, namely, filter, wrapper, and embedded techniques, which involve the integration of filters and wrappers [10], [11]. The use of a filter approach does not depend on a specific machine learning algorithm. The exploitation of low-performance computing capabilities is usually suitable for data sets with fewer features. Filtering techniques in machine learning usually ignore the relationship between classification and characteristics, resulting in failure to detect samples accurately during the learning process.

Many studies have used wrappers as a way to address this problem. The use of wrapper strategies often involves the modification of the training process and the use of classification as an evaluation mechanism [12]. Therefore, the wrapper's strategy for feature selection often affects the training algorithm and provides more accurate results than the filter. Wrapper concentrates its efforts on training the machine learning algorithm by utilizing only a small subset of features that are crucial to assessing the effectiveness of the training model. The wrapper algorithm takes into account the selection accuracy specified at each previous phase to choose whether to add or remove features from the selected feature set. As a result, wrapper methods tend to show higher computational complexity and cost compared to most filtering techniques.

One of the wrapper algorithms is recursive feature elimination (RFE). RFE has done a lot in various fields, either by doing a combination with other features or making modifications of the RFE.

This article presents a systematic literature review (SLR) that focuses on the Recursive Feature Elimination (RFE) algorithm. The aim of this study is to present a comprehensive review of the RFE algorithm, which includes leading academic journals that publish large amounts of RFE-related research. In addition, this review will explore the use of RFE in data mining techniques as well as its integration with machine learning or deep learning. Furthermore, various variations of feature selection techniques used in conjunction with RFE will be examined, along with modifications and integration strategies used to improve feature selection.

MATERIALS AND METHODS

This review is carried out in accordance with the methodology of systematic literature review (SLR) [13], which covers the planning, implementation, and reporting stages. The review methodology covers six main stages, namely: formulation of research questions, development of search strategies, study selection, data extraction, evaluation of study quality, and data synthesis.

A. Research Questions

The purpose of this systematic literature review (SLR) is to provide a brief and comprehensive overview of empirical evidence related to the recursive feature elimination feature selection algorithm. To achieve this goal, a series of five research questions (RQ) is formulated as follows:

1. RQ 1: How many articles have used the recursive feature elimination method, which has emerged in the last few years?
2. RQ 2: Were the data mining techniques chosen by the researchers in using feature selection, in particular recursive feature elimination?
3. RQ 3: What type of feature selection technique is applied in conjunction with recursive feature elimination?
4. RQ 4: What machine or deep learning method is used after recursive feature elimination?
5. RQ 5: How does architecture evolve or modify Recursive feature elimination?

B. Search Strategy

A search strategy has several components, including search terms, literature resources, and the search process. Each of these components will be described separately in the following section.

Search Terms

The search string was formulated using the PICOC criteria (population, intervention, comparison, outcomes, and context) proposed by Kitchenham [13]–[15]. The entire search phrase resulting is feature selection AND (recursive feature elimination OR RFE) AND (predict OR classification OR forecast) AND (important OR relevant OR redundant).

Literature Resources

This investigation has used well-known resources to conduct a comprehensive search for related works to provide input to this review. This database includes IEEE Explore, ScienceDirect, Springer Link, and Scopus (publish and perish). Table 1 is the link and name of the online database.

Table 1. Name and Link Database Resource

Name	Link the database source
IEEE Explore	https://ieeexplore.ieee.org
ScienceDirect	https://www.sciencedirect.com
Springer Link	https://link.springer.com
Scopus (publish and perish)	https://harzing.com/resources/publish-or-perish

Search Process

The process of conducting a systematic literature review (SLR) requires a thorough and in-depth exploration of all relevant resources. Previously formulated search keywords are used to search for scientific articles in four electronic databases. Search phrases are modified to fit different databases because the search engines inside each database use different search string syntax.

C. Study Selection

The research paper for the Systematic Literature Review (SLR) is taken from various sources using the search phrases specified in Section B. The study limits the scope of the search to the period between 2018 and 2023. A total of 1014 items were identified.

In the process of conducting a systematic literature review, researchers use inclusion and exclusion criteria to filter the selection of research papers taken through a database search. These criteria are intended to identify the most relevant studies or studies to be included in the review.

The criteria for inclusion in this SLR are as follows: If there are several publications related to the same research, only more comprehensive and up-to-date articles will be included. Scientific articles or reviews that have been published in renowned scientific journals. The article provides an analysis of feature selection techniques, with a special focus on recursive feature elimination. This SLR focuses on articles published over the period from 2018 to 2023. The article was published in a reputable journal with a scope in engineering or computer science. The article has been published in the journal, categorized as quarterly Q1 or Q2.

The exclusion criteria used in this systematic literature review are as follows: duplicate articles, the article does not use English, full article not available, and sources of academic literature such as conference papers, books, and book chapters.

Out of a total of 1014 primary articles, the study selection method produced 483 articles that were considered relevant and included in the analysis based on titles and abstracts. Relevant to computer or engineering topics are a total of 214 articles, and based on his publication in a reputable journal that is between Quarters 1 and 2, there are 186 articles. Furthermore, we proceeded to obtain the studies above for a comprehensive examination of the overall content, which ultimately culminated in 76 selected relevant articles.

D. Data Extraction

With data extraction, we use the selected research to gather data that contributes to answering the research questions discussed in this review. Guide of data extraction, namely resource library, published year, article title, journal name, quartile, problem, dataset, data mining techniques, feature selection approach, machine learning-related, evaluation performance, and design architecture research.

E. Quality Assessment

The basic purpose of performing quality assessment (QA) on the selected study is to give weight to the quantitative data obtained [16]. This QA technique is considered very important. A series of quality evaluation questions were developed to evaluate the rigor, credibility, and relevance of related research. To ensure the reliability of the findings of this survey, we only conduct in-depth discussions on relevant studies of acceptable quality.

F. Data Synthesis

The primary purpose of data synthesis is to combine evidence collected from research selected to answer research questions. Although one piece of

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

evidence may have limited proof strength, the cumulative effect of some pieces of evidence can significantly enhance the persuasiveness of an argument [16], [17]. Various methodologies are used to combine data collected in relation to research questions (RQ). The narrative synthesis approach is used to analyze data related to research issues. This means that the data is structured consistently in relation to those questions.

G. Threats to Validity

The purpose of this systematic literature review is to ensure the evolution of RFE modifications or integration of the RFE with other feature selection techniques. This review could not identify bias indications in the study selection process. The search does not involve a manual inspection of the title and abstracts of the entire paper loaded in the journal. This SLR does not use conference articles, chapter books, or books to reduce excessive effort. So in this SLR, there may be some papers that are not well filtered.

RESULTS AND DISCUSSION

We have identified 76 papers that are most relevant to the recursive feature elimination (RFE) method, whether modified, strategy of use, or described the method in full. Out of 76 selected papers, the number of articles dealing with the modification of RFE is 6; RFE use strategy is 33; and RFE standard is 37. The list of these papers can be seen in Table 2.

Table 2. Selected Paper Discourse

No	Discourse	Paper
1	Modified	[18]–[23]
2	Strategy	[24]–[56]
3	Standard	[57]–[93]

RQ 1: A number of papers using the recursive feature elimination method, which have been published in 2018–2023, have been published in the journal.

Figure 1. Number of RFE papers without duplicates

Search results using four databases and keywords that have been specified obtained 1041 papers in the journal. Of the 1041 papers, the papers corresponding to the title and abstract about recursive feature elimination did not duplicate 483 papers. The spread of this result can be seen in Figure 1. Figure 1 shows Scopus (Publish and Perish) 165 paper, IEEE Explore 14 paper, ScienceDirect 290 paper, and Springer 14 paper. This suggests that there are still many researchers interested in doing research with or about recursive feature elimination.

Figure 2. Number of Publications 2018–2023

Based on the year of his search, Figure 2 shows the trend of publications related to recursive feature elimination increasing every year. Searches for journals from 2018 to 2022 were carried out from January to December, while searches for journals from 2023 were carried out from January to September. This suggests that research with or about recursive feature elimination is still developing, and there is still a lot to learn about this method.

Figure 3. RFE Publication Place Trends

Such publication trends are published in various reputable international journals. Figure 3 shows the spread of publications in the journal. It shows that recursive feature elimination is applied in various fields of study, including computer science, health, remote sensing, and agriculture. Meanwhile, data about selected articles can be seen in Table 2.

RQ 2: Data mining techniques selected by the researchers include recursive feature elimination.

Based on the paper selected in Table 2, data mining techniques using recursive feature elimination feature selection include two, namely classification and prediction. Figure 4 shows a higher percentage of classification than predicted. The paper that deals with RFE with classification techniques is 58 papers (76%), while the prediction is 18 papers (24%).

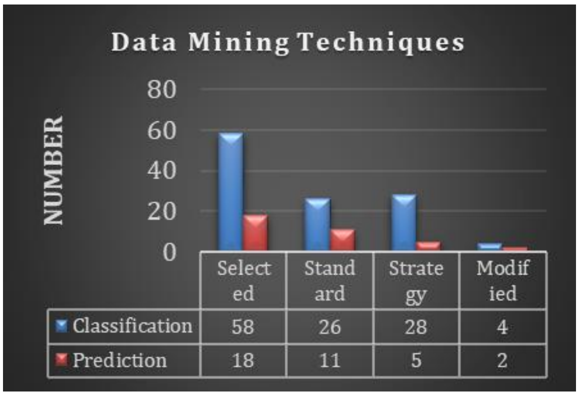

Figure 4. Data Mining Techniques Use RFE

RQ 3: Type of feature selection technique applied jointly, either hybrid or comparison with recursive feature elimination

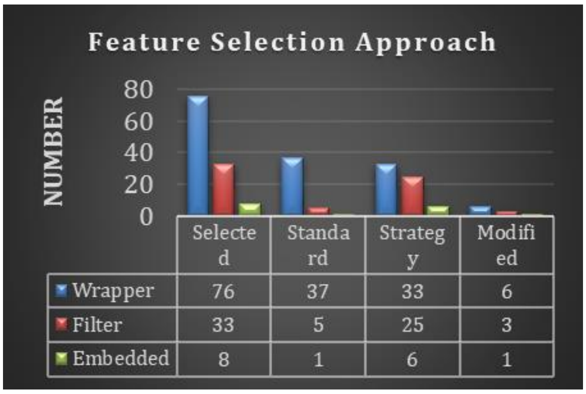

Figure 5. Selected Paper Feature Selection Approach

Figure 5 shows the number of peppers selected against the selection category of features it uses. The number of wrapper categories is equal to the number of papers; this is due to the main keyword, or SLR, being recursive feature elimination, which is part of the wrappers. The number of filter categories took second place because the filter method was able to perform hybrid processes with RFE either before RFE processes like paper [51], or when the RFE modeling process was like paper [23]. The number embedded took the last position because this category is usually used only as a comparison by the RFE method, as in the paper [50].

The filter approach involves direct estimation of feature scores using certain criteria [94]. However, this method does not take into account the impact of the selected feature subset on the performance of the next algorithm. This strategy demonstrates computing efficiency [95]. However, due to the absence of instructions from the learning algorithm, the selected subset may not be ideal. On the selected paper, a feature selection method with a filter approach is obtained from 20 methods. The 20 methods are listed below: ReliefF, ANOVA, pairwise correlations, Spearman correlates, variance thresholds, chi-square, T-set, conditional mutual information maximization (CMIM), fast correlations, and Pearson correlation Fisher's scores, generic univariate select, Hoeffding correlated, impedance correlated, Mann-Whitney U test, Shapiro-Wilk, information gain (IG), and Kappa index.

The wrapper approach involves treating the predictor as a black box, in which the feature selection process relies on a specific learning algorithm to assess the candidate subset [42]. However, this wrapper approach is prone to overfitting because the algorithm tends to study training data overly, which impedes its ability to generalize. The selected articles show that there are 21 feature selection methods with the wrapper approach used in the selected papers. These methods all use the basis of recursive feature elimination (RFE). This is because of these SLRs and keywords about RFE used to find selected papers.

The RF-RFE and SVM-RFE methods are the two most widely used methods on selected paper. The SVM-RFE method became the most widely used as it was the most first method proposed. This method is also relatively simple and easy to apply. The RF-RFE method became the second most widely used method because it had a higher success rate than the SVM-RFE method in some cases, such as in papers [70], [81], [96].

The RFE method still has a great opportunity to do research, especially with the modification of its anchoring method. This is because there are still ways of anchoring that can be done, such as a hybrid of the existing RFE ranking method or with other feature selection methods.

The embedded method is a type of feature selection technique that combines the characteristics of filter and wrapper techniques [62]. This approach combines feature selection as an integrated component of the training process. This allows the integration of algorithm modeling and feature selection simultaneously. Therefore, it can be said that embedded methods show superior computational efficiency and are less susceptible to overfitting when compared with wrapper approaches. Embedded and wrapper methods are two types of subset evaluation approaches that capture dependencies and interactions between features. The above-mentioned capabilities make this method more profitable than the filter method.

The embedded approach is used as a comparison of the RFE method in the selected

VOL. 9. NO. 2 FEBRUARY 2023
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5015

papers of 8 feature selection methods. The most common comparison method of the embedded approach is random forest. (RF). RF is used as a standard RFE comparator, strategy, or modification. This is due to the important features of RF reliability and efficiency [96]. Thus, the RF method can be used as a comparison of the RFE method to provide information about the advantages of the method.

RQ 4: Machine or deep learning method used after recursive feature elimination.

One important step in ML/DL usage is feature selection. Feature selection is the most relevant feature selection process for use in machine learning or deep learning models. Based on the analysis of selected papers, there are 20 ML/DL methods used in selected articles.

The top 10 methods used in the selected paper are machine learning methods. It shows that the process of classification and prediction using machine learning methods is greatly assisted by the existence of feature selection processes. The feature selection process can improve the accuracy of the classification or prediction, as only the relevant features are used in the process. The most commonly used machine learning methods are the support vector machine (SVM) and the random forest (RF). The SVM method is used on 19.5% of the selected paper, while RF is used in 16.7% of the selected paper.

RQ 5: Modification and development strategy Architecture Recursive feature elimination

In the discussion of the selected paper, the selection of RFE features is used in a variety of fields, namely business, computer science, health, environment, engineering, and biology. The field of health is the most frequently used RFE, which is 40 paper, or 51% of the total paper.

The selection of standard RFE features is broadly based on the input of datasets, training with classification and prediction methods, counting the criteria of all features, and deleting features with the smallest ranking [97]. The RFE standard has been applied to six similar fields. Standard RFE performed comparisons to other wrapper approaches [66], [80]; filter approaches [58], [64], [82]; and embedded approaches [57], [64]. In the selection of standard RFE features, the most widely used machine learning method is random forest, which is 20%. Other machine learning techniques that are also commonly used are support vector machines (16%), decision trees (11%), and K-nearest neighbor (9%).

RFE feature selection with strategy is the use of the RFE method combined with other feature selection methods, or what is called a hybrid method. The RFE feature selection process is carried out after the other feature selection processes are complete or is carried out simultaneously with other feature selection methods (ensemble feature selection).

RFE modification is a direct modification of the RFE method in its stages. Modifications architectural are made by adding other feature selection methods at the RFE stage and calculating the latest ranking. The list of feature selection methods used in RFE modification includes RF-RFE [18], [19], [22]; XGB-RFE [19]; SVM-RFE [18], [20], [21], [23]; GBM-RFE [18]; Absolute Cosine [23]; KPCA [21]; dan MI [20]. To combine or figure out ranking criteria between them, you can add weights with a threshold [19]; add weights simply [18], [22]; add weights based on multiplying weights and accuracy [18]; sum with the mRMR method [23]; take the weighted average [21]; and recalculation of the smallest features and input features using the MICBC approach [20]. This provides the opportunity that there are other ways to calculate feature combinations, such as merid values based on correlation, averages based on maximum weight, and so on.

Challenges and Opportunities

The challenge of feature selection with recursive feature elimination can be categorized into three aspects. Firstly, RFE has the potential to delete characteristics that are crucial for categorization. RFE eliminates characteristics based on their influence on classification accuracy. Nevertheless, many characteristics that hold significance for classification might not exert a substantial impact on the accuracy of classification during the initial training phase. Consequently, RFE may exclude these variables, despite their significance for classification.

Additionally, Recursive Feature Elimination (RFE) might be excessively time-consuming when dealing with datasets that include a significant number of features. The RFE algorithm must assess all potential characteristics in order to choose the most effective subset of features. The evaluation of each feature necessitates the training of a classification model utilizing such feature. This process can be time-consuming, particularly when dealing with extensive feature databases.

Furthermore, the RFE has the potential to exhibit instability. The results of RFE may differ based on the sequence of provided features. This phenomenon is attributed to two causes, specifically the utilization of the supplied feature sequence by RFE to determine which characteristics should be eliminated. The arrangement of features can impact the categorization accuracy of the resultant model. The order in which features RFE are presented can cause variations in the outcomes.

The opportunities of applying feature selection through the recursive feature reduction approach may be categorized into three main aspects. The first aspect is the enhancement of classification accuracy. Recursive Feature Elimination (RFE) can be utilized to eliminate irrelevant characteristics, hence enhancing the accuracy of classification. This can be advantageous for various classification applications, including fraud detection, document classification, and medical diagnosis. Furthermore, enhancing the comprehensibility of the model. Recursive Feature Elimination (RFE) can be utilized to decrease the amount of features utilized in a model, thereby enhancing the interpretability of the model. This can be advantageous for situations where the capacity to understand and explain the model's reasoning is crucial, such as making business decisions and developing products. Enhance computational efficiency. Additionally, RFE can be applied to decrease the dimensions of the model, thereby enhancing computing efficacy. This is particularly valuable for applications that prioritize computing performance, such as implementing machine learning algorithms on mobile devices.

CONCLUSION

This research conducted a systematic literature review on the state of the Recursive Feature Elimination (RFE) method from 2018 to 2023. The number of papers obtained was 76 papers, of which 37 papers discussed standard RFE, 33 papers discussed strategies for using RFE, and 6 papers discussed RFE modifications. Research using the RFE feature selection method has increased every year. The RFE method is used in two data mining techniques, namely classification and prediction. Classification was 76 percent, and prediction was 24 percent. A lot of people use filter-based feature selection methods like Pearson correlation, minimum redundant maximum relevant (mRMR), and mutual information (MI) at the same time or for RFE comparisons. Meanwhile, based on the embedded approach, namely random forest, linear regression, and xgboost, The machine learning method most widely used after the feature selection process using RFE is support vector machine at 19.5%, random forest at 16.7%, decision tree at 9%, and k-nearest neighbor at 9%.

The RFE method is used as a hybrid-recursive feature elimination method with two approaches, namely strategy and modification of the RFE steps. RFE strategies are generally divided into two categories: the use of other feature selection before the RFE method and the use of other feature selection simultaneously with RFE. RFE modification, in general, is modifying features by combining feature selection in the RFE process. Modified RFE has two approaches, namely combining feature selection and removing the lowest feature weights, as well as reviewing the lowest feature weights using other feature selections.

This Systematic Literature Review (SLR) presents comprehensive information on recursive feature reduction, with a specific emphasis on algorithmic development. This SLR lacks comprehensive explanations addressing the comparison of recursive feature removal performance across different feature dimensions, namely small, medium, and high. The future objective of SLR in recursive feature elimination is to accurately ascertain the significance of small, medium, and high feature dimensions. Additionally, the SLR has the ability to thoroughly evaluate its performance across different datasets and apply equitable machine learning techniques.

REFERENCE

[1] J. Piri, P. Mohapatra, and R. Dey, "Fetal Health Status Classification Using MOGA - CD Based Feature Selection Approach," in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, IEEE, Jul. 2020, pp. 1–6. doi: 10.1109/CONECCT50063.2020.9198377.

[2] J. Piri *et al.*, "Feature Selection Using Artificial Gorilla Troop Optimization for Biomedical Data: A Case Analysis with COVID-19 Data," *Mathematics*, vol. 10, no. 15, p. 2742, Aug. 2022, doi: 10.3390/math10152742.

[3] J. Piri and P. Mohapatra, "Exploring Fetal Health Status Using an Association Based Classification Approach," in *2019 International Conference on Information Technology (ICIT)*, IEEE, Dec. 2019, pp. 166–171. doi: 10.1109/ICIT48102.2019.00036.

[4] T. Bhattacharyya, B. Chatterjee, P. K. Singh, J. H. Yoon, Z. W. Geem, and R. Sarkar, "Mayfly in Harmony: A New Hybrid Meta-Heuristic Feature Selection Algorithm," *IEEE Access*, vol. 8, pp. 195929–195945, 2020, doi: 10.1109/ACCESS.2020.3031718.

[5] E. Akkur, F. Türk, and O. Erogul, "Breast cancer classification using a novel hybrid feature selection approach," *Neural Netw. World*, vol. 33, no. 2, pp. 67–83, 2023, doi: 10.14311/NNW.2023.33.005.

[6] K. Jermsittiparsert *et al.*, "Pattern recognition and features selection for speech emotion recognition model using deep learning," *Int. J. Speech Technol.*, vol. 23,

VOL. 9. NO. 2 FEBRUARY 2023
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5015

no. 4, pp. 799–806, Dec. 2020, doi: 10.1007/s10772-020-09690-2.

[7] A. Naik, V. Kuppili, and D. Reddy Edla, "Binary Dragonfly Algorithm and Fisher Score Based Hybrid Feature Selection Adopting a Novel Fitness Function Applied to Microarray Data," in *2019 International Conference on Applied Machine Learning (ICAML)*, IEEE, May 2019, pp. 40–43. doi: 10.1109/ICAML48257.2019.00015.

[8] P. Schumann *et al.*, "Detection of Fall Risk in Multiple Sclerosis by Gait Analysis—An Innovative Approach Using Feature Selection Ensemble and Machine Learning Algorithms," *Brain Sci.*, vol. 12, no. 11, p. 1477, Oct. 2022, doi: 10.3390/brainsci12111477.

[9] J. T. Pintas, L. A. F. Fernandes, and A. C. B. Garcia, "Feature selection methods for text classification: a systematic literature review," *Artif. Intell. Rev.*, vol. 54, no. 8, pp. 6149–6200, 2021, doi: 10.1007/s10462-021-09970-6.

[10] M. Rostami, K. Berahmand, E. Nasiri, and S. Forouzandeh, "Review of swarm intelligence-based feature selection methods," *Eng. Appl. Artif. Intell.*, vol. 100, no. September 2020, p. 104210, Apr. 2021, doi: 10.1016/j.engappai.2021.104210.

[11] Q. Al-Tashi, S. J. Abdulkadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Approaches to Multi-Objective Feature Selection: A Systematic Literature Review," *IEEE Access*, vol. 8, pp. 125076–125096, 2020, doi: 10.1109/ACCESS.2020.3007291.

[12] B. Pes, "Ensemble feature selection for high-dimensional data: a stability analysis across multiple domains," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 5951–5973, May 2020, doi: 10.1007/s00521-019-04082-3.

[13] H. A. Mohamed Shaffril, S. F. Samsuddin, and A. Abu Samah, "The ABC of systematic literature review: the basic methodological guidance for beginners," *Qual. Quant.*, vol. 55, no. 4, pp. 1319–1346, 2021, doi: 10.1007/s11135-020-01059-6.

[14] Z. M. Zain, S. Sakri, and N. H. A. Ismail, "Application of Deep Learning in Software Defect Prediction: Systematic Literature Review and Meta-analysis," *Inf. Softw. Technol.*, vol. 158, no. January, p. 107175, Jun. 2023, doi: 10.1016/j.infsof.2023.107175.

[15] D. Bassi and H. Singh, "A Systematic Literature Review on Software Vulnerability Prediction Models," *IEEE Access*, vol. 11, no. October, pp. 110289–110311, 2023, doi: 10.1109/ACCESS.2023.3312613.

[16] J. Paul, W. M. Lim, A. O'Cass, A. W. Hao, and S. Bresciani, "Scientific procedures and rationales for systematic literature reviews (SPAR-4-SLR)," *Int. J. Consum. Stud.*, vol. 45, no. 4, pp. 1–16, Jul. 2021, doi: 10.1111/ijcs.12695.

[17] E. Tieppo, R. R. dos Santos, J. P. Barddal, and J. C. Nievola, *Hierarchical classification of data streams: a systematic literature review*, vol. 55, no. 4. Springer Netherlands, 2022. doi: 10.1007/s10462-021-10087-z.

[18] H. Jeon and S. Oh, "Hybrid-recursive feature elimination for efficient feature selection," *Appl. Sci.*, vol. 10, no. 9, p. 3211, May 2020, doi: 10.3390/app10093211.

[19] S. ur R. Khan *et al.*, "Enhanced Machine-Learning Techniques for Medium-Term and Short-Term Electric-Load Forecasting in Smart Grids," *Energies*, vol. 16, no. 1, p. 276, Dec. 2022, doi: 10.3390/en16010276.

[20] S. Xu, B. Lu, M. Baldea, T. F. Edgar, and M. Nixon, "An improved variable selection method for support vector regression in NIR spectral modeling," *J. Process Control*, vol. 67, pp. 83–93, Jul. 2018, doi: 10.1016/j.jprocont.2017.06.001.

[21] H. Sanz, C. Valim, E. Vegas, J. M. Oller, and F. Reverter, "SVM-RFE: selection and visualization of the most relevant features through non-linear kernels," *BMC Bioinformatics*, vol. 19, no. 1, p. 432, Dec. 2018, doi: 10.1186/s12859-018-2451-4.

[22] R. Su, X. Liu, and L. Wei, "MinE-RFE: determine the optimal subset from RFE by minimizing the subset-accuracy–defined energy," *Brief. Bioinform.*, vol. 21, no. 2, pp. 687–698, Mar. 2020, doi: 10.1093/bib/bbz021.

[23] S. Sahran, D. Albashish, A. Abdullah, N. A. Shukor, and S. Hayati Md Pauzi, "Absolute cosine-based SVM-RFE feature selection method for prostate histopathological grading," *Artif. Intell. Med.*, vol. 87, pp. 78–90, May 2018, doi: 10.1016/j.artmed.2018.04.002.

[24] F. Ali *et al.*, "DBPPred-PDSD: Machine learning approach for prediction of DNA-binding proteins using Discrete Wavelet Transform and optimized integrated features space," *Chemom. Intell. Lab. Syst.*, vol. 182, no. August, pp. 21–30, Nov. 2018, doi: 10.1016/j.chemolab.2018.08.013.

[25] X. Gao and X. Liu, "A novel effective diagnosis model based on optimized least squares support machine for gene microarray," *Appl. Soft Comput.*, vol. 66, pp. 50–59, May 2018, doi: 10.1016/j.asoc.2018.02.009.

[26] E. Valla, S. Nõmm, K. Medijainen, P. Taba, and

A. Toomela, "Tremor-related feature engineering for machine learning based Parkinson's disease diagnostics," *Biomed. Signal Process. Control*, vol. 75, no. October 2021, 2022, doi: 10.1016/j.bspc.2022.103551.

[27] A. Thalor, H. Kumar Joon, G. Singh, S. Roy, and D. Gupta, "Machine learning assisted analysis of breast cancer gene expression profiles reveals novel potential prognostic biomarkers for triple-negative breast cancer," *Comput. Struct. Biotechnol. J.*, vol. 20, pp. 1618–1631, 2022, doi: 10.1016/j.csbj.2022.03.019.

[28] T. Gaber, A. El-Ghamry, and A. E. Hassanien, "Injection attack detection using machine learning for smart IoT applications," *Phys. Commun.*, vol. 52, no. September 2020, p. 101685, 2022, doi: 10.1016/j.phycom.2022.101685.

[29] L. Li, M. Wang, X. Jiang, and Y. Lin, "Universal multi-factor feature selection method for radiomics-based brain tumor classification," *Comput. Biol. Med.*, vol. 164, p. 107122, Sep. 2023, doi: 10.1016/j.compbiomed.2023.107122.

[30] Uzma, U. Manzoor, and Z. Halim, "Protein encoder: An autoencoder-based ensemble feature selection scheme to predict protein secondary structure," *Expert Syst. Appl.*, vol. 213, p. 119081, Mar. 2023, doi: 10.1016/j.eswa.2022.119081.

[31] Y. Zhang, H. Guo, Y. Zhou, C. Xu, and Y. Liao, "Recognising drivers' mental fatigue based on EEG multi-dimensional feature selection and fusion," *Biomed. Signal Process. Control*, vol. 79, no. P2, p. 104237, 2023, doi: 10.1016/j.bspc.2022.104237.

[32] M. G. Albayrak, E. Evin, O. Yiğit, M. Toğaçar, and B. Ergen, "Experimental and artificial intelligence approaches to measuring the wear behavior of DIN St28 steel boronized by the box boronizing method using a mechanically alloyed powder source," *Eng. Appl. Artif. Intell.*, vol. 120, no. January, p. 105910, Apr. 2023, doi: 10.1016/j.engappai.2023.105910.

[33] H. Qi, X. Song, S. Liu, Y. Zhang, and K. K. L. Wong, "KFPredict: An ensemble learning prediction framework for diabetes based on fusion of key features," *Comput. Methods Programs Biomed.*, vol. 231, p. 107378, Apr. 2023, doi: 10.1016/j.cmpb.2023.107378.

[34] R. Alzubi, N. Ramzan, H. Alzoubi, and A. Amira, "A Hybrid Feature Selection Method for Complex Diseases SNPs," *IEEE Access*, vol. 6, pp. 1292–1301, 2018, doi: 10.1109/ACCESS.2017.2778268.

[35] A. Moghimi, C. Yang, and P. M. Marchetto, "Ensemble Feature Selection for Plant Phenotyping: A Journey From Hyperspectral to Multispectral Imaging," *IEEE Access*, vol. 6, pp. 56870–56884, 2018, doi: 10.1109/ACCESS.2018.2872801.

[36] S. Byun *et al.*, "Detection of major depressive disorder from linear and nonlinear heart rate variability features during mental task protocol," *Comput. Biol. Med.*, vol. 112, no. August, p. 103381, Sep. 2019, doi: 10.1016/j.compbiomed.2019.103381.

[37] G. Fang, P. Xu, and W. Liu, "Automated Ischemic Stroke Subtyping Based on Machine Learning Approach," *IEEE Access*, vol. 8, pp. 118426–118432, 2020, doi: 10.1109/ACCESS.2020.3004977.

[38] S. A.-F. Sayed, A. M. Elkorany, and S. Sayed Mohammad, "Applying Different Machine Learning Techniques for Prediction of COVID-19 Severity," *IEEE Access*, vol. 9, pp. 135697–135707, 2021, doi: 10.1109/ACCESS.2021.3116067.

[39] A. Tousi and M. Lujan, "Comparative Analysis of Machine Learning Models for Performance Prediction of the SPEC Benchmarks," *IEEE Access*, vol. 10, pp. 11994–12011, 2022, doi: 10.1109/ACCESS.2022.3142240.

[40] N. Omar, H. Aly, and T. Little, "Optimized Feature Selection Based on a Least-Redundant and Highest-Relevant Framework for a Solar Irradiance Forecasting Model," *IEEE Access*, vol. 10, pp. 48643–48659, 2022, doi: 10.1109/ACCESS.2022.3171230.

[41] U. Ullah, A. G. O. Jurado, I. D. Gonzalez, and B. Garcia-Zapirain, "A Fully Connected Quantum Convolutional Neural Network for Classifying Ischemic Cardiopathy," *IEEE Access*, vol. 10, pp. 134592–134605, 2022, doi: 10.1109/ACCESS.2022.3232307.

[42] Doreswamy, M. K. Hooshmand, I. Gad, M. K. H. Doreswamy, and I. Gad, "Feature selection approach using ensemble learning for network anomaly detection," *CAAI Trans. Intell. Technol.*, vol. 5, no. 4, pp. 283–293, Dec. 2020, doi: 10.1049/trit.2020.0073.

[43] N. Ayub *et al.*, "Big Data Analytics for Short and Medium-Term Electricity Load Forecasting Using an AI Techniques Ensembler," *Energies*, vol. 13, no. 19, p. 5193, Oct. 2020, doi: 10.3390/en13195193.

[44] Z. Chen *et al.*, "Detecting Abnormal Brain Regions in Schizophrenia Using Structural MRI via Machine Learning," *Comput. Intell. Neurosci.*, vol. 2020, pp. 1–13, Mar. 2020, doi: 10.1155/2020/6405930.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 9. NO. 2 FEBRUARY 2023
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5015

[45] H. Lin, Y. Xue, K. Chen, S. Zhong, and L. Chen, "Acute coronary syndrome risk prediction based on gradient boosted tree feature selection and recursive feature elimination: A dataset-specific modeling study," *PLoS One*, vol. 17, no. 11 November, p. e0278217, Nov. 2022, doi: 10.1371/journal.pone.0278217.

[46] S. Bhattacharjee, D. Prakash, C.-H. Kim, H.-C. Kim, and H.-K. Choi, "Texture, Morphology, and Statistical Analysis to Differentiate Primary Brain Tumors on Two-Dimensional Magnetic Resonance Imaging Scans Using Artificial Intelligence Techniques," *Healthc. Inform. Res.*, vol. 28, no. 1, pp. 46–57, Jan. 2022, doi: 10.4258/hir.2022.28.1.46.

[47] M. S. Rahman, M. K. Rahman, S. Saha, M. Kaykobad, and M. S. Rahman, "Antigenic: An improved prediction model of protective antigens," *Artif. Intell. Med.*, vol. 94, no. May 2018, pp. 28–41, Mar. 2019, doi: 10.1016/j.artmed.2018.12.010.

[48] F. Y. Chin, K. H. Lem, and K. M. Wong, "Improving handwritten digit recognition using hybrid feature selection algorithm," *Appl. Comput. Informatics*, Jul. 2022, doi: 10.1108/ACI-02-2022-0054.

[49] D. Papathanasiou, K. Demertzis, and N. Tziritas, "Machine Failure Prediction Using Survival Analysis," *Futur. Internet*, vol. 15, no. 5, 2023, doi: 10.3390/fi15050153.

[50] J. Jemai and A. Zarrad, "Feature Selection Engineering for Credit Risk Assessment in Retail Banking," *Inf.*, vol. 14, no. 3, 2023, doi: 10.3390/info14030200.

[51] H. Castro, J. D. Garcia-Racines, and A. Bernal-Norena, "Methodology for the prediction of paroxysmal atrial fibrillation based on heart rate variability feature analysis," *Heliyon*, vol. 7, no. 11, p. e08244, Nov. 2021, doi: 10.1016/j.heliyon.2021.e08244.

[52] J. Cui, H. Xia, R. Zhang, B. Hu, and X. Cheng, "Optimization scheme for intrusion detection scheme GBDT in edge computing center," *Comput. Commun.*, vol. 168, no. December 2020, pp. 136–145, Feb. 2021, doi: 10.1016/j.comcom.2020.12.007.

[53] B. Fu *et al.*, "Comparison of optimized object-based RF-DT algorithm and SegNet algorithm for classifying Karst wetland vegetation communities using ultra-high spatial resolution UAV data," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 104, p. 102553, Dec. 2021, doi: 10.1016/j.jag.2021.102553.

[54] S. de Roda Husman, J. J. van der Sanden, S. Lhermitte, and M. A. Eleveld, "Integrating intensity and context for improved supervised river ice classification from dual-pol Sentinel-1 SAR data," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 101, no. April, p. 102359, Sep. 2021, doi: 10.1016/j.jag.2021.102359.

[55] S. S. Subbiah and J. Chinnappan, "Deep learning based short term load forecasting with hybrid feature selection," *Electr. Power Syst. Res.*, vol. 210, no. February, p. 108065, 2022, doi: 10.1016/j.epsr.2022.108065.

[56] R. K. Batchu and H. Seetha, "An integrated approach explaining the detection of distributed denial of service attacks," *Comput. Networks*, vol. 216, no. February, p. 109269, 2022, doi: 10.1016/j.comnet.2022.109269.

[57] R. C. Chen, C. Dewi, S. W. Huang, and R. E. Caraka, "Selecting critical features for data classification based on machine learning methods," *J. Big Data*, vol. 7, no. 1, p. 52, Dec. 2020, doi: 10.1186/s40537-020-00327-4.

[58] A. Thakkar and R. Lohiya, "Attack classification using feature selection techniques: a comparative study," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 1, pp. 1249–1266, Jan. 2021, doi: 10.1007/s12652-020-02167-9.

[59] L. B. Zeferino, L. F. T. de Souza, C. H. do Amaral, E. I. Fernandes Filho, and T. S. de Oliveira, "Does environmental data increase the accuracy of land use and land cover classification?," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 91, no. April, p. 102128, Sep. 2020, doi: 10.1016/j.jag.2020.102128.

[60] D. Albashish, A. I. Hammouri, M. Braik, J. Atwan, and S. Sahran, "Binary biogeography-based optimization based SVM-RFE for feature selection," *Appl. Soft Comput.*, vol. 101, 2021, doi: 10.1016/j.asoc.2020.107026.

[61] Q. Zhang, P. Liu, X. Wang, Y. Zhang, Y. Han, and B. Yu, "StackPDB: Predicting DNA-binding proteins based on XGB-RFE feature optimization and stacked ensemble classifier," *Appl. Soft Comput.*, vol. 99, p. 106921, Feb. 2021, doi: 10.1016/j.asoc.2020.106921.

[62] X. Ding, F. Yang, S. Jin, and J. Cao, "An efficient alpha seeding method for optimized extreme learning machine-based feature selection algorithm," *Comput. Biol. Med.*, vol. 134, no. May, p. 104505, Jul. 2021, doi: 10.1016/j.compbiomed.2021.104505.

[63] R. Wazirali, R. Ahmad, and A. A.-K. Abu-Ein, "Sustaining accurate detection of phishing URLs using SDN and feature selection approaches," *Comput. Networks*, vol. 201, no. October, p. 108591, Dec. 2021, doi: 10.1016/j.comnet.2021.108591.

[64] Y. Guo, Z. Zhang, and F. Tang, "Feature

selection with kernelized multi-class support vector machine," *Pattern Recognit.*, vol. 117, p. 107988, Sep. 2021, doi: 10.1016/j.patcog.2021.107988.

[65] C. E. Garcia, M. R. Camana, and I. Koo, "Machine learning-based scheme for multi-class fault detection in turbine engine disks," *ICT Express*, vol. 7, no. 1, pp. 15–22, 2021, doi: 10.1016/j.icte.2021.01.009.

[66] Y. Ding, L. Fan, and X. Liu, "Analysis of feature matrix in machine learning algorithms to predict energy consumption of public buildings," *Energy Build.*, vol. 249, p. 111208, Oct. 2021, doi: 10.1016/j.enbuild.2021.111208.

[67] P. Theerthagiri, "Predictive analysis of cardiovascular disease using gradient boosting based learning and recursive feature elimination technique," *Intell. Syst. with Appl.*, vol. 16, no. September, p. 200121, Nov. 2022, doi: 10.1016/j.iswa.2022.200121.

[68] P. R. Kannari, N. S. Chowdary, and R. Laxmikanth Biradar, "An anomaly-based intrusion detection system using recursive feature elimination technique for improved attack detection," *Theor. Comput. Sci.*, vol. 931, pp. 56–64, Sep. 2022, doi: 10.1016/j.tcs.2022.07.030.

[69] X. Jiang, Y. Zhang, Y. Li, and B. Zhang, "Forecast and analysis of aircraft passenger satisfaction based on RF-RFE-LR model," *Sci. Rep.*, vol. 12, no. 1, p. 11174, Jul. 2022, doi: 10.1038/s41598-022-14566-3.

[70] M. Lee, J. H. Lee, and D. H. Kim, "Gender recognition using optimal gait feature based on recursive feature elimination in normal walking," *Expert Syst. Appl.*, vol. 189, p. 116040, Mar. 2022, doi: 10.1016/j.eswa.2021.116040.

[71] C. Xu, J. Ding, Y. Qiao, and L. Zhang, "Tomato disease and pest diagnosis method based on the Stacking of prescription data," *Comput. Electron. Agric.*, vol. 197, no. 17, p. 106997, 2022, doi: 10.1016/j.compag.2022.106997.

[72] S. Yang *et al.*, "Intelligent multiobjective optimization for high-performance concrete mix proportion design: A hybrid machine learning approach," *Eng. Appl. Artif. Intell.*, vol. 126, no. PB, p. 106868, Nov. 2023, doi: 10.1016/j.engappai.2023.106868.

[73] A. Chandramouli, V. R. Hyma, P. S. Tanmayi, T. G. Santoshi, and B. Priyanka, "Diabetes prediction using Hybrid Bagging Classifier," *Entertain. Comput.*, vol. 47, no. May, p. 100593, 2023, doi: 10.1016/j.entcom.2023.100593.

[74] S. K. Prabhakar and D.-O. Won, "HISET: Hybrid interpretable strategies with ensemble techniques for respiratory sound classification," *Heliyon*, vol. 9, no. 8, p. e18466, Aug. 2023, doi: 10.1016/j.heliyon.2023.e18466.

[75] M. M. Alani, A. Mashatan, and A. Miri, "XMal: A lightweight memory-based explainable obfuscated-malware detector," *Comput. Secur.*, vol. 133, no. November 2022, p. 103409, Oct. 2023, doi: 10.1016/j.cose.2023.103409.

[76] D. H. Djarum, Z. Ahmad, and J. Zhang, "Reduced Bayesian Optimized Stacked Regressor (RBOSR): A highly efficient stacked approach for improved air pollution prediction," *Appl. Soft Comput.*, vol. 144, p. 110466, Sep. 2023, doi: 10.1016/j.asoc.2023.110466.

[77] H. Dong *et al.*, "Non-destructive detection of CAD stenosis severity using ECG-PCG coupling analysis," *Biomed. Signal Process. Control*, vol. 86, no. PC, p. 105328, Sep. 2023, doi: 10.1016/j.bspc.2023.105328.

[78] Q. Shang, L. Feng, and S. Gao, "A Hybrid Method for Traffic Incident Detection Using Random Forest-Recursive Feature Elimination and Long Short-Term Memory Network with Bayesian Optimization Algorithm," *IEEE Access*, vol. 9, pp. 1219–1232, 2021, doi: 10.1109/ACCESS.2020.3047340.

[79] X. Ding, F. Yang, Y. Zhong, and J. Cao, "A Novel Recursive Gene Selection Method Based on Least Square Kernel Extreme Learning Machine," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 19, no. 4, pp. 2026–2038, Jul. 2022, doi: 10.1109/TCBB.2021.3068846.

[80] K. Ren, Y. Zeng, Z. Cao, and Y. Zhang, "ID-RDRL: a deep reinforcement learning-based feature selection intrusion detection model," *Sci. Rep.*, vol. 12, no. 1, p. 15370, Sep. 2022, doi: 10.1038/s41598-022-19366-3.

[81] J. Sung *et al.*, "Classification of Stroke Severity Using Clinically Relevant Symmetric Gait Features Based on Recursive Feature Elimination With Cross-Validation," *IEEE Access*, vol. 10, pp. 119437–119447, 2022, doi: 10.1109/ACCESS.2022.3218118.

[82] S. Haroun, A. N. Seghir, and S. Touati, "Multiple features extraction and selection for detection and classification of stator winding faults," *IET Electr. Power Appl.*, vol. 12, no. 3, pp. 339–346, 2018, doi: 10.1049/IET-EPA.2017.0457.

[83] X. Su, H. Liu, and L. Tao, "TF Entropy and RFE Based Diagnosis for Centrifugal Pumps

VOL. 9. NO. 2 FEBRUARY 2023
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.5015

Subject to the Limitation of Failure Samples," *Appl. Sci.*, vol. 10, no. 8, p. 2932, Apr. 2020, doi: 10.3390/app10082932.

[84] J. L. Ferrando Chacón, T. Fernández de Barrena, A. García, M. Sáez de Buruaga, X. Badiola, and J. Vicente, "A Novel Machine Learning-Based Methodology for Tool Wear Prediction Using Acoustic Emission Signals," *Sensors*, vol. 21, no. 17, p. 5984, Sep. 2021, doi: 10.3390/s21175984.

[85] B. Zhang, Y. Zhang, and X. Jiang, "Feature selection for global tropospheric ozone prediction based on the BO-XGBoost-RFE algorithm," *Sci. Rep.*, vol. 12, no. 1, p. 9244, Jun. 2022, doi: 10.1038/s41598-022-13498-2.

[86] M. Mravik, T. Vetriselvi, K. Venkatachalam, M. Sarac, N. Bacanin, and S. Adamovic, "Diabetes prediction algorithm using recursive ridge regression l2," *Comput. Mater. Contin.*, vol. 71, no. 1, pp. 457–471, 2022, doi: 10.32604/cmc.2022.020687.

[87] M. D. Fathima, S. J. Samuel, and S. P. Raja, "HDDSS: An Enhanced Heart Disease Decision Support System using RFE-ABGNB Algorithm," *Int. J. Interact. Multimed. Artif. Intell.*, vol. In Press, no. In Press, p. 1, 2021, doi: 10.9781/ijimai.2021.10.003.

[88] A. Ebrahimi, U. K. Wiil, R. Baskaran, A. Peimankar, K. Andersen, and A. S. Nielsen, "AUD-DSS: a decision support system for early detection of patients with alcohol use disorder," *BMC Bioinformatics*, vol. 24, no. 1, p. 329, Sep. 2023, doi: 10.1186/s12859-023-05450-6.

[89] F. Soares and M. J. Anzanello, "Support vector regression coupled with wavelength selection as a robust analytical method," *Chemom. Intell. Lab. Syst.*, vol. 172, no. December, pp. 167–173, Jan. 2018, doi: 10.1016/j.chemolab.2017.12.007.

[90] S. Chatterjee, D. Dey, and S. Munshi, "Optimal selection of features using wavelet fractal descriptors and automatic correlation bias reduction for classifying skin lesions," *Biomed. Signal Process. Control*, vol. 40, pp. 252–262, Feb. 2018, doi: 10.1016/j.bspc.2017.09.028.

[91] S. Chatterjee, D. Dey, and S. Munshi, "Integration of morphological preprocessing and fractal based feature extraction with recursive feature elimination for skin lesion types classification," *Comput. Methods Programs Biomed.*, vol. 178, pp. 201–218, Sep. 2019, doi: 10.1016/j.cmpb.2019.06.018.

[92] H. Ebrahimy and M. Azadbakht, "Downscaling MODIS land surface temperature over a heterogeneous area: An investigation of machine learning techniques, feature selection, and impacts of mixed pixels," *Comput. Geosci.*, vol. 124, pp. 93–102, 2019, doi: 10.1016/j.cageo.2019.01.004.

[93] H. Shi, H. Wang, Y. Huang, L. Zhao, C. Qin, and C. Liu, "A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification," *Comput. Methods Programs Biomed.*, vol. 171, pp. 1–10, Apr. 2019, doi: 10.1016/j.cmpb.2019.02.005.

[94] S. Sun, M. Hu, S. Wang, and C. Zhang, "How to capture tourists' search behavior in tourism forecasts? A two-stage feature selection approach," *Expert Syst. Appl.*, vol. 213, no. PA, p. 118895, 2023, doi: 10.1016/j.eswa.2022.118895.

[95] H. H. Htun, M. Biehl, and N. Petkov, "Survey of feature selection and extraction techniques for stock market prediction," *Financ. Innov.*, vol. 9, no. 1, p. 26, Jan. 2023, doi: 10.1186/s40854-022-00441-7.

[96] N. V. Sharma and N. S. Yadav, "An optimal intrusion detection system using recursive feature elimination and ensemble of classifiers," *Microprocess. Microsyst.*, vol. 85, no. March, p. 104293, Sep. 2021, doi: 10.1016/j.micpro.2021.104293.

[97] P. Misra and A. S. Yadav, "Improving the classification accuracy using recursive feature elimination with cross-validation," *Int. J. Emerg. Technol.*, vol. 11, no. 3, pp. 659–665, 2020, [Online]. Available: https://api.elsevier.com/content/abstract/scopus_id/85086443367