ISSN 1870-4069

Differential Evolution for Feature Selection: A Systematic Literature Review

Francisco Javier Hernández-Somohano[1], Luz Ivana Correa-Hernández[2], Jesús Arnulfo Barradas-Palmeros[2],
Hector Gabriel Acosta-Mesa[2], Efrén Mezura-Montes[2]

[1] Universidad Veracruzana,
School of Statistics and Informatics,
Mexico

[2] Universidad Veracruzana,
School of administration and accounting,
Mexico

[3] Universidad Veracruzana,
Artificial Intelligence Research Institute,
Mexico

zS21013274@estudiantes.uv.mx, zS23000652@estudiantes.uv.mx,
heacosta@uv.mx, emezura@uv.mx, zs24010958@estudiantes.uv.mx

Abstract. Feature selection (FS) is an important task in data processing and analysis, aiming to reduce dimensionality and improve the performance of machine learning algorithms such as classification algorithms. Differential evolution (DE) has been successfully used for this purpose. However, a comprehensive assessment of their comparative strengths and weaknesses remains absent. In this systematic review of the literature that analyzes the state-of-the-art of DE for FS, 25 studies were selected for the review. Among the three evaluation criteria approaches (in this study, wrapper, filter, and hybrid approaches), most studies used a wrapper approach, with the k-nearest neighbors (KNN) algorithm being the most implemented. Considering how individuals are encoded, three representations were identified: real-number vectors, binary vectors, and integer-number vectors, with real-number vectors being the most used in DE for feature selection. It was found that most of the works follow a single-objective optimization process, and only a minority uses a multi-objective approach. Finally, for the main field of application, most studies focus on classification tasks using repository datasets from UCI Machine Learning Repository. This research aims to provide new insights into the state-of-the-art DE for FS.

Keywords: Feature selection, differential evolution.

1 Introduction

In various fields, feature selection (FS) plays a crucial role in reducing the dimensionality of datasets. The goal is to select the smallest and most relevant subset of features. The latter improves the interpretability of the data, accelerates model learning, simplifies them, and improves their performance in tasks such as classification [29]. FS becomes a complex problem due to its ample search space where the number of solutions is 2^n for a dataset with n features, as mentioned in [7].

Evolutionary computing (EC) techniques are well known for their ability to perform global optimization, including genetic algorithm (GA), particle swarm optimization (PSO), ant colony optimization (ACO), Artificial Bee Colony (ABC), Forest Optimization Algorithm (FOA), etc. Differential evolution (DE), proposed by Storn and Price in 1997 [23], is a recent and powerful metaheuristic approach that converges quickly and accurately. DE requires few control parameters, is robust, and is easier to use than other global optimization methods. Due to these advantages, differential evolution has been adopted by the FS community and has been successfully implemented in various studies, for example, [5,6], and [22]; however, works on DE for FS are much less than others EC techniques, such as GA and PSO.

Some studies, such as [2], propose a method for DE focused on FS, which consists mainly of four steps: initialization, mutation, crossover, and selection. Possible solutions are generated during initialization, where each solution, called a target vector, is encoded to represent a potential feature subset. After that, the evolutionary process starts, where each iteration is called a generation. In the mutation, a mutant vector is generated for each target vector. Then, the mutant vector is combined with the target vector in the crossover step, generating the trial vector. The target and trial vectors are compared in the selection step, and the one with the highest fitness is maintained in the population. The process is repeated until a stop criterion is met and the solution with the highest fitness in the population (single-objective) or a set of solutions (multi-objective) is returned.

According to [29], FS algorithms are generally classified into two categories: wrapper and filter. However, some studies have combined these two approaches, so a third category called "hybrid" was introduced. These criteria are applied in the selection step to evaluate the potential feature subsets. Wrapper approaches employ a machine learning algorithm, such as a classifier, to assess how well the subset performs within the algorithm. This approach is usually the most computationally expensive but usually gets better performance. Filter approaches evaluate subsets using statistical or theoretical measures to assess feature relevance. While computationally less expensive, this method is less accurate than a wrapper because it does not use a machine learning algorithm in the search process. Hybrid approaches integrate filter and wrapper measures for evaluation.

During the search process step, the goal is to find the optimal subset or subsets that achieve the best performance. However, there are different approaches to address this process. As discussed in [29], some studies combine classification performance and the number of selected features into an aggregate objective function, following a single-objective optimization approach. On the other hand, some studies propose a multi-objective optimization process, where two or more conflicting objectives are optimized simultaneously. Multi-objective approaches often maximize classification performance and minimize the number of selected features.

Diverse works can be found in the literature reviewing Evolutionary Computation and Bio-Inspired methods for FS, such as [1] and [14]. Nonetheless, none of them are focused only on DE. This study aims to analyze the state-of-the-art differential evolution for feature selection, focusing on various characteristics, their popularity, and applications to guide future research.

The rest of this paper is organized into five sections. Section 2 describes the method used, including the research questions, search strategy, and the inclusion and exclusion criteria process. Section 3 presents the results of the selected studies where the proposed method was applied. Section 4 discusses the findings. Finally, Section 5 concludes the study, summarizing the main findings and potential future research.

2 Method

The method used to conduct the systematic literature review is the one proposed by Kitchenham [15]. This method is carried out in three phases: 1) Planning, 2) Conducting the review. 3) Documentation of the review.

This section describes the research questions, the search strategy, and the study selection process.

2.1 Research Questions (RQ)

The research questions formulated to guide the review are:

- **RQ1.** What subset/individual evaluation approaches are used in DE-based algorithms for FS?
- **RQ2.** What representations of solutions are used in DE for FS algorithms?
- **RQ3.** What type of optimization (single-objective or multi-objective) is implemented in DE algorithms for FS?
- **RQ4.** In which applications or domains are DE-based algorithms used for FS?

2.2 Search Strategy

A preliminary search of articles on the topic was conducted to understand it better and formulate an appropriate search string.

Search String A search string consists of keywords related to the study topic. After testing different strings based on the number of found studies, the selected search string was:

```
("Differential Evolution") AND ("feature selection")
```

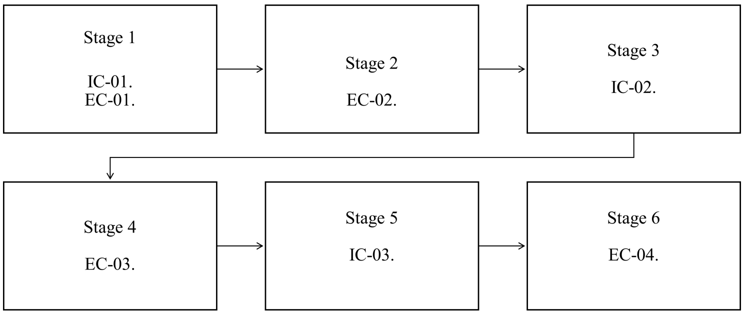

Fig. 1. Search Diagram: Inclusion and Exclusion Criteria by Phases.

Source Selection Initially, five sources were proposed (IEEE, ACM, SpringerLink, ScienceDirect, and Willey). However, Willey was discarded because no results were found during the period specified later. Finally, the search string was applied to the four remaining sources.

2.3 Study Selection

The study selection process was carried out in six phases. In the first phase, initial inclusion criteria (IC) and exclusion criteria (EC) were applied, while in the following phases, inclusion and exclusion criteria were alternated, as shown in Figure 1.

Inclusion Criteria

- **IC-01.** Studies must have been published between 2019 and 2024.
- **IC-02.** Titles must contain the terms *"Differential Evolution"* and *"Feature Selection"* to ensure topic relevance.
- **IC-03.** Studies must address at least two of the research questions in their abstract.

Exclusion Criteria

- **EC-01.** Studies written in a language other than English are excluded.
- **EC-02.** Studies that do not belong to the following categories are excluded: surveys, research articles, review articles, journals, or conference papers.
- **EC-03.** Duplicate studies found in the search are removed.
- **EC-04.** Studies that do not exclusively use differential evolution-based algorithms for feature selection are excluded.

2.4 Threats to Validity

It is important to acknowledge certain limitations that may affect the validity of this systematic review. First, our temporal scope (2019-2024) captures only recent developments in DE for FS, potentially missing foundational work published before 2019 as well as emerging research published after our cutoff date in 2024. Second, by restricting our search to four academic databases (ScienceDirect, SpringerLink, IEEE, and ACM), we may have overlooked relevant studies published in other repositories or specialized venues. Finally, our methodology lacks quantitative measures to evaluate the quality of selected studies or assess the comprehensiveness of our search string, which limits our ability to objectively evaluate the completeness of this review. Despite these limitations, we believe our findings provide valuable insights into current trends and characteristics of DE methods for FS, while acknowledging that a more comprehensive analysis could be performed in future research.

3 Results

This section describes the selected studies, their characteristics, and the answers to the research questions.

3.1 Study Selection

After applying the method mentioned in the selected sources, 25 studies were obtained.

DE has recently been implemented in FS, which should be considered when evaluating the number of studies on this topic. Despite the reduction in the number of studies during the first two phases, a sufficient number of studies were collected.

3.2 Study Characteristics

Publication Sources Of the 25 selected studies from the four consulted sources, a similar number of studies were found in all four sources: 7 in ACM, 9 in IEEE, 8 in ScienceDirect, and 7 in SpringerLink, indicating that the topic is present similarly in the selected sources.

Additionally, 52% (13 studies) were published in journals, while the remaining 48% (12 studies) were presented at conferences. This balanced distribution indicates that the topic has been explored both in journal articles and conference communications, reflecting sustained research interest.

Publication Years Within the study period, Figure 2 shows that the years with the highest number of published studies were 2020 (4 studies), 2023 (9 studies), and 2024 (5 studies). Contrarily, the years with the fewest published studies were 2019 and 2022, with 2 studies. Although studies on the topic were published throughout the period, the number of publications increased in only three years (from 2020 to 2023), suggesting a growing interest in applying DE to FS.

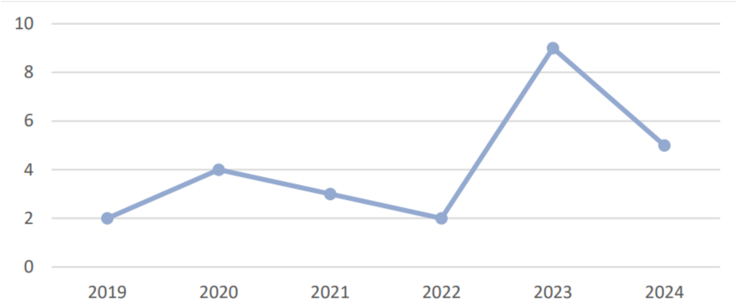

Fig. 2. Number of studies published per year.

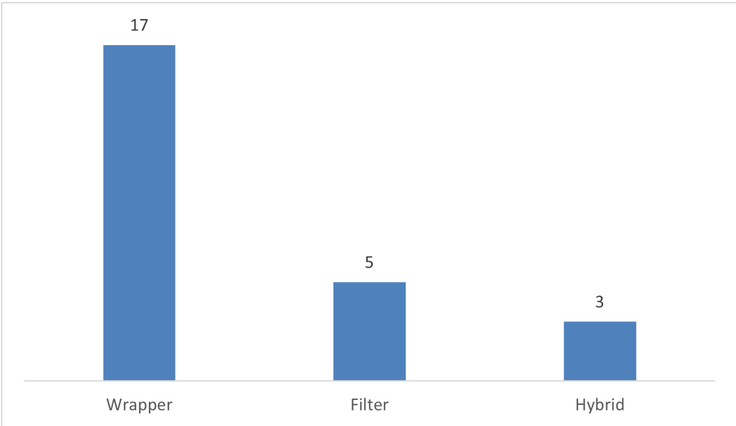

Fig. 3. Number of studies by evaluation approach.

3.3 Answers to Research Questions

Since the research questions are central to this study, all selected studies address at least one of them.

RQ1: What subset/individual evaluation approaches are used in DE-based FS algorithms? Three evaluation approaches were identified among the 25 selected studies. The most popular approach was the wrapper method, implemented in 17 studies. In contrast, filter and hybrid approaches were less common, with five and three studies, respectively, as shown in Figure 3.

As shown in Figure 4, seven classification algorithms were identified within the most popular approach (wrapper). The most commonly used were k-Nearest Neighbors (KNN) and Support Vector Machine (SVM). Three different techniques and measures were found for filter-based approaches, with correlation measures being the most frequently used. In hybrid approaches, two

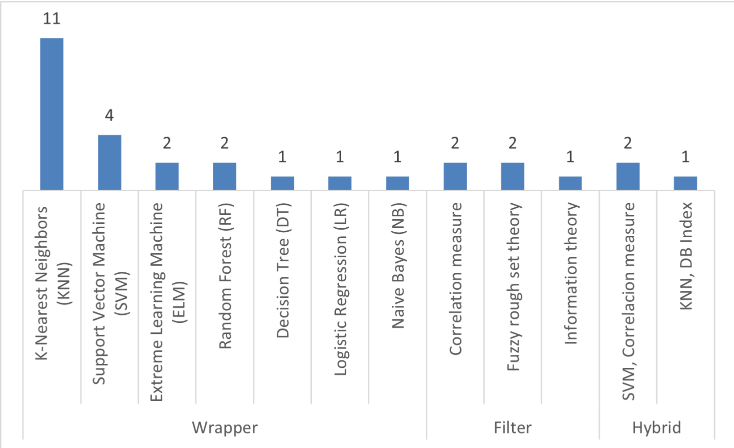

Fig. 4. Number of techniques by evaluation approach.

studies implemented correlation measures and an SVM model, while one used KNN and the Davies–Bouldin (DB) index.

RQ2: What representations of solutions are used in DE for FS algorithms? Three types of solution representations were identified. Among the 25 selected studies, the most commonly used representation was real-valued vectors, followed by binary-valued vectors. In contrast, integer-valued representations were the least used, as shown in Figure 5.

RQ3: What type of optimization (single-objective or multi-objective) is implemented in DE algorithms for FS? Among the two main types of optimization, most selected studies employed single-objective optimization. However, the number of studies using multi-objective optimization was only four fewer than those employing single-objective optimization, see Figure 6.

RQ4. In which applications or domains are DE-based algorithms used for FS? Based on the analysis of the selected studies, the application or domain where DE is implemented for FS was categorized into four main categories: classification, health and bioinformatics, security and informatics systems, and images and sensors. Table 1 specifies which study falls into each category and the specific case in which it is applied.

According to this categorization, most studies use DE for FS in classification with datasets from repositories such as the UCI Machine Learning Repository. These datasets come from various contexts and are used to evaluate the performance of the proposed algorithm across different datasets, comparing it with variations of the same algorithm or other metaheuristics.

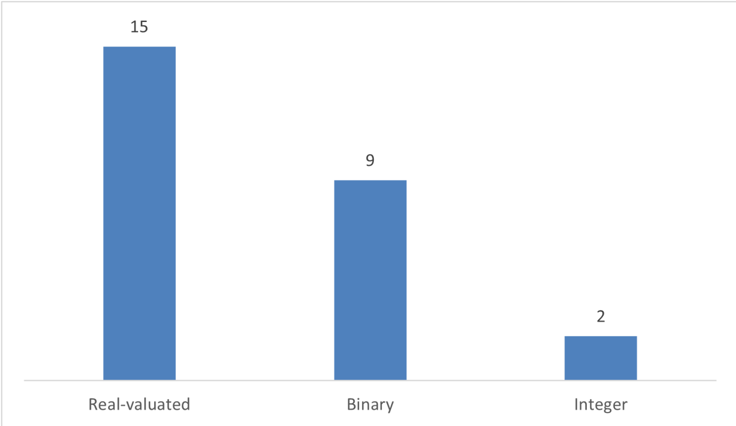

Fig. 5. Number of studies by solution representation.

Table 1. Applications and domains of DE for FS with corresponding studies.

Domain / Application	Case	Study
Classification	General classification tasks using standard benchmark datasets (UCI Machine Learning Repository and similar repositories)	[5],[6],[22],[30],[18],[12], [13],[16],[11],[27],[8],[25], [26],[4]
Health and Bioinformatics	Microarray data analysis for disease diagnosis, stroke prediction, tuberculous pleural effusion diagnosis, high-dimensional medical dataset classification	[28],[24],[20],[31],[17],[10]
Security Systems and Informatics	Network intrusion detection systems (IDS), software fault prediction	[2],[19],[9]
Images and Remote Sensors	Spectral feature selection of hyperspectral remote sensing images, Hand gesture classification using sEMG and motion sensor data	[3],[21]

However, there are studies with specific application contexts where the algorithm is implemented in datasets from a particular domain. In this aspect, most studies focus on health and bioinformatics, with six studies. Additionally, there are three studies in security and informatics systems and only two in the context of images and sensors. See Figure 7.

4 Discussion

Between 2019 and 2024, a regular ascent has been observed in the number of studies that employ DE for FS, with proportional information in the sources consulted, including survey articles and conference papers, demonstrating a growing interest in this field. This increase suggests that the implementation of this algorithm has started to be treated with more seriousness, supported by the quality and prestige of the sources where the studies are published.

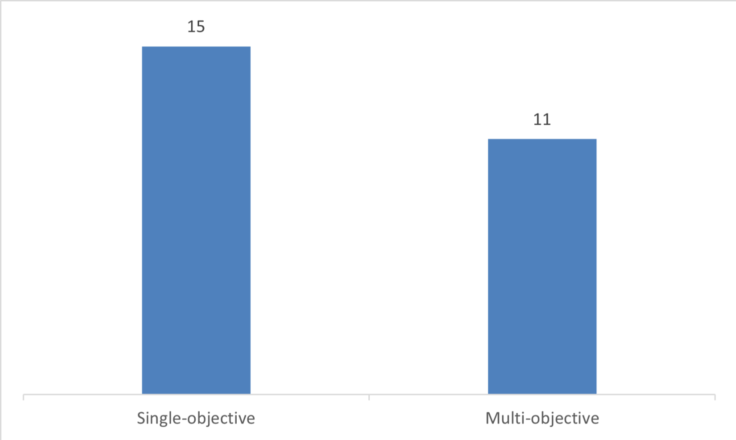

Fig. 6. Number of studies by optimization type.

Most of the selected studies use real-valued representations to encode feature subsets. The latter can be explained by the fact that DE was initially designed for continuous spaces, and working in this space allows the original structure of DE to be maintained without significant modifications. However, some studies, such as [2], have adopted a binary representation, which is appropriate for FS because the goal is to determine whether a feature is selected, transforming the search space into a binary space. This representation needs modifications to the algorithm's structure due to the alteration of the search space. On the other hand, a minority of studies (only 3) use an integer vector representation, where the selected feature index is directly utilized in a vector, such as [4], where, through a permutation strategy, use a permutational-based apace. Although this last strategy is less explored, it represents an opportunity for future work.

Regarding subset/individual evaluation approaches, most studies decide on wrapper approaches despite their higher computational cost. This popularity may be because most studies apply DE for FS in classification problems. Using a classifier as the evaluation criterion is suitable in this context, as it allows the performance of subsets to be measured in terms of their ability to predict adequately. Among the classifiers identified in the selected studies, the most common is KNN, which is less computationally expensive than more complex and robust algorithms like SVM. Despite their significantly smaller presence, Extreme Learning Machine (ELM), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), and Naive Bayes (NB) are employed in the studies. The preceding points to a possible opportunity to explore using a more complex and robust algorithm that could improve performance while also addressing the challenge of the high computational cost. On the other hand, filter-based approaches, which use measures like correlation, fuzzy set theory, or information theory, are less frequent but offer the advantage of being more computationally efficient.

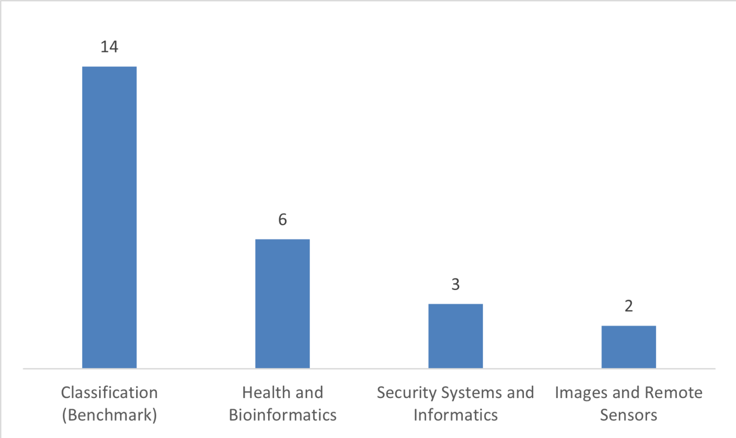

Fig. 7. Number of studies by applications or domains.

A small number of studies use a hybrid approach. In these studies, such as [22], where use a filter approach and wrapper approach in distinct stages, or such as[28] where each subset implements a redundancy measure (correlation) and the classification precision of the subset, looking to get solutions minimally redundant and predictive. Since only three studies tackle them, this field has great potential for future work. In terms of objective optimization, most studies focus on optimizing a single objective function. However, the difference in the number of studies addressing multi-objective optimization is not significant (only four fewer studies in comparison). Although multi-objective optimization involves a more complex process and generates non-dominated solutions, it can lead to better results in terms of solution diversity. The aforementioned provides a series of subsets for the user to choose the most convenient one.

Finally, in terms of applications, most studies implement DE in classification tasks using datasets from standard repositories such as UCI Machine Learning Repository, aiming to assess the algorithm's performance in various contexts and compare it with the performance of other algorithms applied to the same datasets. However, specialized applications were also identified, mainly in areas like health and bioinformatics, as well as (though less frequent) cybersecurity, and images and remote sensors. This focus on more specific applications demonstrates how DE for FS is maturing and offers valuable solutions to concrete relevant problems across various disciplines.

5 Conclusions

All research questions were successfully addressed through a systematic literature review. Trends and characteristics of DE applied to FS were identified, such as the representations used, evaluation methods, techniques employed within the approaches, the number of studies implementing multi-objective or single-objective optimization, and the various application domains involved. Due to the nature of the research questions, all selected studies provided relevant information to address the objectives posed.

Future research could focus on exploring new representations, distinct approaches, and strategies of DE for FS, as well as on how the algorithm's performance is evaluated and compared with other algorithms. Investigating new DE adaptations with integer-based representations and novel strategies for permutation-based search spaces appears to be a particularly promising area. Furthermore, identifying additional optimization objectives, developing more complex and robust algorithms within wrapper frameworks or hybrid strategies, and applying DE-based FS methods to datasets from different disciplines should also be considered. Additionally, given the popularity and high computational demands of wrapper methods, future research should explore strategies to reduce their computational cost.

Moreover, future studies could investigate more specific aspects of DE algorithms, such as the design and integration of novel operators (e.g., mutation, crossover, and selection mechanisms) specifically tailored for FS tasks. Additionally, the development of new DE-based strategies for FS represents a promising research direction. Finally, further studies could provide a more detailed examination of the factors that influence DE's performance in FS, offering deeper insights into its underlying mechanisms.

References

1. Agrawal, P., Abutarboush, H. F., Ganesh, T., Mohamed, A. W.: Metaheuristic Algorithms on Feature Selection: A Survey of One Decade of Research (2009-2019). IEEE Access 9, 26766–26791 (2021) https://doi.org/10.1109/ACCESS.2021.3056407
2. Almasoudy, F.H., Al-Yaseen, W.L., Idrees, A.K.: Differential Evolution Wrapper Feature Selection for Intrusion Detection System. Procedia Comput. Sci. 167, 1230–1239 (2020) https://doi.org/10.1016/j.procs.2020.03.438
3. Anamika, Gupta, R., Singh, G.: Binary Differential Evolution-Based Feature Selection for Hand Gesture Classification. In: Agrawal, R., Kishore Singh, C., Goyal, A. (eds.) Advances in Smart Communication and Imaging Systems, LNCE, vol. 721, pp. 221–232. Springer, Singapore (2021) https://doi.org/10.1007/978-981-15-9938-5_22
4. Barradas-Palmeros, J.-A., Mezura-Montes, E., Rivera-López, R., Acosta-Mesa, H.-G.: Computational cost reduction in wrapper approaches for feature selection: A case study using permutational-based differential evolution. In: IEEE Congress on Evolutionary Computation (CEC), pp. 124–131, IEEE (2024)
5. Bidgoli, A.A., Ebrahimpour-Komleh, H., Rahnamayan, S.: A novel multi-objective binary differential evolution algorithm for multi-label feature selection. In: IEEE Conference on Computational Intelligence, pp. 1–6. IEEE, (2019)
6. Bidgoli, A.A., Rahnamayan, S., Ebrahimpour-Komleh, H.: Opposition-based multi-objective binary differential evolution for multi-label feature selection. In: Springer Conference on Computational Intelligence, pp. 15–29. Springer, (2019)
7. Dash, M., Liu, H.: Feature selection for classification. Intell. Data Anal. 1(1–4), 131–156 (1997)

8. Dominico, G., Bernardes, J. S., Dorneles, L. L., Dorn, M.: Multi-Objective Wrapper Differential Evolution with Guided Initial Population for Feature Selection. In: 2023 IEEE Congress on Evolutionary Computation (CEC), pp. 1–8. IEEE (2023)
9. Faris, M., Mahmud, M.N., Salleh, M.F.M., Alsharaa, B.: A differential evolution-based algorithm with maturity extension for feature selection in intrusion detection system. Future Generation Computer Systems 137, 201–216 (2023)
10. Gudadhe, S., Thakare, A.: Differential evolution wrapper-based feature selection method for stroke prediction. In: Springer Advances in Data Science and Engineering, pp. 389–402. Springer, (2024)
11. Hancer, E., Xue, B., Zhang, M.: Fuzzy filter cost-sensitive feature selection with differential evolution. Knowledge-Based Systems 241, 108259 (2022) `https://doi.org/https://doi.org/10.1016/j.knosys.2022.108259`
12. Hancer, E.: New filter approaches for feature selection using differential evolution and fuzzy rough set theory. In: Proceedings of Springer Conference on Computational Intelligence, pp. 220–234. Springer, (2020)
13. Hu, X.-M., Guo, Z.-W.: Multimodal Bare-Bone Niching Differential Evolution in Feature Selection. In: 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1553–1558. IEEE, Melbourne, Australia (2021) `https://doi.org/10.1109/SMC52423.2021.9658633`
14. Jiao, R., Nguyen, B. H., Xue, B., Zhang, M.: A Survey on Evolutionary Multiobjective Feature Selection in Classification: Approaches, Applications, and Challenges. IEEE Transactions on Evolutionary Computation 28(4), 1156–1176 (2024) `https://doi.org/10.1109/TEVC.2023.3292527`
15. Kitchenham, B.A., Budgen, D., Brereton, P.: Evidence-Based Software Engineering and Systematic Reviews. Chapman & Hall/CRC, (2015)
16. Li, H., He, F., Chen, Y. et al.: MLFS-CCDE: multi-objective large-scale feature selection by cooperative coevolutionary differential evolution. Memetic Comput. 13, 1–18 (2021) `https://doi.org/10.1007/s12293-021-00328-7`
17. Mostafa, R.R., Khedr, A.M., Al Aghbari, Z., Afyouni, I., Kamel, I., Ahmed, N.: An adaptive hybrid mutated differential evolution feature selection method for low and high-dimensional medical datasets. Knowledge-Based Systems 283, 111218 (2024)
18. Nayak, S. K., Rout, P. K., Jagadev, A. K., Swarnkar, T.: Elitism based multi-objective differential evolution for feature selection: A filter approach with an efficient redundancy measure. *Journal of King Saud University - Computer and Information Sciences* 32(2), 174–187 (2020) `https://doi.org/10.1016/j.jksuci.2017.08.001`
19. Pethe, Y.S., Das, H.: Software fault prediction using a differential evolution-based wrapper approach for feature selection. Neural Computing and Applications 35, 457–473 (2023)
20. Prajapati, S., Das, H., Gourisaria, M.K.: Feature selection using differential evolution for microarray data classification. In: Springer Advances in Computational Intelligence, pp. 215–227. Springer, (2023)
21. Qian, Y.: Hyperspectral Feature Selection Algorithm Based on Differential Evolution and Multivariate Mutual Information. In: 2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI), pp. 1821–1826 (2023) `https://doi.org/10.1109/ICETCI57876.2023.10176658`
22. Qiu, C.: A hybrid two-stage feature selection method based on differential evolution. J. Intell. Fuzzy Syst. 39(1), 871–884 (2020) `https://doi.org/10.3233/JIFS-191765`
23. Storn, R., Price, K.: Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J. Glob. Optim. 11(4), 341–359 (1997)

24. Sudhakar, P., Satapathy, S.C.: Feature selection with binary differential evolution for microarray datasets. In: Springer Advances in Intelligent Systems and Computing, pp. 341–354. Springer, (2023)
25. Wang, P., Xue, B., Liang, J., Zhang, M.: Differential Evolution With Duplication Analysis for Feature Selection in Classification. IEEE Transactions on Cybernetics 53(10), 6676–6689 (2023) https://doi.org/10.1109/TCYB.2022.3213236
26. Wang, P., Xue, B., Liang, J., Zhang, M.: Feature selection using diversity-based multi-objective binary differential evolution. Applied Intelligence 53, 1543–1562 (2023)
27. Wang, P., Xue, B., Liang, J., Zhang, M.: Multiobjective Differential Evolution for Feature Selection in Classification. IEEE Transactions on Cybernetics 53(7), 4579–4593 (2023) https://doi.org/10.1109/TCYB.2021.3128540
28. Xie, W., Li, W., Fang, Y., Chi, Y., Yu, K.: A Hybrid Feature Selection Method Based on Binary Differential Evolution and Feature Subset Correlation for Microarray Data. In: 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, pp. 3193–3200 (2022) https://doi.org/10.1109/BIBM55620.2022.9995617
29. Xue, B., Zhang, M., Browne, W.N., Yao, X.: A survey on evolutionary computation approaches to feature selection. IEEE Trans. Evol. Comput. 20(4), 606–626 (2016). https://doi.org/10.1109/TEVC.2015.2504420
30. Zhang, Y., Gong, D., Gao, X., Tian, T., Sun, X.: Binary differential evolution with self-learning for multi-objective feature selection. Inf. Sci. 507, 67–85 (2020) https://doi.org/10.1016/j.ins.2019.08.040
31. Zhou, X., Chen, Y., Gui, W., Heidari, A.A., Cai, Z., Wang, M., Chen, H., Li, C.: Enhanced differential evolution algorithm for feature selection in tuberculous pleural effusion clinical characteristics analysis. Computers in Biology and Medicine 170, 106914 (2024)