Multicriteria Algorithms with Applications

Journal Homepage: sciencesforce.com/mawa

Multicriteria Algo. Appl. Vol. 4 (2024) 28–52

Paper Type: Review Article

Data Mining Problems Optimization by using Metaheuristic Algorithms: A Survey

Samia Mandour [1,*], Abduallah Gamal [1], Ahmed Sleem [2] and Mohamed Belal [3]

[1] Department of Computer Science, Faculty of Computers and Informatics, Zagazig University, Zagazig, Sharqiyah, 44519, Egypt; Emails: samia.rmdan@fci.zu.edu.eg; abduallahgamal@zu.edu.eg.

[2] Department of Computer Science, Faculty of Computers and Informatics, Tanta University, Tanta, Egypt; ahmed.selim@ics.tanta.edu.eg.

[3] Department of Computer Science, Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt; belal@fci.helwan.edu.eg.

Received: 02 Jan 2024 **Revised**: 15 May 2024 **Accepted**: 17 Jun 2024 **Published**: 20 Jun 2024

Abstract

Big data refers to large, diverse, and complicated data sets that are challenging to store, analyze, and visualize for use in subsequent operations or outcomes. Exploring and analyzing vast amounts of data in order to find significant patterns and principles is called data mining. Data mining is crucial to many human endeavors because it uncovers previously undiscovered patterns that are helpful. There are several main tasks of data mining, including Clustering, feature selection, and association rules. Several data mining techniques are employed to handle these significant duties. Metaheuristic algorithms are currently regarded as one of the most efficient methods for handling data mining issues. Black boxes like metaheuristics can offer distinct solutions regardless of the problem's nature. These algorithms treat data mining problems as combinatorial optimization problems. Numerous research papers are published in this area each year, which is why we decided to give a survey study on the topic. Consequently, this paper provides a thorough literature review on using metaheuristic algorithms to solve data mining issues that have emerged in the last five years (2019-2023).

Keywords: Big Data; Data Mining; Clustering; Feature Selection; Association Rules; Metaheuristic; Combinatorial Optimization Problems.

1 | Introduction

Large-scale datasets have grown quickly over the past few decades as a result of the fast development of computer and database technologies. On the other hand, there is a sharp increase in high-dimensional datasets and high-speed and high-accuracy data mining apps. Data mining is used to extract usable patterns from huge data repositories, and is a crucial and crucial step in knowledge discovering in databases (KDD) [2], as shown in Figure 1. To find and extract intriguing patterns coming from the vast data repository, data mining uses a variety of methodologies, and algorithms [3]. Data mining has gained significant attention in the last two decades due to its significance for different fields such as decision-making[4], healthcare [5-8], education[9, 10], chemical engineering [11], climate[12, 13], kidney failure[14], recognition hand gesture [15], COVID-19

Corresponding Author: samia.rmdan@fci.zu.edu.eg

https://doi.org/10.61356/j.mawa.2024.4301

[16-18], criminology [19-21], banking [22, 23], business[24, 25], marketing [26, 27], agriculture [28, 29], medical diagnosis [30], and other applications.

Figure 1. Data mining process.

Data mining involves a number of primary duties, such as clustering, feature selection, and association rules. These important tasks are handled by a number of data mining methods. One of the most effective approaches for solving data mining problems at the moment is thought to be using metaheuristic algorithms. These algorithms are typically referred to as the quickest method of problem-solving because they have high capabilities for selecting the best and most practical solution from among all feasible options. Combining optimization methods with metaheuristic algorithms can help us select the best options from a large pool of viable ones with the least amount of numerical work [31]. There are numerous metaheuristic algorithms accessible, and these algorithms are categorized into five categories according to some papers [32], as shown in Figure 2. In recent years, a large number of metaheuristics (MHs) have appeared, grabbing the interest of many scholars. They have been demonstrated to be successful in addressing a variety of optimization problems, such as scheduling issues [33], parameter extraction from solar photovoltaic models [34], milling manufacturing optimization issues [35], green coal production issues [36], feature selection issues [37], optimum power flow issues [38], etc.

A lot of researchers were strongly motivated to use and adapt metaheuristic algorithms to solve problems relating to data mining because of their capacity to handle a wide variety of complex problems, including continuous optimization problems[39], discrete optimization problems[40], and others. Two decades ago, metaheuristics were frequently used to address the most pressing data mining problems such as feature selection, clustering, association rules, and others. This article's goal is to show how data mining and optimization are related and to present some of the most current research on the topic. This article tracks the publications on this topic from 2019 through 2023. This article's main contributions can be summed up as follows:

- Introducing some major data mining problems optimization based on meta-heuristic algorithms from 2019 to 2023.
- Review of the frequency of meta-heuristic algorithms employed by the methods under study.
- Review of the frequency of meta-heuristic algorithms employed for feature selection.
- Review of the frequency of meta-heuristic algorithms employed for feature selection with different data mining applications.

- Review of the frequency of meta-heuristic algorithms employed for clustering.
- Review of the frequency of meta-heuristic algorithms employed for association rules.

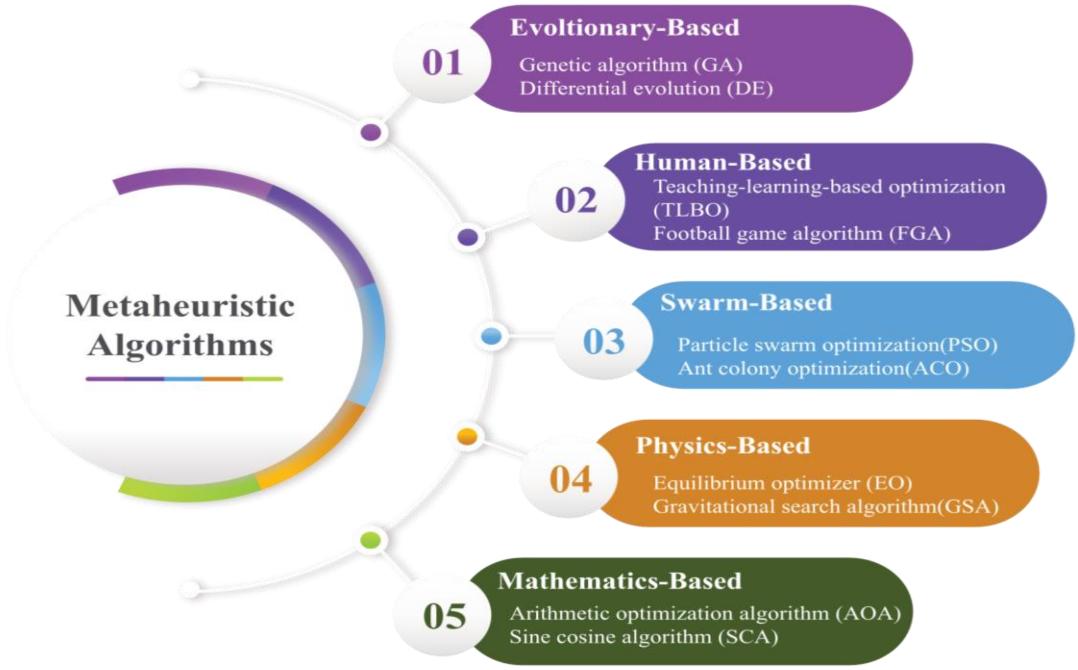

Figure 2. Classification of Metaheuristic algorithms.

2 | Research Trends in Data Mining

The literature review covers some significant advancements in metaheuristics, as well as how they have been successfully used in three selected data mining tasks (feature selection, clustering, and association rules).

2.1 | Data Mining Tasks

The kinds of patterns or data to be found throughout the data mining process can be specified using data mining functions or tasks. Association, clustering, and classification are some of the most important data mining tasks.

2.2 | Data Mining Methods

Based on a variety of data mining methodologies or approaches, data mining tasks are accomplished. The researchers have so far looked into a variety of data mining techniques. Currently, it is believed that applying metaheuristic algorithms is one of the best methods for resolving data mining issues.

2.3 | Data Mining Tasks using Metaheuristic Algorithms

This subsection will deal with three main data mining tasks, which are arranged as follows: (2.3.1) feature selection, (2.3.2) clustering, and (2.3.3) association rules.

2.3.1 | Feature Selection

Dealing with huge datasets can impede data mining due to their high dimensionality, so, it is a critical problem with machine learning methods [41]. A minimum of $10 \times n \times c$ training data are needed for a classification issue with n dimensions and C class, in accordance to a general rule[42]. Applications that use datasets with a lot of dimensions must therefore raise the classification parameters. Consequently, the classifier's performance considerably deteriorates. According to this principle, there is an urgent need to use methods for dimensionality reduction. Dimensional reduction is one well-liked method to get rid of noise and unnecessary features. It is a useful technique for increasing model generalization, reducing computational complexity, increasing precision, and reducing the amount of storage needed [43]. Dimension reduction has been suggested using two main methods: feature extraction [44], and feature selection[45]. In the fields of data mining, recognition of patterns, and statistics, feature selection has become a popular study topic. The primary goal of feature selection is to pick a subset of the available features by excluding those that have little to no predictive value and unnecessary, highly correlated features [46]. Subset generation, subset evaluation, stopping criteria, and result validation are the four major stages of the feature selection procedure. A subset of the candidate feature set arises from the initial features in each iteration of the finding process and the suitability of each subset is assessed using an assessment criterion. Up until the specified stop criterion is met, the subset generation procedure and its assessment are repeated. The best subset of the chosen feature has been verified on the test dataset after this procedure[47], as shown it Figure 3.

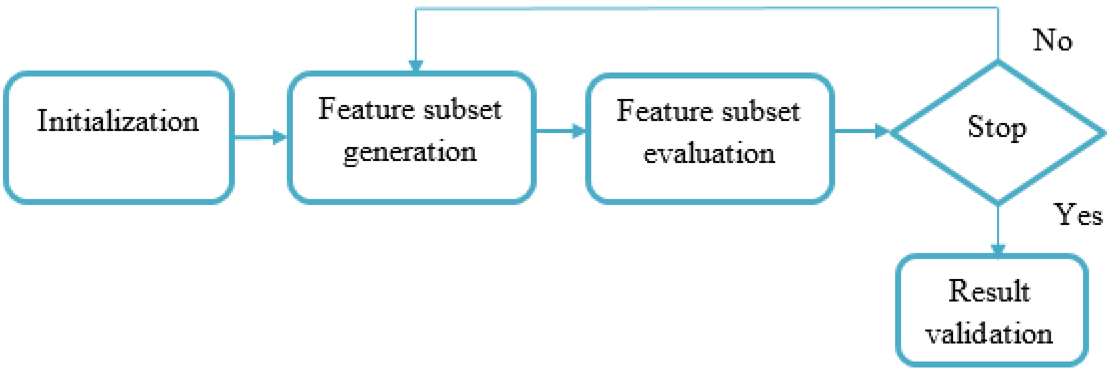

Figure 3. Feature selection process [1].

In general, there are two types of feature selection techniques: supervised and unsupervised feature selection methods [48]. A collection of train data is available for supervised methods, and each of these data sets is characterized by taking the values of the features with the class label. In contrast, train data for unsupervised methods lack class labels. Because of the use of labels for classes, it can generally be said that feature selection techniques perform more effectively and consistently in the supervised mode [49]. To find the best subset of features, several supervised feature selection techniques have been created. The methods are typically categorized into three groups: filter, wrapper, and embedded methods[50]. The process of learning or classification algorithms has no bearing on filter approaches. It always concentrates on the data's broad characteristics[51]. Wrapper methods continually interact with the classifier and contain the classification method. These techniques are more computationally costly than filters while also producing more precise results. Filters and wrapping methods combine to form embedded methods. In embedded techniques, the feature selection takes place during the training phase, which is conducted alongside the classifier[52]. Wrapper approaches produce outcomes that are superior to those of filter methods, but they are slower. The modeling technique that generates and then evaluates each subset is what determines how wrapper methods work. One of the most important methods of wrapper is called the randomized search method. Randomness is incorporated into randomized methods to help them avoid getting stuck in local optima and to help them explore the search area. The term "population-based approaches" refers to randomized algorithms[53]. Figure 4 displays a flow chart that classifies the approaches to solve feature selection.

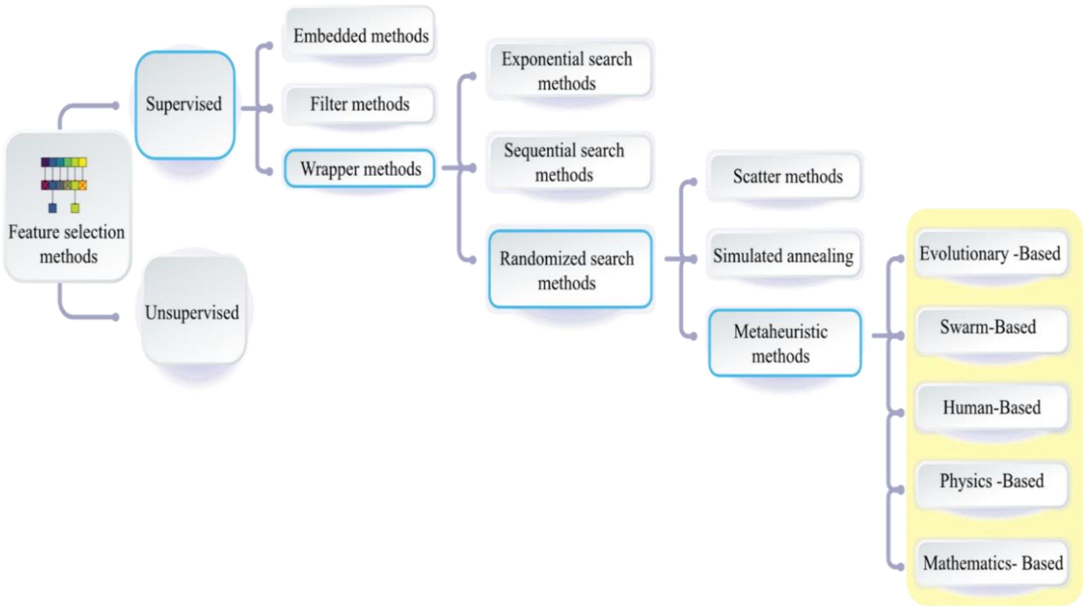

Figure 4. Feature selection methods techniques.

The task of finding the ideal subset of features is NP-Hard [54]. One of the best tools for solving combinatorial issues is the use of metaheuristic algorithms[55]. Furthermore, research demonstrates that metaheuristic algorithms outperform exhaustive or greedy methods[56]. Modern metaheuristic algorithms are heavily influenced by nature, and they are frequently employed in the field of feature selection today[57]. In this part of the study, we concentrate on the metaheuristics that have been suggested in the previous five years (2019- May 2023) for the feature selection issue.

2.3.1.1 | Wrapper-based Metaheuristic for Feature Selection

Meta-heuristics procedures are one method for resolving complex optimization and NP-Hard problems. Instead of searching for the best solution, meta-heuristic algorithms can uncover workable solutions in a reasonable amount of time. These algorithms belong to a class of approximate optimization algorithms that have methods for escaping local optima as well as applying them to a variety of optimization issues[58]. To prevent adding to the high dimensional dataset's computational complexity, many feature selection techniques use meta-heuristics[59, 60]. These algorithms address optimization problems and iteratively search for the best answer using simple principles and operations[61].

The general flowchart model of the primary tasks carried out by metaheuristic algorithms is shown in Figure 5. The fitness values of candidates are first determined once the initial sample has been established. The iterations begin later. The exploration and exploitation operators of the metaheuristics produce new candidate solutions given a termination condition. During the optimization process, it's crucial to avoid continually evaluating the same options. Since the metaheuristics' recombination operators would likely produce the same candidates repeatedly, there is no need to waste time recalculating them. Additionally, because these algorithms need a lot of computing, their quicker iterations, like parallel or dynamic programming, might produce superior results because they perform more fitness evaluations in a shorter amount of time. Figure. 5 represents the Feature selection cycle using metaheuristic algorithms[62].

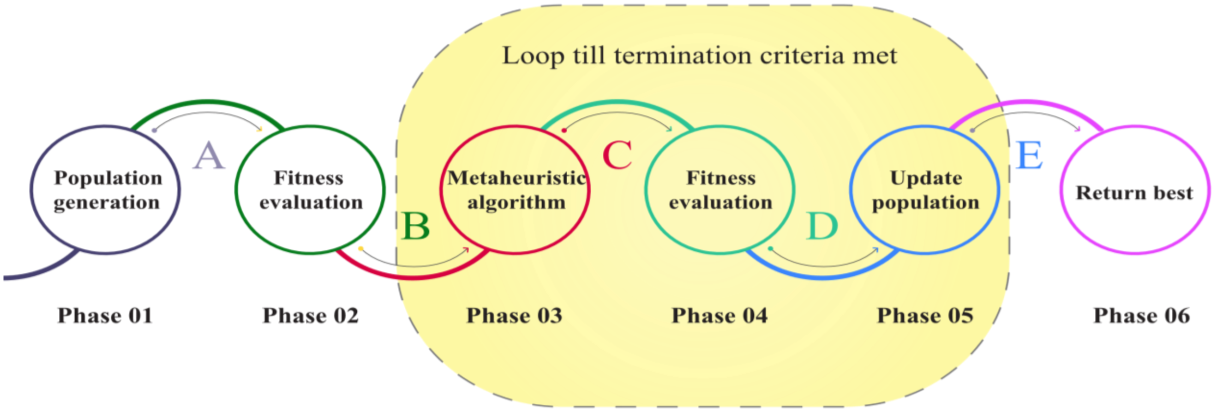

Figure 5. Feature selection cycle using metaheuristic algorithms.

The population of potential solutions is used in metaheuristic optimization techniques. Typically, the solutions are shown as a sequence of values (vector). A representation of a solution for metaheuristic feature selection algorithms is often a binary representation of a chosen collection of features. Therefore, each potential solution can be represented by $d$dimensions; each solution is initially set up using binary numbers (0 or 1). By choosing a handful of the potential features (one value) and excluding other features (zero value); the feature selection problem for classification purposes is summed up. A proposed solution with its chosen features can be seen in Figure 6. Four out of the eight characteristics in this solution have been chosen (green ones).

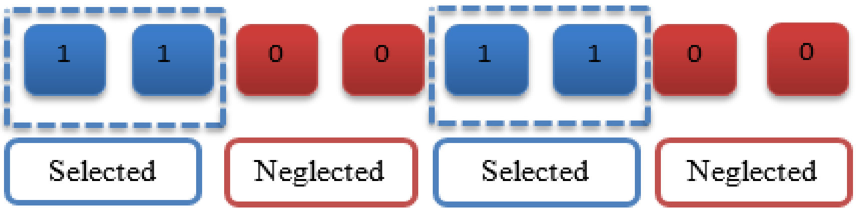

Figure 6. Solution's binary encoding.

A wrapper feature selection method optimizes an objective function to choose the optimum feature subset. Depending on the classification issue, different objective functions for feature selection are constructed. An objective function that maximizes accuracy in classification or minimizes the number of selected characteristics was previously developed. Additionally, the multi-objective function was developed to merge the two opposing objectives for solving the feature selection problem. By giving weights for each of the objectives and running the learning method, the multi-objective function issue was reduced to a single objective. It is important to note that numerous metaheuristic algorithms have been created since 1966. Between 2019 and 2023, a number of research articles that were submitted in this regard throughout five years are covered in Table 1.

Table 1. An overview of some wrapper-based metaheuristic for feature selection.

No	Ref	year	Methodology	Dataset	Advantages	Shortcomings
1	[63]	2022	A new optimizing approach is being proposed within which the grasshopper's position is represented by binary values and its values are modified using operators.	20 different-sized data sets from the University of California, Irvine (UCI) machine learning library [64]	Increasing exploration potential and cutting down on overall computation time	-
2	[65]	2020	A new wrapper-based metaheuristic selection approach is being introduced By describing the population as a set of quantum bits, to	Fourteen difficult datasets through the areas of face image detection, microarray gene expression(ASU datasets), high dimensional text(Text Datasets),	Establishing an effective decision-making system and increasing accuracy	Difficulties with both, continuous and multi-objective optimization

			enhance the exploration and exploitation of feature selection.	and a variety of datasets from the UCI library (UCI datasets)		have not been solved
3	[66]	2021	Suggested an innovative algorithm to enhance the algorithm's capacity for exploration based on the behavior of a butterfly.	21 feature selection problems,23 benchmark test functions, 30 benchmarks from CEC2014, and 30 benchmark functions from CEC2017, the proposed method was assessed	Providing an excellent equilibrium in the search space and preventing stagnation into local minima.	Computational time is high
4	[67]	2021	A powerful wrapper-based approach called (BBO-SVM-RFE) increases the balance between exploration and exploitation in the original BBO by combining the embedded support vector machine recursive feature elimination (SVM-RFE) for the Feature Selection process with the Binary Biogeography Based Optimization (BBO) based optimizer.	On the basis of 18 benchmark datasets[68], the suggested BBO-SVM-RFE approach was evaluated.	The appropriateness of the features is the main focus of BBOSVM-RFE	Redundancy Issue with SVM-RFE methodology. In other words, more features are chosen compared to other approaches, which creates
5	[69]	2021	Three new operators the correlation-based particle swarm optimization (PSO) initialization method, the relevance redundancy-based local search, and the adaptable flip mutation are combined in a unique feature selection algorithm based on bare bones PSO (BBPSO) with mutual information.	In regard of 16 well-known datasets, BBPSO is validated [70-72]	Better classification accuracy using fewer features	Computational cost is the main drawback.
6	[73]	2023	A new variant of grasshopper optimization algorithm(GOA) based on incorporating an elite opposition-based learning method called EGOA to strengthen the global optimization ability of GOA	To verify EGOA, 21 different publicly accessible data sets are employed. [74]	EGOA gains precision, choosing the best feature, optimizing search, and achieves better values in terms of cost evaluation indices	-
7	[75]	2021	Employing the binary crow search algorithm (BCSA) with time varying flight length, named (BCSA-TVFL) to identify new features and determine the flight length parameter. Eight different transfer functions are then investigated to discover the best fit for the suggested strategy.	20 common test sets from the UCI library were utilized to compare the algorithm's performance.	Addressing the issues with dimensionality reduction, and enhances the accuracy of feature selection in a more significant way.	High computational cost
8	[76]	2021	In order to calculate the final dimension of the optimum feature selection (OFS) for the subsequent OFS search, a two-stage hybrid ant colony optimization (ACO) for high-dimensional feature selection (TSHFS-ACO) leverages the interval technique.	Arcene from the NIPS feature selection contest[77] ,and ten publicly available gene expression datasets[78]	Minimizing the algorithm's run time, escaping from a local optimum, and identifying the feature subset with the best fitness value, are the main advantages of the proposed strategy	Due to an excess of hyper-parameters, the algorithm lacks stability and has redundant features.

9	[79]	2019	Brain Storm Optimization (BSO) along with the Fuzzy Min-Max (FMM) neural network is used to tackle picking features and categorizing issues.	Ten benchmark data sets are selected, through the UCI machine learning library[80]	Enhancing classification precision while reducing model complexity	Compared to some comparable methods, the suggested one demands longer execution times due to poor investigation.
9	[81]	2021	With the use of a time-varying transferred function as well as the Binary Jaya approach, an innovative mixed feature selection strategy that incorporates five filters and a wrapper method has been suggested, helping to balance the proposed method's trade-off between diversification and intensification.	The performance is evaluated using 10 benchmarks micro-array datasets http://csse.szu.edu.cn/staff/zhuzx/Datasets.html	Efficient in terms of classification accuracy ,and execution time	-
10	[82]	2022	Feature selection for better Alzheimer's classification accuracy utilizing effective Fisher Score and greedy searching	The datasets used for the experiments are ADNI-TADPOLE [83] and AIBL[84].	A very effective minimal feature set for SVM and KNN is discovered, revealing a superior minimum feature set that can enhance the model's performance.	Discovering a better minimum set of features, though not the greatest set, testing on a small sample size, and Missing data frequently has a negative impact on a classifier's performance.
11	[85]	2020	Introduced a hybrid approach that combines swarm grasshopper intelligence with quantum computing to improve both the consistency of the local search space and exploration ability.	Twenty datasets acquired from the UCI repository are used to run the algorithm[86]	The best possible subset is sought out.so, it provides a better accuracy	The issues of multi-objective feature selection, applying to big datasets are not addressed.
12	[87]	2020	Two supervised and unsupervised heuristic functions are introduced, Through numerous rounds, a multi-label ant colony optimization (MLACO) seeks the most exciting attributes in the domain of the feature with the lowest duplication and highest relevance with class labels.	The effectiveness of this strategy is evaluated using nine well-known datasets (Corel5k, Scene, 20NG, Image, Chemistry, Chess, Cooking, CS, and Bibtex).	The key benefits are an improved rate of classification using a relevance and redundancy analysis and the quickest average execution time	The primary challenge is to improve similarity metrics between class labels.
13	[88]	2020	using a wrapper-based feature selection approach and the weighing operator, smart crossover, and mutations operators	Five datasets (Lung, Dermatology, Arrhythmia, WDBC, and Hepatitis) are utilized to verify this strategy using datasets from the UCI public repository.	Generalization ability, efficient accuracy for two-class and multi-class data on a broad scale, the minimum amount of features, minimizing the processing time	Highly unbalanced datasets and optimizing with additional hyper-parameters are not handled.
14	[89]	2019	A new hybrid optimization method that manages the trade-off between inquisitive	The approach is evaluated using 18 benchmark datasets gathered	Accurate in terms of precision, the quantity	SVM and Artificial Neural Networks (ANN),

			and exploitative behaviors during optimization iterations by combining the strengths of both grey wolf optimization (GWO) and particle swarm optimization (PSO).	from the UC Irvine Machine Learning Repository[90]	of features, and the execution duration	two fierce rivals of KNN, didn't compare with it.
15	[91]	2020	Utilizing new transfer functions, Binary Grey Wolf Optimizer addresses the discretization issues with feature selection.	Twelve datasets are taken from the UCI machine learning repository and used to verify the performance of the suggested approach[92]	Good accurate classification	The paper did not implement feature selection using a neural network.
16	[93]	2021	A hybrid technique using the simulated annealing (SA) algorithm and the Harris Hawks optimization (HHO) algorithm by using two bitwise operations is designed to find the best subset of characteristics.	To validate this strategy, 24 standard datasets and 19 artificial datasets with dimension sizes that can exceed hundreds were used. Data set can be found in[80] and https://www.openml.org/search ?type=data	Escape from local optimum, increasing population diversity, transferring the desirable traits to the population's members, and doing so in a suitable amount of time.	Computational cost and the time commitment.
17	[94]	2022	Combining seeding and chaos population techniques in a binary dandelion algorithm (DA) for feature selection	From the UCI Machine Learning Archive, 15 datasets used as benchmarks (http://archive.ics.uci.edu/ml)	Achieving smaller feature subsets with outstanding classification accuracy	inadequate search performance as a result of the early iterations' low convergence ability
18	[95]	2022	In this paper, the feature selection method uses three States based on a hybrid chaotic/vortex search algorithm (VSA).	From the UCI machine learning repository, 24 datasets were gathered[64]	The key benefits in terms of classification performance overall include the number of FS, outstanding stability, and quick convergence.	-
19	[96]	2022	Utilizing the greedy crossover technique, a new metaheuristic known as the coronavirus herd immunity optimizer (CHIO) was developed to address FS issues in medical diagnosis.	A COVID-19 dataset from the real world from the written link and 23 medical benchmark datasets from (UCI ,Kaggle ,and KEEL)https://github.com/Atharv aPeshkar/ Covid-19-Patient-Health-Analytics.	convergence speed and classification accuracy	weak capacity to leverage search results (exploitation ability)
20	[97]	2020	A new swarm intelligence system inspired by the behavior of coyotes called the binary coyote optimizing algorithm (BCOA) has been suggested.	The UCI Machine Learning Repository contains seven datasets that are used in this study, including "Statlog," "Spect," "Breast Cancer," "Sonar," "Soybean," "Arrhythmia," and "Zoo."they are found in [64]	Good results in terms of training accuracy on average a proper equilibrium in search strategy, avoiding haphazard searches and avoiding local optima	sensitivity to huge data sets
21	[98]	2021	HLBDA, a new Binary Dragonfly Algorithm with a hyper learning approach, was suggested as a wrapper-based approach for FS.	A COVID-19 application and 21 datasets gathered from Arizona State University and the repository of UCI are utilized for assessment.	The principal merits are the best subset of highly discriminative features and the maximum accuracy.	A weakness in initialization strategies
22	[99]	2022	A binary version of an improved whale optimization algorithm with three effective search strategies, migrating, selective selection, and enriched encircling prey is proposed.	typical medical data set from the machine learning library at the University of California, Irvine [64]	Comparable in terms of precision, sensibility, and accuracy to the most recent high-performing binary optimization algorithms.	-

23	[100]	2020	An enhanced binary variant of the salp swarm method by adding inertia weight parameter for wrapper approach feature selection problems	There were 23 UCI benchmark datasets used[101].	Classification precision and a small number of attributes are advantages.	-
24	[102]	2020	an enhanced version of the SSA method that addresses the feature selection issues utilizing the opposition-based learning (OBL) technique and the local search algorithm	All experiments used 18 UCI benchmark datasets from the UCI datasets source[64].	Fast convergence to best subset, and efficient accuracy performance	choosing additional features above other comparable optimization algorithms
25	[103]	2022	A proposed algorithm using binary artificial algae to solve classification issues	25 public datasets with varying degrees of difficulty were chosen from the well-known data resource[104]	Results stable and successful.	lackage of optimum solutions with parameter tuning
26	[105]	2022	A different FS approach built on the Group Search Optimizer (GSO), along with the Logistic, Piecewise, Singer, Sinusoidal, and Tensor maps of Chaos	Twenty common datasets with various size and dimensions descriptions	Classification error rate, classification precision, and processing time are the major merits	Multi-objective optimization issues that have not been solved
27	[106]	2023	Using transfer functions to create five alternative versions of the binary greater cane rat (BGCRA) algorithm, which was motivated by the GCR's understandable nocturnal behavior, it was possible to choose the most affordable and efficient version among them.	12 benchmark datasets (lonosphere, CongressEW, SpectEW,Breastcancer,Pima,StatlogHeart,Exactly,Exactly2,M-of-n,Vote,WineEW,Zoo) were taken from UCI sources.	lower the dimensionality, choose useful feature sets, and produce more accurate results	Extended computation time.
28	[107]	2023	Gorilla troops optimizer (GTO), a new metaheuristic algorithm, has been improved to get over the drawbacks of the original GTO by integrating three techniques: Tangent Flight, Cauchy Inverse Cumulative (CICD) Distribution Operator, and Opposition Based Learning (EOBL)	Sixteen benchmark datasets namely, BreastEW, CongressEW, Exactly, Exactly2, HeartEW, IonosphereEW, Lymphography, M-of-n, PenglungEW, SonarEW, SpectEW, Vote, WineEW, and Zoo were used in this article.	Regarding precision, and the amount of features, there is quick convergence, the lowest cost, and strong stability	High computation time.
29	[108]	2023	To handle FS challenges, a new equilibrium optimizer version improved with a self-adapting mechanism, and the theory of quantum physics is integrated with an artificial bee colony algorithm.	Twenty-five datasets were selected from the UCI repository[64]	Improving the classification accuracy	It doesn't address large-scale data problems and multi-objective criteria
30	[109]	2023	As a feature selection method, a novel binary Colony Predation Algorithm version relying on the Gaussian Cuckoo Variable Dimensional Strategy is presented.	The UCI machine learning repository contains 12 high-dimensional biomedical data sets that are utilized for validation.	smallest feature subset with the highest feature selection classification accuracy	High computational cost Limitations while tackling discrete and multi-objective optimization issues

2.3.1.2 | Wrapper-based Metaheuristic for Feature Selection in Machine Learning Applications

Data mining and machine learning researchers and engineers have a hurdle when analyzing high-dimensional data. By eliminating duplicate and irrelevant data, the Wrapper-based metaheuristic for feature selection offers a practical solution to this issue, as shown in Figure 7. This can increase learning accuracy and enable a deeper comprehension of the learning model or data. Wrapper-based metaheuristic for feature selection has been applied effectively in different sectors of machine learning (ML), such as the internet of things (IoT) in its different sectors, online fraud detection, email spam and malware filtering, image recognition, speech recognition, and automatic text categorization. Some of the research articles submitted on this subject over the last five years will be covered in Table 2.

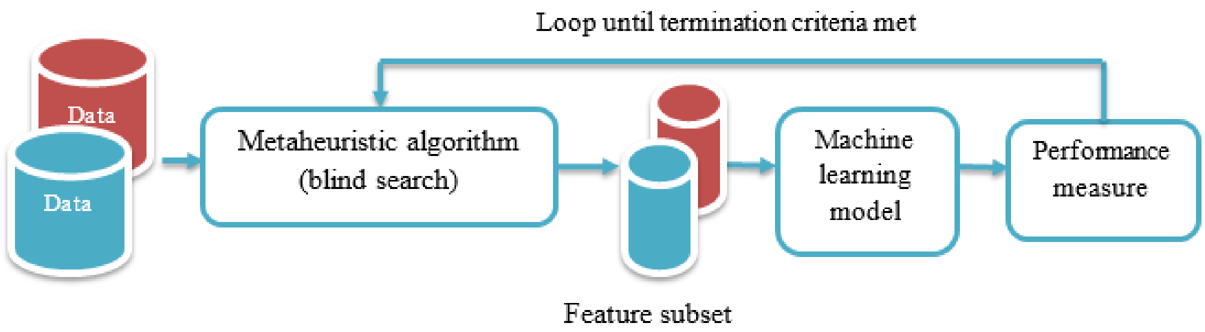

Figure 7. Wrapper-based metaheuristic for feature selection in ML.

Table 2. A summary of wrapper-based metaheuristic for feature selection in different machine learning sectors.

	Machine learning sectors									
	IOT					Online fraud detection	Email - spam and malware-filtering	Image - recognition	Speech recognition	Automatic text categorization
	Health-care	Self-Driving-Cars	Smart-Cities	Smart-Agriculture	Financial					
Publication	[110, 111]	[112, 113]	[114, 115]	[116, 117]	[118, 119]	[120, 121]	[122, 123]	[124, 125]	[126, 127]	[128, 129]

2.3.2 | Clustering

Clustering is the approach of grouping things into units that share qualities. A bunch of objects is divided up into groupings termed "clusters" to separate them apart such that they are more similar to one another compared to items in other clusters [130]. The clustering job is regarded as an unsupervised learning instance if the objects don't have any external information. The most popular unsupervised learning technique, known as cluster analysis, is used to discover hidden patterns or clustering in data. Data clustering identifies a collection of homogeneous patterns in the data set. So, the goal is to create an algorithm that can correctly divide a dataset that has not been leveled into groups. Numerous clustering techniques have been developed recently, and they can be divided into hierarchical and partitional techniques [131], as shown in Figure 8. In hierarchical clustering approaches, cluster size and shape are not taken into account; instead, data are sorted in a hierarchical tree structure based on how similar the data points are. In other words, because only one cluster may be selected at a time during the clustering analysis process, this approach results in static cluster formation. In the context of partitional approaches, the dataset within a collection of distinct clusters is directly analyzed to minimize intra-cluster dissimilarity and maximize inter-cluster dissimilarity. Even though these two clustering techniques persist in use today, their effectiveness depends on knowing in advance how many clusters a dataset has. Because prior knowledge of the number of groups that naturally occur in the data is frequently unavailable and calculating the ideal number of clusters for such datasets becomes highly challenging, the present methods cannot be used to tackle problems in the real world. This is the main justification for the data clustering problem's classification as an NP-hard task. Datasets were automatically clustered to help with this problem. Having no prior knowledge of the dataset's values of attributes, automatic clustering determines the number and structure of clusters in a dataset spontaneously [132].

The K-mean algorithm is one of the most effective algorithms for solving clustering problems, due to the effectiveness of its time complexity, because it relies on the deterministic local search method. This algorithm might fall into local optima, which led researchers to consider employing meta-heuristic algorithms. Meta-heuristic algorithms are widely used due to their random search nature, leading to faster convergence times and capacity to produce high-quality solutions, so metaheuristic algorithms are chosen over classically based ones for resolving large-scale data clustering issues. Many algorithms have emerged in this regard. We will review several algorithms in Table 3.

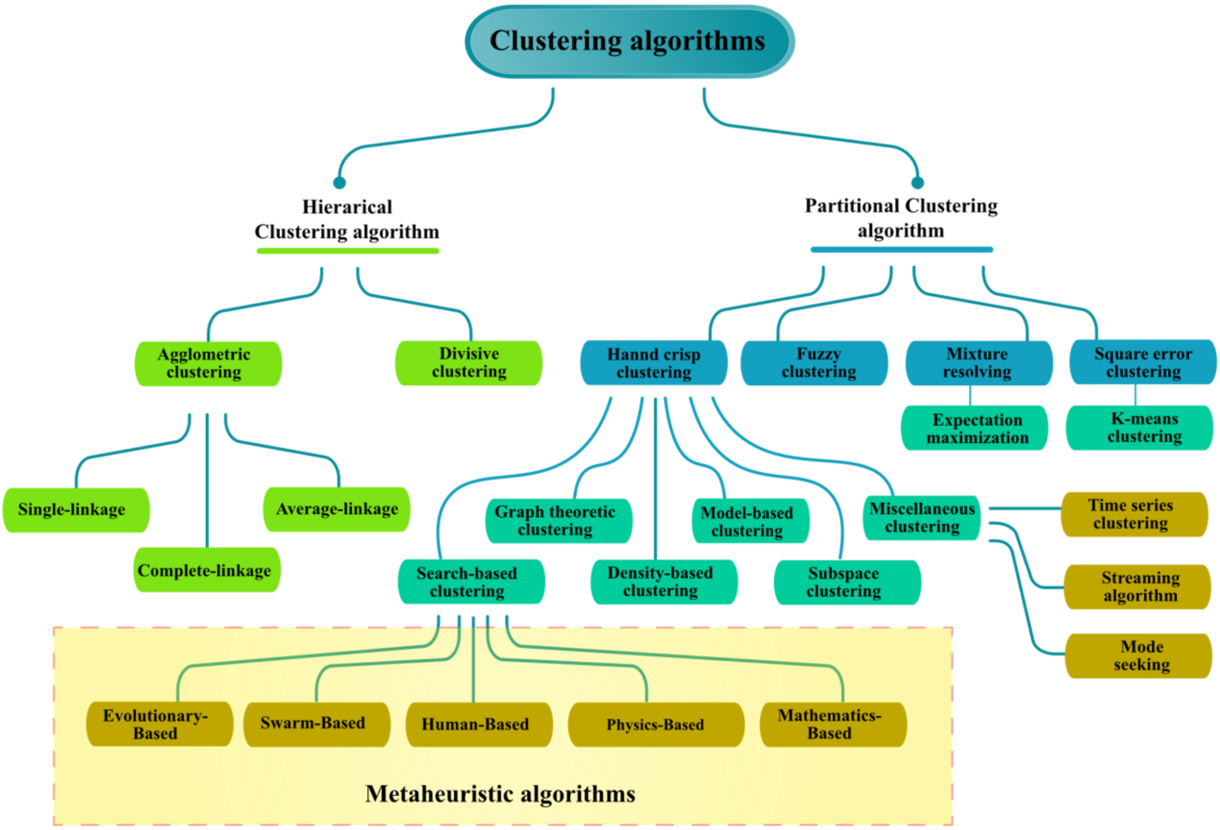

Figure 8. Clustering techniques.

Table 3. An overview of some metaheuristic algorithms for clustering.

No	Ref	year	Methodology	Dataset	Advantages	Shortcomings
1	[133]	2020	A new optimizing approach for data clustering based on Harris Hawks Optimizer algorithm enhanced with Chaotic sequences called (CHHO)	A total number of 12 datasets vary between Shape and UCI datasets are used for validation: Shape datasets (Flame,Jain,Aggregation,Compound,R15,D31,Spiral, Path based), and UCI datasets(Glass, Iris, Wine, and Yeast)	It is quite effective at resolving the data clustering problem and minimizing the risk of local entrapment while doing so.	The difficulty of convergence towards optimal points in the case of large scale
2	[134]	2019	Data clustering methodology suggested combining memetic algorithm steps with a new variant of differential evolution based on a mutation operator and a neighborhood selection heuristic.	A number of six datasets are used for assessment (Iris ,Wine ,Vowel ,CMC ,Glass ,Cancer)	consistency in performance for the F-measure validity measures, accuracy, and the average of the intra-cluster distances	-

3	[135]	2022	An enhanced variant of electromagnetic field optimizer called electromagnetic clustering algorithm is proposed for determining the optimal centroid for performing optimal data clustering.	Eight datasets from UCI repository including 2 IOT datasets(Gas, Human Activity), and other datasets(IRIS, LONO,CMC,CRUDE-OIL,THYROID ,and Vowel)	High efficiency in solving data clustering problems, and achieves more stability in results	-
4	[136]	2019	The proposed algorithm was inspired by three other algorithms, including the local search, the ant colony, and the ant lion for data clustering, and enhanced with the Cauchy mutation operator	Four datasets are used: IRIS, Glass, Wine, and ZOO.	High performance	The algorithm needs to firstly adjustment parameters depending on the issue.
5	[137]	2019	The suggested technique for data clustering is based on a novel type of metaheuristic algorithm called coral reef optimizer (CRO) with substrate layers (SL) of PSO and an adapted version of GKA depends on mutation operator as a local search strategy	Seven real datasets(IRIS ,Wine , Breast Cancer Wisconsin, HTRU2, Spambase , User locations Finland ,and Abalone), and two Synthetic datasets (c20d6n2000, and c20d6n200000)	Good technique for addressing high dimensional data clustering issues	Results are impacted by the CRO's parameter settings.
6	[138]	2020	Three improvements to the Bat algorithm are used to give an enhanced clustering method: First, the Gaussian convergence factor and other convergence factors are added to improve global search capability. Next, the hunting mechanism of the whale optimization algorithm is incorporated to improve the local search capability of the bat algorithm. Finally, the sine strategy is added to enhance the updating mechanism of solutions.	Seven UCI datasets are used for validation (Heartstatlog, WDBC, Iris, Wine, Bupa, Seeds, Heartstat, and Wisconsin breast cancer)	Strengthen the global search ability and enhance the accuracy rate of data clustering	The proposed algorithm suffers from unstable searching.
7	[139]	2019	The data clustering issue is resolved using the symbiotic organism search (SOS) technique.◎ For the phases of mutualism and cooperation, novel formulas have been presented.in the parasitism stage, adopted parasite vector	Ten UCI datasets are used from calefornia university(Artificial dataset one, Artificial dataset two, Iris, Breast cancer Wisconsin (Original), Balance scale, Seeds, Statlog (Heart), CMC, Haberman's Survival ,and Wine.)	Superior accuracy and high level of stability.	Complex clustering problems are not addressed
8	[140]	2020	A new clustering method is presented by hybridization gray wolf optimizer (GWO) with Tabu search (TS) strategy. The fundamental principle of the hybrid GWOTS is to first determine each leader's neighborhood before adjusting the positions of the other members of the pack.	Iris ,Blood ,Breast cancer ,Seeds, Wine ,Diabetes , Australian , Haberman ,Heart ,Liver , Planning Relax, and Tic-tac-toe.	results that are ideal and quickly converge	High computational time.
9	[141]	2021	Hybrid algorithm recently created The crossover operator, polygamy (a particular form of elitism), and the PSO principle are presented as components of PSOPC, an efficient clustering algorithm. This hybridization relies on the use of polygamy as a unique form of elitism for crossover to improve the exploration and exploitation approach, as well as PSO as a global search method.	Wine, Haberman, Glass ,Buba ,Iris ,CMC ,and Cancer	convergence speed, and solution optimality	Dynamic data clustering is not handled
10	[142]	2022	In the context of clustering, two adapted firefly algorithms, the crazy firefly method and the variable step size firefly algorithm, are each hybridized with a traditional particle swarm optimization (PSO) technique.	Ten UCI datasets are used(Iris, Wine, Yeast, Thyroid, Hepatitis, Heart ,Glass ,Breast ,Wdbc , and Leaves) , and eight Shape sets (Spiral, Path based, Jain, Flame, Compound, R15, D31, and Aggregation)	Determining the ideal number of clusters and effectively addressing problems with	-

					artificial data clustering	
11	[143]	2022	As a clustering method, an improved Black Hole algorithm (IBH) is suggested. It developed adaptability to various circumstances it would encounter as it progressed. With this upgrade, BH is better able to take use of the algorithm's recent advancements to produce better solutions and avoid becoming stuck in local optima.	Iris ,Glass ,Wine ,Cancer, and CMC	A high convergence speed, simplicity and freedom of hyper-parameters, and high efficiency in context of data clustering.	-
12	[144]	2022	A new hybrid intelligence swarm method called WOATS combines the Tabu Search (TS) and Whale Optimization method (WOA) algorithms. To record the top solutions, WOATS utilized an Elite List (EL) memory component of TS. These options were utilized by WOATS to direct the swarm's members during the search phase. To guarantee the diversity of solutions, WOATS used the crossover operator.	Iris ,Glass , Balance ,Seed ,Mouse, Vary Density, Magic, Electricity, CoverType ,Poker ,Wine ,Cancer ,CMC ,Ecoli , and Survival	High clustering efficiency in analyzing medium and large scale problems.	-
13	[145]	2022	A new automatic clustering method based on enhancing the fundamental effectiveness of the Barnacles Mating Optimizer (BMO) with regard to of convergence trends and stagnation avoidance by integrating the elements of the Sine Cosine Algorithm (SCA) with a disruption operator.	Wine, Iris, Ecoli, Haberman's Survival,Glass, Liver Disorders, IonosphereEW, Lymphography, M-of-n, PenglungEW3, Brain-T21, CongressEW1, KrvskpEW4, Gesture.	Superior solutions, a fair trade-off between exploration and exploitation, and higher convergence rates	No exceptional ability to prevent stagnation and a delayed convergence of the approach while dealing with larger-scale issues
14	[146]	2023	A strong hybrid method with no control parameters is suggested to combine the leaders and followers optimization algorithm (LaF) and differential evolution (DE), balancing exploration and exploitation in optimization-based partitional data clustering.	Glass, Iris, Wine, Yeast, Flame, Jain, R15, D31, Aggregation, Compound, Path-based, Spiral	More effective on datasets with spherical data distributions.	Poorer clustering effectiveness for datasets with u/v distributions of data. Lacking to a fully automated data clustering method because the number of clusters is input into the algorithm as a parameter.
15	[147]	2023	This study introduces MHTSASM, a novel technique that combines K-Means clustering with the Tabu search (TS). The benefits of both the TS and K-Means algorithms are completely utilized in this algorithm. With the use of adaptive search memory ASM, it utilizes TS to create economic data exploration, striking a balance between the intensification and diversification tactics that are employed to improve the search process.	Iris, Glass, Cancer, Contraceptive Method Choice (CMC), Wine, Bavarian postal zones dataset, Germany postal zones dataset, and the Fisher´s iris dataset	High clustering efficiency in solving different clustering problems.	Addressing and detecting clusters with non-convex geometries.

16	[148]	2022	A new hybridization optimization search technique for tough optimization issues. The suggested approach, known as HRSA, combines the original Remora Optimization Algorithm (ROA) and Reptile Search Algorithm (RSA) and manages the search process using a novel transition method. The proposed HRSA technique discovers superior solutions and addresses the primary issues that the original methods raised.	23 benchmark problems and eight data clustering problems (Iris, Glass, Cancer, Contraceptive Method Choice (CMC), Wine, Seeds, Heart, Water, and Vowels) are used for comparison.	Due to the mathematical issues, the suggested strategy produces overwhelmingly favorable results, and a great effectiveness when used to solve different clustering problems.	The usage of real data sets in the future for data clustering is a limitation of this paper. Furthermore, the suggested solution occasionally requires longer execution times.
17	[149]	2022	The Arithmetic Optimization Algorithm (AOA) and Opposition-based Learning method are hybridized with the Flow Direction Algorithm (FDA) to take advantage of the arithmetic operators in AOA to enhance the performance of the Flow Direction Algorithm and get around early convergence, trapping in the local region, and an imbalance between the exploration and exploitation search processes.	In the first set, 23 benchmark functions are used, including 7 unimodal functions, 6 multimodal functions, and 10 fixed dimension functions. Eight typical data clustering problems are used to evaluate the suggested approach.	Address the shortcomings of the original approaches, such as the local search area trap, early convergence, and the search process equilibrium.	-
18	[150]	2022	For issues involving global optimization and tacking data clustering problems, this research paper presents a nebulous LA-based hybrid optimizing technique. The artificial Jellyfish search algorithm (JS) and the Marine Predator Algorithm (MPA) are improved in the suggested approach to lessen their computing complexity while maintaining their benefits. The probability vector of the LA is also enhanced to improve efficiency. Additionally, the roulette wheel selection technique is used to choose the best solution, while the greedy selection method is used to maintain elitism between alternatives.	CMC, Banknote authentication, Glass identification, Iris, Liver disorders, Wine, Breast cancer Coimbra, Divorce predictors, Hepatitis C virus, and Blood Transfusion Service Center	Eliminate the flaws, such as premature converge, local optimal trapping, and sluggish global as well as local search	-
19	[151]	2022	In order to improve the classic arithmetic optimization algorithm (AOA) search mechanism for dealing with global optimization and data clustering, a new variant of AOA is presented in this paper. This new variant uses Lévy Flight distribution opposition-based learning (OLB).	23 benchmark problems and 8 UCI datasets: Cancer, CMC, Glass, Iris, Seeds, Heart, Vowels, and Water are used for research.	Balancing between exploration and exploitation makes the proposed method a potential optimization technique for addressing various global optimization issues and data clustering issues with high dimensions because of the it's high stability.	-
20	[152]	2020	An algorithm that combines Chaos Optimization and Flower Pollination over K-means is developed to increase the	D1, D2, D3, D4, D5, D6, D7, D8, D9, D10, D11, D12, D13, D14, D15, D16. These datasets	superior in terms of convergence level, execution	Stability issues

			effectiveness of minimizing the cluster integrity	have been collected from [68, 153-158]	duration, and cluster integrity	
21	[159]	2019	The memetic particle gravitation optimization (MPGO) algorithm, which relies on PSO and the gravitation search algorithm (GSA), is proposed in this study as a memetic clustering technique with effective search and quick convergence, accordingly. In order to choose the optimum partition for partitioning each pattern with an appropriate clustering centre, MPGO uses hybrid operator and diversification improvement as its two key techniques.	Six UCI machine learning benchmarks (car evaluation, wine, yeast, iris, statlog, breast cancer) are utilized for validation together with 52 benchmark test routines.	Efficacy in terms of accuracy rate, and fitness value.	Computing performance, Population diversity
22	[160]	2019	Two Firefly Algorithm (FA) versions were proposed: (i) internal intensified exploration (IIEFA) and (ii) compound intensified exploration (CIEFA).◎ Incorporating matrix-based search criteria and dispersal methods improves exploration and exploitation.◎ a minimum redundancy policy◎ Method for choosing features for lowering feature dimensions based on Maximum Relevance.	Acute Lymphoblastic Leukaemia (ALL) from ALL-IDB2 database [161], and nine UCI datasets :Wisconsin breast cancer diagnostic (Wbc1), Wisconsin breast cancer original (Wbc2), Wine, Iris, Thyroid, Sonar, Balance, Ecoli, and Ozone,	Addressing issues including multi-dimensional clustering, converging speed, and clustering efficiency in terms of average accuracy rates sensibility, and sensitivity	Dealing with complex and irregular data distribution problems are not handled
23	[162]	2020	The Best Worst Mean Harmony Search (BWM_HS) algorithm, a new variation of the harmony search (HS) algorithm proposed in this paper, makes better use of the useful data kept in the Harmony Memory (HM) to direct the search process. The BWM_HS employs an altered memory concern technique for this reason, replacing the random harmonic selection scheme by three brand-new pitch selections along with production criteria. Using these three principles, the BWM_HS algorithm generates two new harmonies at each iteration, further using the data from the HM.	The algorithm is investigated on CEC 2017 test suite's benchmark functions with 10D, 30D, and 50D, and ten well-known clustering problems (Iris, CMC, Breast cancer, Wine, Vowel, Glass, Aggregation, Balance, D31, R15)	Better outcomes in terms of precision, rapid convergence, adaptability, and time complexity	-
24	[163]	2023	The HSGS algorithm, a new variation of the Harmony Search (HS) algorithm, is suggested in this study as a solution to clustering optimizing issues. The proposed HSGS enhances the exploitation potential of the HS algorithm by utilizing the mathematical aspects of Golden Search (GS). To do this, two new harmonies are created at the end of each HSGS cycle. The canonical HS algorithm's search operators, which value exploration ability, are used to construct the first one. One of the two new phases used by the HSGS to create the second. The exploitation potential of the HSGS increases with a varied ratio based on the phase that is used.	Five datasets concerns with gene expression (YS [164], RCNS [165],AT[166], HFS[167], YCC[168]) , and some of UCI datasets (Iris, Glass, Cancer, Seed, Vowel, Newthyroid, WDBC,A1,A2,A3, D31,R15, Aggregation, Compund, Pathbased, and S2)	Better tradeoff between exploration and exploitation	High computational time
25	[169]	2021	The water cycle algorithm, which depends on the rate of evaporation, is utilized in this study conjointly with the Hookes and Jeeves method, a local search	Fisher Iris, Wisconsin Breast Cancer, Glass	Better convergence as well as improving the optimization results.	-

			technique, to increase the algorithm's robustness and searching performance.			
26	[170]	2019	An alternative hybrid technique called ASOSCA for automatic clustering based on the integration of two types of metaheuristics, namely the sine-cosine algorithm (SCA) and atom search optimization (ASO). ASOSCA uses SCA's operator as a local search strategy for enhancing the convergence speed.	BreastTissu, CMC, Transfusion, Seeds, Balancescale, Wine, Iris, Glass, Ecoli, BreastW, MuskClean1, LiverDis, Vote, Spectew, SonarEW, Vowel.	Robustness and effectiveness in solving data clustering problems	-

2.3.3 |Association Mining Rules

One of the primary tasks of data mining is association rule mining, or ARM. For data mining, it is a crucial task for identifying common patterns. In huge datasets, ARM looks for close correlations between the elements. Finding Association Rules within a huge database is considered an NP-Hard task. So, the processing time required by conventional ARM techniques is substantial. They also rely on data preparation prior to executing the algorithm, which results in information loss. Two other shortcomings of standard ARM approaches are a strong boundary between intervals in numeric characteristics and differentiating the membership degree for the interval in fuzzy sets. High-dimensional spaces are challenging to solve because of their nature. Because standard heuristic methods are unable to produce intricate solutions, metaheuristic algorithms, have become more and more popular. These methods search the issue space using an iterative methodology in order to find a sufficiently effective solution. Metaheuristic algorithms can be used to discover association rules without the need for the frequent itemset generation stage. As a result, computing time is reduced. So, metaheuristics techniques are used for discovering association rules for addressing the conventional ARM algorithms' limitations. Several algorithms are published in this regard as presented in Table 4.

Table 4. An overview of some metaheuristic algorithms for association mining rules.

No	Ref	year	Methodology	Dataset	Advantages	Shortcomings
1	[171]	2020	An association rule mining is proposed based on two steps, a reduction in dimensions by using low-variance and hashing table methods, and fuzzy logic and the whale optimization algorithm are suggested in the second step for item recognition and association rule generation.	Food mart, Chain store, Connect, Mushroom, and Chess	An effective mining strategy that minimizes the amount of memory needed, the time it takes to execute, and the frequency of the items it finds.	-
2	[172]	2022	Multi-objective orthogonal mould algorithm (MOOSMA) with numerical association rule mining (NARM) was presented in this study. The primary goal is based on four effectiveness indicators for each association: support, confidence, comprehension, and interest.	Twenty test functions form CEC09, and ten datasets from Bilkent University's function approximation (libraryBasketball, Bodyfat, Bolts, Longley, Pollution, Pwlinear, Quake, Stock price, Stulong, and Vineyard) are used for assessment.	With this variation, the four association rules of support, confidence, comprehension, and interestingness are maximized.	Fuzzy ARM, and sparse ARM are not handled.

3	[173]	2023	In this study, a new hybrid ARM technique based on Levy flight and water wave optimization (LWWO) is suggested. Three variants of the suggested LWWO are made for ARM by integrating it with the three algorithms ant colony, bat algorithm, and cuckoo search. In order to maximize the global optimal solution throughout the search process, these algorithms mix the search tactics of many algorithms.	The datasets (Iris, Heart, Ecoli, Breast, Flare, and Led7) were downloaded from "LUCS-KDD Discretised/Normalized" database	It is ideal for practical applications by increasing the effectiveness of mining and keeping an adequate equilibrium between rule qualities with mining efficiency.	The particular connections between the items in the derived association rules are not examined.
4	[174]	2020	This paper proposes novel hybrid multi-objective evolutionary optimization techniques based on differential evolution and sine cosine algorithm for rapidly mining the reduced high-quality numerical association rules by simultaneously altering appropriate intervals of associated attributes without discovering the frequent itemsets.	Ailerons, Bodyfat, Bolts, Elevators, House_16h, Quake, Stulong, Longley, Pollution, and Vineyard	High stability, and works effectively with datasets that contain various types of characteristics, including nominal, binary, discrete, and numerical.	The suggested algorithm's parameters have to be established beforehand, and the high level of complexity of algorithmic
5	[175]	2021	The privacy-protected ARM is demonstrated in this study using a constraint-based objective function and the GA in two parts. The association rules in the database are first mined using the FP-Growth algorithm, and then the privacy-saved ARM is carried out by the GA in the second phase.	Two datasets are used (T10I4D100K, and retail)	Top-notch rules	Loss of data is the primary drawback.
6	[176]	2021	Using an innovative Biogeography Based Optimization (BBO) algorithm, this study provides an efficient rule-based technique that predicts credit risk. This is used to find the best rule set including the highest level of predictive precision of a dataset with both categories and continuous features.	German and Australian credit datasets are utilized.	High precision and a straightforward method	Only handle single-objective issues
7	[177]	2021	This research suggests a novel strategy based on association rule mining and artificial immune systems. The suggested method provides the best interpretation of the desired word according to context in addition to indicating the existence of a lexically ambiguous phrase in the written content.	The method was used on a corpus of publicly accessible documents.	Outstanding efficiency and accuracy	The quantity of iterations is strongly connected with accuracy.
8	[178]	2019	In this study, a multi-objective particle swarm optimizer (MOPSO) algorithm with a discretization method for mining numerical association rules is proposed.	Basketball, Body Fat, and Quake.	No data preparation is necessary as the optimal time between datasets is automatically determined.	Less targets are taken in consideration
9	[179]	2020	This study proposes a grey wolf optimizer algorithm, an	Chess, Mushroom, Accident, and Connect	Minimal delay complexity	Excessive memory requirements

			optimization approach founded on the natural behavior of the grey wolf, which is utilized to address the problem of mining high utility itemset utilizing five separate Boolean methods.			
10	[180]	2020	This work suggests an improved binary Artificial Bee Colony (IBABC) technique that strikes an equitable equilibrium between exploration and exploitation for a novel rule-hiding mechanism. To choose transactions that are sensitive to change, the suggested rule-hiding algorithm is combined with the IBABC method.	Multidimensional knapsack problems, Chess, Retail, BMS-1, BMS-2, and Mushroom	Very effective, stable, and capable of addressing multifaceted problems.	high level of complexity in computing

3 | Conclusions and Future Work

Data mining must advance in order to analyze the massive amount of data effectively and extract insights from it. The areas, in which data mining is used, are also continuously expanding. Finding robust algorithms that can be applied to a wide range of tasks without or with minor modifications are therefore an indispensable step. Metaheuristics are considered strong techniques that could find acceptable solutions for several complex optimization problems in a reasonable amount of time. Therefore, the researchers have directed their attention to those algorithms for solving the data mining tasks more accurately and quickly. The No-Free-Lunch (NFL) theorem states that no heuristic is sufficient to address every optimization issue. The majority of metaheuristics excel in at least one particular domain. However, adopting a single metaheuristic to discover the optimal solution across many domains is not always guaranteed. Given these problems, it remains possible to use new metaheuristics on various data mining tasks to provide better outcomes. Each year, dozens of new studies using metaheuristics for feature selection, clustering, and association rules are published, producing solutions of outstanding quality. Many researchers are drawn to this fervent curiosity. Even with massive datasets, the reported outcomes of these methods are astounding. In this regard, we examined the research of eminent academics whose work has gotten numerous citations and whose articles have appeared in prestigious publications and conferences. - Future work will examine how one or more metaheuristics can be used with data mining to solve one or more of the aforementioned concerns.

Acknowledgments

The author is grateful to the editorial and reviewers, as well as the correspondent author, who offered assistance in the form of advice, assessment, and checking during the study period.

Author Contributions

All authors contributed equally to this work.

Funding

This research has no funding source.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to the privacy-preserving nature of the data but are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Dash, M. and H. Liu, Feature selection for classification. Intelligent data analysis, 1997. 1(1-4): p. 131-156.

[2] Mannila, H. Data mining: machine learning, statistics, and databases. in Proceedings of 8th International Conference on Scientific and Statistical Data Base Management. 1996. IEEE.

[3] Arora, R.K. and M.K. Gupta, e-Governance using data warehousing and data mining. International Journal of Computer Applications, 2017. 169(8): p. 975-8887.

[4] Vercellis, C., Business intelligence: data mining and optimization for decision making. 2011: John Wiley & Sons.

[5] Munoz-Gama, J., et al., Process mining for healthcare: Characteristics and challenges. Journal of Biomedical Informatics, 2022. 127: p. 103994.

[6] Tang, X., et al., Image pattern recognition combined with data mining for diagnosis and detection of myocardial infarction. IEEE Access, 2020. 8: p. 146085-146092.

[7] Jayasri, N. and R. Aruna, Big data analytics in health care by data mining and classification techniques. ICT Express, 2022. 8(2): p. 250-257.

[8] Santos-Pereira, J., L. Gruenwald, and J. Bernardino, Top data mining tools for the healthcare industry. Journal of King Saud University-Computer and Information Sciences, 2022. 34(8): p. 4968-4982.

[9] Romero, C. and S. Ventura, Data mining in education. Wiley Interdisciplinary Reviews: Data mining and knowledge discovery, 2013. 3(1): p. 12-27.

[10] Shafiq, D.A., et al., Student Retention Using Educational Data Mining and Predictive Analytics: A Systematic Literature Review. IEEE Access, 2022.

[11] Schweidtmann, A.M., et al., Machine learning in chemical engineering: A perspective. Chemie Ingenieur Technik, 2021. 93(12): p. 2029-2039.

[12] Farrokhi, A., S. Farzin, and S.-F. Mousavi, Meteorological drought analysis in response to climate change conditions, based on combined four-dimensional vine copulas and data mining (VC-DM). Journal of Hydrology, 2021. 603: p. 127135.

[13] Hannachi, A., Patterns identification and data mining in weather and climate. Vol. 10. 2021: Springer.

[14] Pramanik, R., S. Khare, and M.K. Gourisaria. Inferring the Occurrence of Chronic Kidney Failure: A Data Mining Solution. in Proceedings of Second Doctoral Symposium on Computational Intelligence: DoSCI 2021. 2022. Springer.

[15] Khalaf, L.I., et al. Survey On Recognition Hand Gesture By Using Data Mining Algorithms. in 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). 2022. IEEE.

[16] Leung, C.K., et al. Revealing COVID-19 data by data mining and visualization. in Advances in Intelligent Networking and Collaborative Systems: The 13th International Conference on Intelligent Networking and Collaborative Systems (INCoS-2021) 13. 2022. Springer.

[17] Zahra, S.R., et al., Detecting Covid-19 chaos driven phishing/malicious URL attacks by a fuzzy logic and data mining based intelligence system. Egyptian Informatics Journal, 2022. 23(2): p. 197-214.

[18] Salman, A.D., H.A.D. AL-farttoosi, and A.J. Kadhim. Study impact the latitude on Covid-19 spread virus by data mining algorithm. in Journal of Physics: Conference Series. 2020. IOP Publishing.

[19] Iglesias, E.P. and R. Badiang, Predicting the Assessment Course Performance of Criminology Students Using Data Mining. International Journal, 2023. 12(1).

[20] Dutta, S., et al. Application of social networks and data mining on crime victims. in Proceedings of International Conference on Advanced Computing Applications: ICACA 2021. 2022. Springer.

[21] Thota, L.S., et al. Demarcation of Indian Child Crimes using Data Mining Approach. in 2022 International Conference for Advancement in Technology (ICONAT). 2022. IEEE.

[22] Lin, M., Innovative risk early warning model under data mining approach in risk assessment of internet credit finance. Computational Economics, 2022. 59(4): p. 1443-1464.

[23] Ledhem, M.A., Data mining techniques for predicting the financial performance of Islamic banking in Indonesia. Journal of Modelling in Management, 2022. 17(3): p. 896-915.

[24] Al-Hashedi, K.G. and P. Magalingam, Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review, 2021. 40: p. 100402.

[25] Li, T. and C. Zhang, Research on the application of multimedia entropy method in data mining of retail business. Scientific Programming, 2022. 2022.

[26] Saura, J.R., Using data sciences in digital marketing: Framework, methods, and performance metrics. Journal of Innovation & Knowledge, 2021. 6(2): p. 92-102.
[27] Suma, V. and S.M. Hills, Data mining based prediction of demand in Indian market for refurbished electronics. Journal of Soft Computing Paradigm (JSCP), 2020. 2(02): p. 101-110.
[28] Issad, H.A., R. Aoudjit, and J.J. Rodrigues, A comprehensive review of Data Mining techniques in smart agriculture. Engineering in Agriculture, Environment and Food, 2019. 12(4): p. 511-525.
[29] Navghare, A., et al., Analysis of Agriculture Data using Data Mining Techniques. 2020.
[30] Ghorbani, R. and R. Ghousi, Predictive data mining approaches in medical diagnosis: A review of some diseases prediction. International Journal of Data and Network Science, 2019. 3(2): p. 47-70.
[31] Abdel-Basset, M., L. Abdel-Fatah, and A.K. Sangaiah, Metaheuristic algorithms: A comprehensive review. Computational intelligence for multimedia big data on the cloud with engineering applications, 2018: p. 185-231.
[32] Abdel-Basset, M., et al., Exponential distribution optimizer (EDO): a novel math-inspired algorithm for global optimization and engineering problems. Artificial Intelligence Review, 2023: p. 1-72.
[33] Zhao, F., et al., An ensemble discrete differential evolution for the distributed blocking flowshop scheduling with minimizing makespan criterion. Expert Systems with Applications, 2020. 160: p. 113678.
[34] Xiong, G., et al., Parameter extraction of solar photovoltaic models using an improved whale optimization algorithm. Energy conversion and management, 2018. 174: p. 388-405.
[35] Yıldız, A.R., et al., A new hybrid Harris hawks-Nelder-Mead optimization algorithm for solving design and manufacturing problems. Materials Testing, 2019. 61(8): p. 735-743.
[36] Cui, Z., et al., Hybrid many-objective particle swarm optimization algorithm for green coal production problem. Information Sciences, 2020. 518: p. 256-271.
[37] Alweshah, M., et al., The monarch butterfly optimization algorithm for solving feature selection problems. Neural Computing and Applications, 2020: p. 1-15.
[38] Naderi, E., M. Pourakbari-Kasmaei, and H. Abdi, An efficient particle swarm optimization algorithm to solve optimal power flow problem integrated with FACTS devices. Applied Soft Computing, 2019. 80: p. 243-262.
[39] Shayanfar, H. and F.S. Gharehchopogh, Farmland fertility: A new metaheuristic algorithm for solving continuous optimization problems. Applied Soft Computing, 2018. 71: p. 728-746.
[40] Eusuff, M., K. Lansey, and F. Pasha, Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization. Engineering optimization, 2006. 38(2): p. 129-154.
[41] Rostami, M., et al., Integration of multi-objective PSO based feature selection and node centrality for medical datasets. Genomics, 2020. 112(6): p. 4370-4384.
[42] Cadenas, J.M., M.C. Garrido, and R. MartíNez, Feature subset selection filter–wrapper based on low quality data. Expert systems with applications, 2013. 40(16): p. 6241-6252.
[43] Wang, H., et al., A factor graph model for unsupervised feature selection. Information Sciences, 2019. 480: p. 144-159.
[44] Xu, Y., D. Zhang, and J.-Y. Yang, A feature extraction method for use with bimodal biometrics. Pattern recognition, 2010. 43(3): p. 1106-1115.
[45] Farahat, A.K., A. Ghodsi, and M.S. Kamel, Efficient greedy feature selection for unsupervised learning. Knowledge and information systems, 2013. 35: p. 285-310.
[46] Li, J., et al., Exclusive feature selection and multi-view learning for Alzheimer's Disease. Journal of Visual Communication and Image Representation, 2019. 64: p. 102605.
[47] Renuka Devi, D. and S. Sasikala, Online feature selection (OFS) with accelerated bat algorithm (ABA) and ensemble incremental deep multiple layer perceptron (EIDMLP) for big data streams. Journal of Big Data, 2019. 6(1): p. 1-20.
[48] Solorio-Fernández, S., J.A. Carrasco-Ochoa, and J.F. Martínez-Trinidad, A review of unsupervised feature selection methods. Artificial Intelligence Review, 2020. 53(2): p. 907-948.
[49] Ding, D., et al., Unsupervised feature selection via adaptive hypergraph regularized latent representation learning. Neurocomputing, 2020. 378: p. 79-97.
[50] Ghosh, A., A. Datta, and S. Ghosh, Self-adaptive differential evolution for feature selection in hyperspectral image data. Applied Soft Computing, 2013. 13(4): p. 1969-1977.
[51] Xu, Z., et al., Discriminative semi-supervised feature selection via manifold regularization. IEEE Transactions on Neural networks, 2010. 21(7): p. 1033-1047.
[52] Tang, J., S. Alelyani, and H. Liu, Feature selection for classification: A review. Data classification: Algorithms and applications, 2014: p. 37.
[53] Liu, H. and H. Motoda, Feature extraction, construction and selection: A data mining perspective. Vol. 453. 1998: Springer Science & Business Media.
[54] Kohavi, R. and G.H. John, Wrappers for feature subset selection. Artificial intelligence, 1997. 97(1-2): p. 273-324.
[55] Miao, J. and L. Niu, A survey on feature selection. Procedia Computer Science, 2016. 91: p. 919-926.
[56] Deniz, A., et al., Robust multiobjective evolutionary feature subset selection algorithm for binary classification using machine learning techniques. Neurocomputing, 2017. 241: p. 128-146.
[57] Emary, E., H.M. Zawbaa, and A.E. Hassanien, Binary ant lion approaches for feature selection. Neurocomputing, 2016. 213: p. 54-65.

[58] Zhang, S., et al., Swarm intelligence applied in green logistics: A literature review. Engineering Applications of Artificial Intelligence, 2015. 37: p. 154-169.

[59] Hancer, E., B. Xue, and M. Zhang, Fuzzy filter cost-sensitive feature selection with differential evolution. Knowledge-Based Systems, 2022. 241: p. 108259.

[60] Ain, Q.U., et al., Genetic programming for automatic skin cancer image classification. Expert Systems with Applications, 2022. 197: p. 116680.

[61] Barak, S., J.H. Dahooie, and T. Tichý, Wrapper ANFIS-ICA method to do stock market timing and feature selection on the basis of Japanese Candlestick. Expert Systems with Applications, 2015. 42(23): p. 9221-9235.

[62] Dokeroglu, T., A. Deniz, and H.E. Kiziloz, A comprehensive survey on recent metaheuristics for feature selection. Neurocomputing, 2022.

[63] Hichem, H., et al., A new binary grasshopper optimization algorithm for feature selection problem. Journal of King Saud University-Computer and Information Sciences, 2022. 34(2): p. 316-328.

[64] Dua, D. and C. Graff, UCI machine learning repository. 2017.

[65] Agrawal, R., B. Kaur, and S. Sharma, Quantum based whale optimization algorithm for wrapper feature selection. Applied Soft Computing, 2020. 89: p. 106092.

[66] Long, W., et al., Pinhole-imaging-based learning butterfly optimization algorithm for global optimization and feature selection. Applied Soft Computing, 2021. 103: p. 107146.

[67] Albashish, D., et al., Binary biogeography-based optimization based SVM-RFE for feature selection. Applied Soft Computing, 2021. 101: p. 107026.

[68] Blake, C., UCI repository of machine learning databases. http://www. ics. uci. edu/~ mlearn/MLRepository. html, 1998.

[69] Song, X.-f., et al., Feature selection using bare-bones particle swarm optimization with mutual information. Pattern Recognition, 2021. 112: p. 107804.

[70] Dua, D. and C. Graff, UCI machine learning repository [http://archive. ics. uci. edu/ml]. Irvine, CA: University of California, School of Information and Computer Science. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.

[71] Mohammad, R., F. Thabtah, and T. McCluskey, Phishing websites dataset. 2015.

[72] Tan, C.L., Phishing dataset for machine learning: Feature evaluation. Mendeley Data, 2018. 1: p. 2018.

[73] Xu, Z., et al., Enhanced Gaussian bare-bones grasshopper optimization: Mitigating the performance concerns for feature selection. Expert Systems with Applications, 2023. 212: p. 118642.

[74] Pietruszkiewicz, W., Application of discrete predicting structures in an early warning expert system for financial distress. 2004, Ph. d. thesis, Faculty of Computer Science and Information Technology ….

[75] Chaudhuri, A. and T.P. Sahu, Feature selection using Binary Crow Search Algorithm with time varying flight length. Expert Systems with Applications, 2021. 168: p. 114288.

[76] Ma, W., et al., A two-stage hybrid ant colony optimization for high-dimensional feature selection. Pattern Recognition, 2021. 116: p. 107933.

[77] Tabakhi, S. and P. Moradi, Relevance–redundancy feature selection based on ant colony optimization. Pattern recognition, 2015. 48(9): p. 2798-2811.

[78] Tran, B., B. Xue, and M. Zhang, Variable-length particle swarm optimization for feature selection on high-dimensional classification. IEEE Transactions on Evolutionary Computation, 2018. 23(3): p. 473-487.

[79] Pourpanah, F., et al., A hybrid model of fuzzy min–max and brain storm optimization for feature selection and data classification. Neurocomputing, 2019. 333: p. 440-451.

[80] Lichman, M., UCI machine learning repository. 2013, Irvine, CA, USA

[81] Chaudhuri, A. and T.P. Sahu, A hybrid feature selection method based on Binary Jaya algorithm for micro-array data classification. Computers & Electrical Engineering, 2021. 90: p. 106963.

[82] KP, M.N. and P. Thiyagarajan, Feature selection using efficient fusion of Fisher Score and greedy searching for Alzheimer's classification. Journal of King Saud University-Computer and Information Sciences, 2022. 34(8): p. 4993-5006.

[83] Zimmermann, J., et al., Differentiation of Alzheimer's disease based on local and global parameters in personalized Virtual Brain models. NeuroImage: Clinical, 2018. 19: p. 240-251.

[84] Ellis, K.A., et al., The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. International psychogeriatrics, 2009. 21(4): p. 672-687.

[85] Wang, D., et al., A novel quantum grasshopper optimization algorithm for feature selection. International Journal of Approximate Reasoning, 2020. 127: p. 33-53.

[86] Blake, C.L. and C.J. Merz, UCI repository of machine learning databases, 1998. 1998.

[87] Paniri, M., M.B. Dowlatshahi, and H. Nezamabadi-Pour, MLACO: A multi-label feature selection algorithm based on ant colony optimization. Knowledge-Based Systems, 2020. 192: p. 105285.

[88] Sahebi, G., et al., GeFeS: A generalized wrapper feature selection approach for optimizing classification performance. Computers in biology and medicine, 2020. 125: p. 103974.

[89] Al-Tashi, Q., et al., Binary optimization using hybrid grey wolf optimization for feature selection. Ieee Access, 2019. 7: p. 39496-39508.

[90] Merz, C.J., UCI repository of machine learning databases. URL: http://www. ics. uci. edu/~ mlearn/MLRepository. html, 1998.

[91] Hu, P., J.-S. Pan, and S.-C. Chu, Improved binary grey wolf optimizer and its application for feature selection. Knowledge-Based Systems, 2020. 195: p. 105746.
[92] Asuncion, A. and D. Newman, UCI machine learning repository. 2007, Irvine, CA, USA.
[93] Abdel-Basset, M., W. Ding, and D. El-Shahat, A hybrid Harris Hawks optimization algorithm with simulated annealing for feature selection. Artificial Intelligence Review, 2021. 54: p. 593-637.
[94] Zhao, Y., et al., A binary dandelion algorithm using seeding and chaos population strategies for feature selection. Applied Soft Computing, 2022. 125: p. 109166.
[95] Gharehchopogh, F.S., I. Maleki, and Z.A. Dizaji, Chaotic vortex search algorithm: metaheuristic algorithm for feature selection. Evolutionary Intelligence, 2022. 15(3): p. 1777-1808.
[96] Alweshah, M., et al., Coronavirus herd immunity optimizer with greedy crossover for feature selection in medical diagnosis. Knowledge-Based Systems, 2022. 235: p. 107629.
[97] de Souza, R.C.T., et al., Binary coyote optimization algorithm for feature selection. Pattern Recognition, 2020. 107: p. 107470.
[98] Too, J. and S. Mirjalili, A hyper learning binary dragonfly algorithm for feature selection: A COVID-19 case study. Knowledge-Based Systems, 2021. 212: p. 106553.
[99] Nadimi-Shahraki, M.H., H. Zamani, and S. Mirjalili, Enhanced whale optimization algorithm for medical feature selection: A COVID-19 case study. Computers in Biology and Medicine, 2022. 148: p. 105858.
[100] Hegazy, A.E., M. Makhlouf, and G.S. El-Tawel, Improved salp swarm algorithm for feature selection. Journal of King Saud University-Computer and Information Sciences, 2020. 32(3): p. 335-344
[101] Dua, D. and C. Graff, UCI machine learning repository, 2017. URL http://archive. ics. uci. edu/ml, 2017. 7(1).
[102] Tubishat, M., et al., Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection. Expert Systems with Applications, 2020. 145: p. 113122.
[103] Turkoglu, B., S.A. Uymaz, and E. Kaya, Binary artificial algae algorithm for feature selection. Applied Soft Computing, 2022. 120: p. 108630.
[104] Dataset, A.R., URL https://archive. ics. uci. edu/ml/datasets. REALDISP+ Activity+ Recognition+ Dataset, 2014.
[105] Abualigah, L. and A. Diabat, Chaotic binary group search optimizer for feature selection. Expert Systems with Applications, 2022. 192: p. 116368.
[106] Agushaka, J.O., et al., A novel binary greater cane rat algorithm for feature selection. Results in Control and Optimization, 2023: p. 100225.
[107] Mostafa, R.R., et al., An improved gorilla troops optimizer for global optimization problems and feature selection. Knowledge-Based Systems, 2023: p. 110462.
[108] Zhong, C., et al., A self-adaptive quantum equilibrium optimizer with artificial bee colony for feature selection. Computers in Biology and Medicine, 2023: p. 106520.
[109] Xu, B., et al., Dimensional decision covariance colony predation algorithm: global optimization and high− dimensional feature selection. Artificial Intelligence Review, 2023: p. 1-57.
[110] Tahir, M., et al., A novel binary chaotic genetic algorithm for feature selection and its utility in affective computing and healthcare. Neural Computing and Applications, 2020: p. 1-22.
[111] Mohammadi, A., et al., Efficient deep steering control method for self-driving cars through feature density metric. Neurocomputing, 2023. 515: p. 107-120.
[112] Birchler, C., et al., Single and multi-objective test cases prioritization for self-driving cars in virtual environments. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2022.
[113] Almutairi, M.S., K. Almutairi, and H. Chiroma, Selecting Features That Influence Vehicle Collisions in the Internet of Vehicles Based on a Multi-Objective Hybrid Bi-Directional NSGA-III. Applied Sciences, 2023. 13(4): p. 2064.
[114] Qureshi, K.N., et al., Nature-inspired algorithm-based secure data dissemination framework for smart city networks. Neural Computing and Applications, 2021. 33: p. 10637-10656.
[115] Mafarja, M., et al., Augmented whale feature selection for IoT attacks: Structure, analysis and applications. Future Generation Computer Systems, 2020. 112: p. 18-40.
[116] He, B., et al., Estimate soil moisture of maize by combining support vector machine and chaotic whale optimization algorithm. Agricultural Water Management, 2022. 267: p. 107618.
[117] Hussain, N., et al., Multiclass cucumber leaf diseases recognition using best feature selection. Comput. Mater. Contin, 2022. 70(2): p. 3281-3294.
[118] Simumba, N., et al., Multiple objective metaheuristics for feature selection based on stakeholder requirements in credit scoring. Decision Support Systems, 2022. 155: p. 113714.
[119] Venkateswarlu, Y., et al., An Efficient Outlier Detection with Deep Learning-Based Financial Crisis Prediction Model in Big Data Environment. Computational Intelligence and Neuroscience, 2022. 2022.
[120] Han, S., et al., Competition-driven multimodal multiobjective optimization and its application to feature selection for credit card fraud detection. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022. 52(12): p. 7845-7857.
[121] Ileberi, E., Y. Sun, and Z. Wang, A machine learning based credit card fraud detection using the GA algorithm for feature selection. Journal of Big Data, 2022. 9(1): p. 1-17.
[122] Jacob, W.S., Multi-objective genetic algorithm and CNN-based deep learning architectural scheme for effective spam detection. International Journal of Intelligent Networks, 2022. 3: p. 9-15.

[123] Sharaff, A., et al., Spam message detection using Danger theory and Krill herd optimization. Computer Networks, 2021. 199: p. 108453.

[124] Samee, N.A., et al., Metaheuristic optimization through deep learning classification of COVID-19 in chest X-ray images. Computers, Materials and Continua, 2022: p. 4193-4210.

[125] Canayaz, M., MH-COVIDNet: Diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images. Biomedical Signal Processing and Control, 2021. 64: p. 102257.

[126] Manohar, K. and E. Logashanmugam, Hybrid deep learning with optimal feature selection for speech emotion recognition using improved meta-heuristic algorithm. Knowledge-Based Systems, 2022. 246: p. 108659.

[127] Dey, A., et al., A hybrid meta-heuristic feature selection method using golden ratio and equilibrium optimization algorithms for speech emotion recognition. IEEE Access, 2020. 8: p. 200953-200970.

[128] Alyasiri, O.M., et al., Wrapper and Hybrid Feature Selection Methods Using Metaheuristic Algorithms for English Text Classification: A Systematic Review. IEEE Access, 2022.

[129] Abiodun, E.O., et al., A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities. Neural Computing and Applications, 2021. 33(22): p. 15091-15118.

[130] Jain, A.K., M.N. Murty, and P.J. Flynn, Data clustering: a review. ACM computing surveys (CSUR), 1999. 31(3): p. 264-323.

[131] Rokach, L. and O. Maimon, Clustering methods. 2005.

[132] Aliniya, Z. and S.A. Mirroshandel, A novel combinatorial merge-split approach for automatic clustering using imperialist competitive algorithm. Expert Systems with Applications, 2019. 117: p. 243-266.

[133] Singh, T., A chaotic sequence-guided Harris hawks optimizer for data clustering. Neural Computing and Applications, 2020. 32: p. 17789-17803.

[134] Mustafa, H.M., et al., An improved adaptive memetic differential evolution optimization algorithms for data clustering problems. PloS one, 2019. 14(5): p. e0216906.

[135] Kushwaha, N., M. Pant, and S. Sharma, Electromagnetic optimization-based clustering algorithm. Expert Systems, 2022. 39(7): p. e12491.

[136] Mageshkumar, C., S. Karthik, and V. Arunachalam, Hybrid metaheuristic algorithm for improving the efficiency of data clustering. Cluster Computing, 2019. 22: p. 435-442.

[137] Tsai, C.-W., et al., A high-performance parallel coral reef optimization for data clustering. Soft computing, 2019. 23: p. 9327-9340.

[138] Zhu, L., et al., Data clustering method based on improved bat algorithm with six convergence factors and local search operators. IEEE Access, 2020. 8: p. 80536-80560.

[139] Zhou, Y., et al., Automatic data clustering using nature-inspired symbiotic organism search algorithm. Knowledge-Based Systems, 2019. 163: p. 546-557.

[140] Aljarah, I., et al., Clustering analysis using a novel locality-informed grey wolf-inspired clustering approach. Knowledge and Information Systems, 2020. 62: p. 507-539.

[141] Sharma, M. and J.K. Chhabra, An efficient hybrid PSO polygamous crossover based clustering algorithm. Evolutionary Intelligence, 2021. 14(3): p. 1213-1231.

[142] Behera, M., et al., Automatic Data Clustering by Hybrid Enhanced Firefly and Particle Swarm Optimization Algorithms. Mathematics, 2022. 10(19): p. 3532.

[143] Deeb, H., et al., Improved Black Hole optimization algorithm for data clustering. Journal of King Saud University-Computer and Information Sciences, 2022. 34(8): p. 5020-5029.

[144] Ghany, K.K.A., et al., A hybrid modified step whale optimization algorithm with tabu search for data clustering. Journal of King Saud University-Computer and Information Sciences, 2022. 34(3): p. 832-839.

[145] Abd Elaziz, M., et al., Sine–Cosine-Barnacles Algorithm Optimizer with disruption operator for global optimization and automatic data clustering. Expert Systems with Applications, 2022. 207: p. 117993.

[146] Zorarpacı, E., Data clustering using leaders and followers optimization and differential evolution. Applied Soft Computing, 2023. 132: p. 109838.

[147] Alotaibi, Y., A new meta-heuristics data clustering algorithm based on tabu search and adaptive search memory. Symmetry, 2022. 14(3): p. 623.

[148] Almotairi, K.H. and L. Abualigah, Hybrid reptile search algorithm and remora optimization algorithm for optimization tasks and data clustering. Symmetry, 2022. 14(3): p. 458.

[149] Abualigah, L., et al., Enhanced Flow Direction Arithmetic Optimization Algorithm for mathematical optimization problems with applications of data clustering. Engineering Analysis with Boundary Elements, 2022. 138: p. 13-29.

[150] Barshandeh, S., R. Dana, and P. Eskandarian, A learning automata-based hybrid MPA and JS algorithm for numerical optimization problems and its application on data clustering. Knowledge-Based Systems, 2022. 236: p. 107682

[151] Abualigah, L., et al., Augmented arithmetic optimization algorithm using opposite-based learning and lévy flight distribution for global optimization and data clustering. Journal of Intelligent Manufacturing, 2022: p. 1-39.

[152] Kaur, A., S.K. Pal, and A.P. Singh, Hybridization of Chaos and Flower Pollination Algorithm over K-Means for data clustering. Applied Soft Computing, 2020. 97: p. 105523.

[153] Senthilnath, J., S. Omkar, and V. Mani, Clustering using firefly algorithm: performance study. Swarm and Evolutionary Computation, 2011. 1(3): p. 164-171.

[154] Younsi, R. and W. Wang. A new artificial immune system algorithm for clustering. in Intelligent Data Engineering and Automated Learning–IDEAL 2004: 5th International Conference, Exeter, UK. August 25-27, 2004. Proceedings 5. 2004. Springer.
[155] Karaboga, D. and C. Ozturk, A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Applied soft computing, 2011. 11(1): p. 652-657.
[156] Binu, D., M. Selvi, and A. George, MKF-cuckoo: hybridization of cuckoo search and multiple kernel-based fuzzy C-means algorithm. AASRI Procedia, 2013. 4: p. 243-249.
[157] Elfarra, B.K., T.J. El Khateeb, and W.M. Ashour, BH-centroids: A new efficient clustering algorithm. work, 2013. 1(1).
[158] Tang, R., et al. Integrating nature-inspired optimization algorithms to K-means clustering. in Seventh International Conference on Digital Information Management (ICDIM 2012). 2012. IEEE.
[159] Huang, K.-W., et al., Memetic particle gravitation optimization algorithm for solving clustering problems. Ieee Access, 2019. 7: p. 80950-80968.
[160] Xie, H., et al., Improving K-means clustering with enhanced Firefly Algorithms. Applied Soft Computing, 2019. 84: p. 105763.
[161] Labati, R.D., V. Piuri, and F. Scotti. All-IDB: The acute lymphoblastic leukemia image database for image processing. in 2011 18th IEEE international conference on image processing. 2011. IEEE.
[162] Talaei, K., A. Rahati, and L. Idoumghar, A novel harmony search algorithm and its application to data clustering. Applied Soft Computing, 2020. 92: p. 106273.
[163] Talaei, K., A. Rahati, and L. Idoumghar, HSGS: A hybrid of harmony search algorithm and golden section for data clustering. Expert Systems with Applications, 2023. 224: p. 119954.
[164] Chu, S., et al., The transcriptional program of sporulation in budding yeast. Science, 1998. 282(5389): p. 699-705.
[165] Wen, X., et al., Large-scale temporal gene expression mapping of central nervous system development. Proceedings of the National Academy of Sciences, 1998. 95(1): p. 334-339.
[166] Reymond, P., et al., Differential gene expression in response to mechanical wounding and insect feeding in Arabidopsis. The Plant Cell, 2000. 12(5): p. 707-719.
[167] Iyer, V.R., et al., The transcriptional program in the response of human fibroblasts to serum. science, 1999. 283(5398): p. 83-87.
[168] Cho, R.J., et al., A genome-wide transcriptional analysis of the mitotic cell cycle. Molecular cell, 1998. 2(1): p. 65-73.
[169] Taib, H. and A. Bahreininejad, Data clustering using hybrid water cycle algorithm and a local pattern search method. Advances in Engineering Software, 2021. 153: p. 102961.
[170] Abd Elaziz, M., et al. Automatic data clustering based on hybrid atom search optimization and sine-cosine algorithm. in 2019 IEEE congress on evolutionary computation (CEC). 2019. IEEE.
[171] Sharmila, S. and S. Vijayarani, Association rule mining using fuzzy logic and whale optimization algorithm. Soft Computing, 2021. 25: p. 1431-1446.
[172] Yacoubi, S., et al., A modified multi-objective slime mould algorithm with orthogonal learning for numerical association rules mining. Neural Computing and Applications, 2023. 35(8): p. 6125-6151.
[173] He, Q., et al., Association Rule Mining through Combining Hybrid Water Wave Optimization Algorithm with Levy Flight. Mathematics, 2023. 11(5): p. 1195.
[174] Altay, E.V. and B. Alatas, Differential evolution and sine cosine algorithm based novel hybrid multi-objective approaches for numerical association rule mining. Information Sciences, 2021. 554: p. 198-221.
[175] Menaga, D. and S. Saravanan, GA-PPARM: constraint-based objective function and genetic algorithm for privacy preserved association rule mining. Evolutionary Intelligence, 2021: p. 1-12.
[176] Giri, P.K., et al., Biogeography based optimization for mining rules to assess credit risk. Intelligent Systems in Accounting, Finance and Management, 2021. 28(1): p. 35-51.
[177] Husain, M.S., Exploiting Artificial Immune System to Optimize Association Rules for Word Sense Disambiguation. International Journal of Intelligent Systems and Applications in Engineering, 2021. 9(4): p. 184-190.
[178] Kuo, R.-J., M. Gosumolo, and F.E. Zulvia, Multi-objective particle swarm optimization algorithm using adaptive archive grid for numerical association rule mining. Neural Computing and Applications, 2019. 31: p. 3559-3572.
[179] Pazhaniraja, N., S. Sountharrajan, and B. Sathis Kumar, High utility itemset mining: a Boolean operators-based modified grey wolf optimization algorithm. Soft Computing, 2020. 24: p. 16691-16704.
[180] Telikani, A., et al., Privacy-preserving in association rule mining using an improved discrete binary artificial bee colony. Expert Systems with Applications, 2020. 144: p. 113097.