MDPI

Literature Review on Hybrid Evolutionary Approaches for Feature Selection

Jayashree Piri [1]**, Puspanjali Mohapatra** [2]**, Raghunath Dey** [3]**, Biswaranjan Acharya** [4,]*****,
Vassilis C. Gerogiannis [5] **and Andreas Kanavos** [6,]*****

1   Department of CSE, GITAM Institute of Technology (Deemed to be University), Visakhapatnam 530045, India
2   International Institute of Information Technology, Bhubaneswar 751003, India
3   School of Computer Engineering, KIIT (Deemed to be University), Bhubaneswar 751024, India
4   Department of Computer Engineering-AI, Marwadi University, Rajkot 360003, India
5   Department of Digital Systems, University of Thessaly, 382 21 Larissa, Greece
6   Department of Informatics, Ionian University, 491 00 Corfu, Greece
*   Correspondence: biswaranjan.acharya@marwadieducation.edu.in (B.A.); akanavos@ionio.gr (A.K.)

Abstract: The efficiency and the effectiveness of a machine learning (ML) model are greatly influenced by feature selection (FS), a crucial preprocessing step in machine learning that seeks out the ideal set of characteristics with the maximum accuracy possible. Due to their dominance over traditional optimization techniques, researchers are concentrating on a variety of metaheuristic (or evolutionary) algorithms and trying to suggest cutting-edge hybrid techniques to handle FS issues. The use of hybrid metaheuristic approaches for FS has thus been the subject of numerous research works. The purpose of this paper is to critically assess the existing hybrid FS approaches and to give a thorough literature review on the hybridization of different metaheuristic/evolutionary strategies that have been employed for supporting FS. This article reviews pertinent documents on hybrid frameworks that were published in the period from 2009 to 2022 and offers a thorough analysis of the used techniques, classifiers, datasets, applications, assessment metrics, and schemes of hybridization. Additionally, new open research issues and challenges are identified to pinpoint the areas that have to be further explored for additional study.

Keywords: metaheuristics; feature selection; hybridization; evolutionary methods; classification

1. Introduction

Feature selection (FS) is a method that aims to choose the minimum required features that can represent a dataset by selecting those features that add the most to the estimation variable that falls within the user's field of interest [1]. The volume of data available has risen significantly in recent years due to advancements in data gathering techniques in different fields, resulting in increased processing time and space complexity needed for the implementation of architectures in the realms of machine learning (ML). The collected data in many domains typically is of high dimensionality, making it impossible to select an optimum range of features and exclude unnecessary ones. The employed ML models are forced to learn insignificantly as a result of inappropriate features in the dataset, which leads to a poor recognition rate and a large drop in outcomes. By removing unnecessary and outdated features, FS reduces the dimensionality and improves the quality of the resulting attribute vector [2–4]. FS has been used for various purposes, including cancer classification (e.g., to improve the diagnosis of breast cancer and diabetes [5]), speech recognition [6], gene prediction [7], gait analysis [8], and text mining [9], etc.

FS has a pair of essential opposing goals, namely, reducing the number of needed characteristics and maximizing the performance of classification to overcome the curse of dimensionality. The three principal kinds of any FS strategy are filter, wrapper, and embedded methods, which integrate both filters and wrappers [10,11]. A filter technique is

independent of any ML algorithm. It is appropriate for datasets containing fewer features, and it often requires low-performance computing capabilities. In filtering approaches, the association among classifiers and attributes is not considered, and thus filters often fail to detect the samples correctly during the learning process.

Many studies have used wrappers to address these problems. A wrapper technique frequently alters the training process and uses classifiers as assessment mechanisms. Thus, wrapper techniques for FS often affect the training algorithm and produce more precise results than filters. Wrappers put effort into training the employed ML algorithm by using only a subset of the features that are also needed for determining the training model performance. Depending on the selection accuracy determined in each preceding phase, a wrapper algorithm considers either adding or removing a feature from the selected number of features. As a result, wrapper methods are often more computationally complex and more expensive than most filtering techniques.

Conventional wrapper approaches [12] take a set of attributes and require the user to include arguments as parameters, after which the most informative attributes are chosen from a set of features proportional to the arguments provided by the user. The limitations of such techniques are that the selected feature vector is recursively evaluated, in which case certain characteristics are not included at the first level for assessment. In addition, arguments are specified by the user, and thus certain feature mixtures cannot be taken into account even with more precision. These issues may cause searching overhead along with overfitting. Evolutionary wrapper approaches, which are more common when the search area is very broad, have been created to address the drawbacks of classic wrapper methods. These approaches have many benefits over conventional wrapper methods, including the fact that they need fewer domain details. Evolutionary optimization techniques are population-based metaheuristic strategies that can solve a problem with multiple candidate solutions described by a group of individuals. Each entity in the FS tasks represents a part of the feature vector. An objective (target) function is employed to evaluate and assess the consistency of every candidate solution. The chosen individuals are exposed to the intervention of genetic operators in order to produce new entities that comprise the next generation [13].

A plethora of variations of metaheuristic methods has already been developed to support the FS tasks. When defining a metaheuristic approach, exploration and exploitation are two opposing aspects to take into account. In order to increase the effectiveness of these algorithms, it is essential to establish a good balance between these two aspects. This is because the algorithms perform well in some situations but poorly in others. Every nature-inspired approach has advantages and disadvantages of its own; hence it is not always practical to predict which algorithm is best for a given situation [14].

Researchers [15] now face a hurdle in the implementation and high-precision suggestion of modern metaheuristics for real-world applications. As a result, several researchers are working to solve FS challenges by using hybrid metaheuristics. By merging and coordinating the exploration and exploitation processes, hybridization aims to identify compatible alternatives to ensure the best possible output of the applied optimization methods [16]. A typical strategy for addressing such issues is to combine the advantages of various independent architectures through the hybridization of metaheuristic methods [15,17].

This review paper extends our previous work presented in [18]. The reasons for broadening and extending the research on hybrid FS are highlighted as follows.

- The initial work [18] only focused and reviewed a limited number of papers (only 10 in number) published from 2020–2021. In order to provide a more comprehensive overview of the field, the additional relevant research on hybrid FS from 2009–2022 is extremely important to include in the review study.
- The current review paper deepens the scope of our research on multiple domains covering a wide range of metaheuristic approaches and wrapper classifiers.
- The literature review presented in the current paper aims to fulfill the highly evolving nature of research in the field of FS, and it is very important to stay up to date with the

latest developments in order to provide the most accurate and relevant information to the readers.

- Therefore, we believed it was important to design the current updated and extended review paper, which will be of interest to researchers in the FS domain.

We intend to address research issues and challenges that are open and interesting in terms of further research, and to provide a thorough overview of hybrid evolutionary techniques used to solve FS problems. This review draws the attention of scholars working with various metaheuristic frameworks, enabling them to further investigate enlightened approaches for tackling the complex FS problems often encountered in big data applications across many application domains.

The remaining parts of this review article include Section 2, which gives an outline of feature-collection processes and important contextual information. The details of the applied literature review on hybrid evolutionary methods for FS are presented in Section 3. Section 4 provides analysis and guidance for future research based on the literature studies. The last section contains the conclusions of this study. Table 1 summarizes the acronyms of all terms used in this paper (i.e., Table 1 presents the names of all presented FS selection methods, ML models, parameters, and corresponding evaluation metrics).

Table 1. Acronyms of the reviewed FS methods and respective evaluation metrics.

Searching Techniques	
ABC	Artificial Bee Colony Algorithm
ACO	Ant Colony Optimization
AFSA	Artificial Fish-Swarm Algorithm
AJA	Adaptive Jaya Algorithm
ALO	Ant Lion Optimization
Ant–Cuckoo	Ant Colony Optimization-Cuckoo Search
ASO	Atom Search Optimization
BALO	Binary Ant Lion Optimization
BBBC	Big Bang Big Crunch
BGWO	Binary Grey Wolf Optimization
BGWOPSO	Binary Grey Wolf Optimization-Particle Swarm Optimization
BHHO	Binary Harris Hawks Optimization
BPSO	Binary Particle Swarm Optimization
BSA	Backtracking Optimization Search Algorithm
BSO	Brain Storm Optimization
BTLBO	Binary Teaching Learning-Based Optimization
CFO	Central Force Optimization
ChOA	Chimp Optimization Algorithm
CRO	Chemical Reaction Optimization
CS	Cuckoo Search
CSA	Crow Search Algorithm
CSO	Curved Space Optimization
CSS	Charged System Search
DA	Dragon Algorithm

Table 1. *Cont.*

DE	Differential Evolution
DPO	Dolphin Partner Optimization
DSA	Differential Search Algorithm
FA	Firefly Algorithm
FLA	Frog Leaping Algorithm
FSJaya	FS-Based on Jaya optimization
GA	Genetic Algorithm
GOA	Grasshopper Optimization Algorithm
GSA	Gravitational Search Algorithm
GSO	Group Search Optimizer
GWO	Grey Wolf Optimization
HBBEPSO	Hybrid Binary Bat Enhanced Particle Swarm Optimization Algorithm
IBHHO	Improved Binary Harris Hawks Optimization
HBO	Heap-Based Optimizer
HBPSOSCA	Hybrid Binary Particle Swarm Optimization and Sine Cosine Algorithm
HHO	Harris Hawk Optimization
HS	Harmony Search
ISA	Interior Search Algorithm
JA	Jaya Algorithm
KHA	Krill Herd Algorithm
LCA	League Championship Algorithm
MA	Monkey Algorithm
MAKHA	Monkey–Krill Herd Algorithm
MBA	Mine Blast Algorithm
MFO	Moth–Flame Optimization
MOChOA	Multiobjective Chimp Optimization
MPA	Marine Predators Algorithm
MVA	Multiverse Algorithm
PSO	Particle Swarm Optimization
QE	Queen Bee Evolution
RTEA	Ring Theory-Based Evolutionary Algorithm
RTHS	Ring Theory-Based Harmony Search
SA	Simulated Annealing
SBS	Sequential Backward Selection
SCA	Sine Cosine Algorithm
SDO	Supply–Demand-Based Optimization
SFLA	Shuffled Frog Leaping Algorithm
SFS	Sequential Forward Selection
SHO	Spotted Hyena Optimization
SHO-SA	Spotted Hyena Optimization-Simulated Annealing
SOA	Seagull Optimization Algorithm

Table 1. *Cont.*

SPO	Stochastic Paint Optimizer
SSA	Salp Swarm Algorithm
TEO	Thermal Exchange Optimizer
TLBO	Teaching Learning-Based Optimization
TS	Tabu Search
VSA	Vortex Search Algorithm
WOA	Whale Optimization Algorithm
Machine Learning Algorithms	
ANN	Artificial Neural Network
CART	Classification And Regression Tree
DT	Decision Tree
KNN	k-Nearest Neighbor
LDA	Linear Discriminant Analysis
LR	Logistic Regression
NB	Naive Bayes
RF	Random Forest
SVM	Support Vector Machine
Performance Metrics	
ACC	Accuracy
AUROC	Area Under the Receiver Operating Characteristic
BF	Best Fitness
CA	Classification Accuracy
CE	Classification Error
FDR	False Discovery Rate
FNR	False Negative Rate
FPR	False Positive Rate
FSc	F-Score
IGD	Inverted Generational Distance
MCC	Matthews Correlation Coefficient
MF	Mean Fitness
MSE	Mean Square Error
NFE	Number of Function Evaluations
NPV	Negative Predictive Value
NSF	Number of Selected Features
PA	Predictive Accuracy
PPV	Positive Predictive Value
PR	Precision
RE	Recall
RT	Running Time
SN	Sensitivity

Table 1. *Cont.*

Parameters	
DBI	Davies–Bouldin Index
DI	Dunn Index
SI	Silhouette Index
Others	
EC	Evolutionary Computation
FS	Feature Selection
ML	Machine Learning
MO	Multiobjective
MRMR	Minimum Redundancy Maximum Relevance
OBL	Opposition-Based Learning
SO	Single Objective
WRS	Wilcoxon's Rank Sum

2. Related Work

It is unusual for all properties in a considered dataset to be useful when designing an ML platform in real life. The inclusion of unwanted and redundant attributes lessens the model's classification capability and accuracy. As more factors are added to an ML framework, its complexity increases [19,20]. By finding and assembling the ideal set of features, FS in ML aims to produce useful models of a problem under study and consideration [21]. Some important advantages of FS are [10,12]:

- reducing overfitting and eliminating redundant data,
- improving accuracy and reducing misleading results, and
- reducing the ML algorithm training time, dropping the algorithm complexity, and speeding up the training process.

The prime components of an FS process are presented in Figure 1 and they are [18] as follows.

1. Searching Techniques: To obtain the best features with the highest accuracy, searching approaches are required to be applied in an FS process. Exhaustive search, heuristic search, and evolutionary computation are a few popular searching methods. An exhaustive search is explored in a few works [19,20]. Numerous heuristic strategies and greedy techniques, such as sequential forward selection (SFS) [22], and sequential backward selection (SBS), have therefore been used for FS [23]. However, in later parts of the FS process, it could be impossible to select or delete eliminated or selected features because both SFS and SBS suffer from the "nesting effect" problem. After being selected, features in the SFS method cannot be discarded later, while the features discarded in the SBS cannot be selected again. These two approaches can be compromised by using SFS *l* times and then applying SBS *r* times [24]. The nesting effect can be reduced by such a method, but the correct values of *l* and *r* must be determined carefully. Sequential backward and forward floating methods were presented to avoid this problem [22]. A two-layer cutting plane approach was recently suggested in [23] to evaluate the best subsets of characteristics. In [24], an exhaustive FS search with backtracking and a heuristic search was proposed.
Various EC approaches have been proposed in recent years to tackle the challenges of the FS problems successfully. Some of them are differential evolution (DE) [25], genetic algorithms (GAs) [26], grey wolf optimization (GWO) [27,28], ant colony optimization (ACO) [29–31], binary Harris hawks optimization (BHHO) [32,33] and improved BHHO (IBHHO) [34], binary ant lion optimization (BALO) [35,36], salp swarm algo-

rithm (SSA) [37], dragon algorithm (DA) [38], multiverse algorithm (MVA) [39], Jaya optimization algorithms such as the FS based on the Jaya optimization algorithm (FSJaya) [40] and the FS based on the adaptive Jaya algorithm (AJA) [41], grasshopper swarm intelligence optimization algorithm (GOA) and its binary versions [42], binary teaching learning-based optimization (BTLBO) [43], harmony search (HS) [44], and the vortex search algorithm (VSA) [45], etc. All these techniques have been applied for performing FS on various types of datasets, and they have been demonstrated to achieve high optimization rates and to increase the CA. EC techniques require no domain knowledge and do not presume whether the training dataset is linearly separable or not. Another valuable aspect of EC methods is that their population-based process can deliver several solutions in one cycle. However, EC approaches often entail considerable computational costs because they typically include a wide range of assessments. The stability of an EC approach is also a critical concern, as the respective algorithms often pick different features from various rounds. Further research study is required as the growing number of characteristics in large-scale datasets also raises computational costs and decreases the consistency of EC algorithm application [13] in certain real-world FS activities. A high-level description of the most used EC algorithms is given below.

- Genetic Algorithm (GA): A GA [46] is a metaheuristic influenced by natural selection that belongs to the larger class of evolutionary algorithms in computer science and operations research. GA relies on biologically inspired operators, such as mutation, crossover, and selection to develop high-quality solutions to optimization and search challenges. The GA is a mechanism that governs biological evolution and for tackling both constrained and unconstrained optimization issues. The GA adjusts a population of candidate solutions on a regular basis.

- Particle Swarm Optimization (PSO): PSO is a bioinspired algorithm that is straightforward to use while looking for the best alternative in the solution space. It differs from other optimization techniques in that it simply requires the objective function and is unaffected by the gradient or any differential form of the objective. It also has a small number of hyperparameters. Kennedy and Eberhart proposed PSO in 1995 [47]. Sociobiologists think that a school of fish or a flock of birds moving in a group "may profit from the experience of all other members", as stated in the original publication. In other words, while a bird is flying around looking for food at random, all of the birds in the flock can share what they find and assist the entire flock to get the best hunt possible. While we may imitate the movement of a flock of birds, we can also assume that each bird is assisting us in locating the best solution in a high-dimensional solution space, with the flock's best solution being the best solution in the space. This is a heuristic approach because we can never be certain that the true global optimal solution exists, and it rarely does. However, we frequently discover that the PSO solution is very close to the global optimum.

- Grey Wolf Optimizer (GWO): Mirjalili et al. [48] presented GWO as a new metaheuristic in 2014. The grey wolf's social order and hunting mechanisms inspired the algorithm. First, there are four wolves, or degrees of the social hierarchy, to consider when creating GWO.

  – The α wolf: the solution having best fitness value;
  – the β wolf: the solution having second-best fitness value;
  – the δ wolf: the solution having third-best fitness value; and
  – the ω wolf: all other solutions.

  As a result, the algorithm's hunting mechanism is guided by the first three appropriate wolves, α, β, and δ. The remaining wolves are regarded as ω and follow them. Grey wolves follow a set of well-defined steps during hunting: encircling, hunting, and attacking.

- Harris Hawk Optimization (HHO): Heidari and his team introduced HHO as a new metaheuristic algorithm in 2019 [49]. HHO uses Harris hawk principles to investigate the prey, surprise pounce, and diverse assault techniques used by Harris hawks in the environment. Hawks reflect alternatives in HHO, whereas prey represents the best solution. The Harris hawks use their keen vision to follow the target and then conduct a surprise pounce to seize the prey they have spotted. In general, HHO is divided into two phases: exploitation and exploration. The HHO algorithm can be switched from exploration to exploitation, and the exploration behaviour can then be adjusted depending on the fleeing prey's energy.

2. Criteria for Evaluation: The common evaluation criteria for wrapper FS techniques are the classification efficiency and effectiveness by using the selected attributes. Decision trees (DTs), support vector machines (SVMs), naive Bayes (NB), k-nearest neighbor (KNN), artificial neural networks (ANNs), and linear discriminant analysis (LDA) are just a few examples of common classifiers that have been used as wrappers in FS applications [50–52]. In the domain of filter approaches, measurements from a variety of disciplines have been incorporated, particularly information theory, correlation estimates, distance metrics, and consistency criteria [53]. Individual feature evaluation, relying on a particular aspect, is a basic filter approach in which only the best tier features are selected [50]. Relief [54] is a distinctive case in which a distance metric is applied to assess the significance of features. Filter methods are often computationally inexpensive, but they do not consider attribute relationships, which often leads to complicated problems in case of repetitive feature sets, such as in the case of microarray gene data, where the genes are intrinsically correlated [21,53]. To overcome these issues, it is necessary to perform proper filter measurements to choose a suitable subset of relevant features in order to evaluate the whole feature set. Wang et al. [55] recently published a distance measure to assess the difference between the chosen feature space and the space spanned by all features in order to locate a subset of features that approximates all features. Peng et al. [56] introduced the minimum redundancy maximum relevance (MRMR) approach based on shared information, and recommended measures were added to the EC because of their powerful exploration capability [57,58]. A unified selection approach was proposed by Mao and Tsang [23], which optimizes multivariate performance measures but also results in an enormous search area for high-dimensional data, a problem that requires strong heuristic search methods for finding the best output. There are several relatively straightforward statistical methods, such as *t*-testing, logistic regression (LR), hierarchical clustering, and classification and regression tree (CART), which can be applied jointly to produce better classification results [59]. Recently, authors of [60] have applied sparse LR for FS problems including millions of features. Min et al. [24] developed a rough principle procedure to solve FS tasks under budgetary and schedule constraints. Many experiments show that most filter mechanisms are inefficient for cases with vast numbers of features [61].

3. Number of Objectives: Single-objective (SO) optimization frameworks are techniques which combine the classifier's accuracy and the features quantity into a single optimization function. On the contrary, multiobjective (MO) optimization approaches entail techniques designed to find and balance the tradeoffs among alternatives. In an SO situation, a solution's superiority over other solutions is determined by comparing the resulting fitness values, while in an MO optimization, the dominance notion is employed to get the best results [62]. In particular, to determine the significance of the derived feature sets, in an MO situation, multiple criteria need to be optimized by considering different parameters. MO strategies thus may be used to solve some challenging problems involving multiple conflicting goals [63], and MO optimization comprises fitness functions that minimize or maximize multiple conflicting goals.

For example, a typical MO problem with minimization functions can be expressed mathematically as follows,

$$
\begin{aligned}
min\ O(x) &= [o1(x), o2(x), \ldots, o_n(x)] \\
Sub\ to:\ g_i(x) &\leq 0, \quad i = 1, 2, \ldots\ldots, m\ , \\
h_i(x) &= 0, \quad i = 1, 2, \ldots, l
\end{aligned} \tag{1}
$$

where x is the decision variables vector, n is the number of objectives, $o_n(x)$ is the nth objective function and $g_n(x)$ and $h_n(y)$ are the problem constraints.

Finding the balance among the competing objectives is the process that identifies the dominance of an MO optimization approach. For example, a solution s_1 dominates another solution s_2 in a minimization problem if and only if

$$
\forall : O_k(s_1) \leq O_k(s_2)\ and\ \exists_l : O_l(s_1) < O_l(s_2)\ , \tag{2}
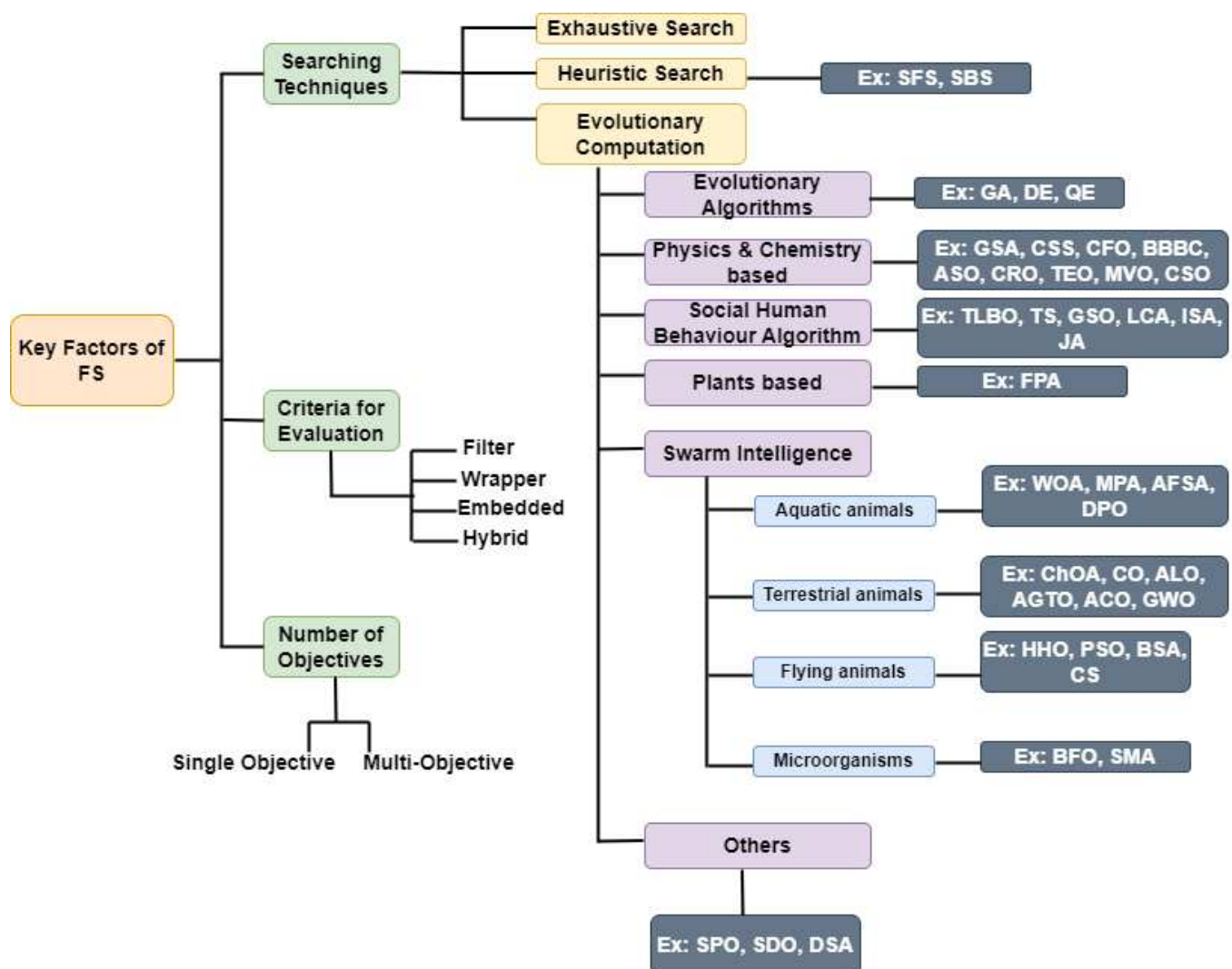$$

where $k, l \in \{1, 2, \ldots, n\}$.

Figure 1. Key factors of feature selection.

3. A Brief Survey

Search Procedure

We adhere to the PRISMA principles for systematic reviews in our work (www.prisma-statement.org (accessed on 19 March 2023)). The relevant research questions are developed in accordance with these standards:

1. What are the search approaches that were utilised to find the best features?
2. What are the search algorithms utilised to choose the best features for classification?
3. What hybrid search approaches have been utilised to choose the best characteristics for classification?

The review began by searching for relevant research on Internet sites and in the University Teknologi PETRONAS online library. The Internet search was guided by the use of search engines to explore the electronic libraries and databases depicted in Figure 2. The terms "hybrid + feature selection", "hybrid + search technique + feature selection", and "hybrid + search technique + feature selection + classification" were the search parameters employed. There have been several studies on hybrid evolutionary FS. To ensure that the search was concentrated and controllable, the following inclusion and exclusion criteria were defined to select the publications for further study:

- Inclusion Criteria:
    - Research articles on hybrid evolutionary FS must have been published between 2009 and 2022.
    - Only research that has been published in peer-reviewed publications is included.
    - If the study had been published in more than one journal, we select the most complete version for inclusion.
    - Only related works utilised for classification are included.

- Exclusion Criteria:
    - Research articles prior to 2009 are not included.
    - Papers that are unrelated to the search topic are rejected.
    - Only items written in English are considered. Other languages are removed.

The papers chosen by the abovementioned search procedure were reviewed by title and abstract in accordance with the inclusion and exclusion criteria. Then, all of the studies identified as relevant to our topic were downloaded for information extraction and additional investigation. Figure 2 provides information on the number of research studies discovered during the search of the most popular computerised libraries and databases.
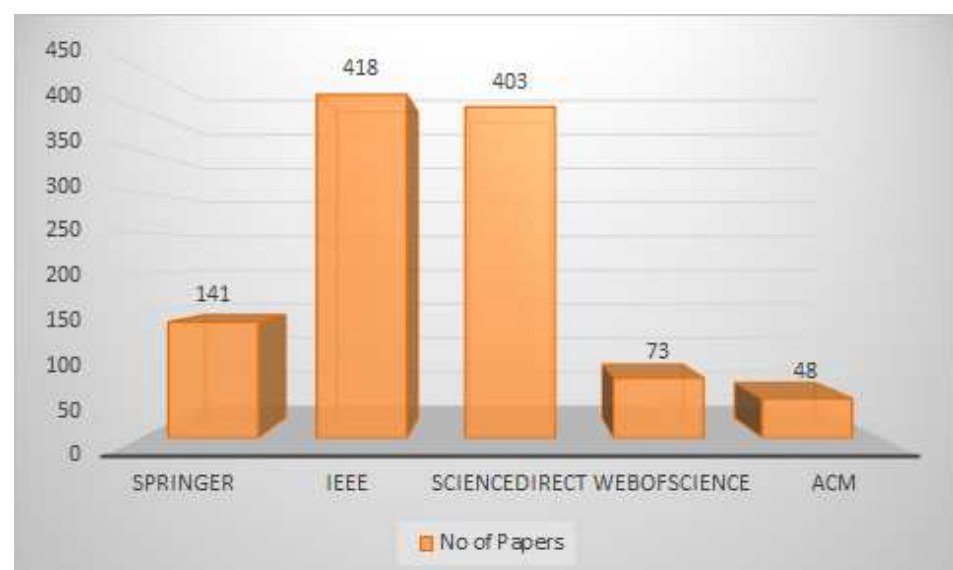

Figure 2. Number of papers identified.

The next step was to prescreen the abstracts of the returning results. The primary goal of prescreening was to eliminate redundant data (some papers were returned in multiple databases) as well as incorrect findings. Improper findings were found in some studies where authors claimed to have employed the hybrid idea, but our research demonstrated that they hybridize filter and wrapper criteria rather than multiple search techniques.

Finally, studies of 35 publications on hybrid metaheuristic approaches that were presented between 2009 and 2022 are covered in this review report. Figure 3 presents the number of papers collected for each year.
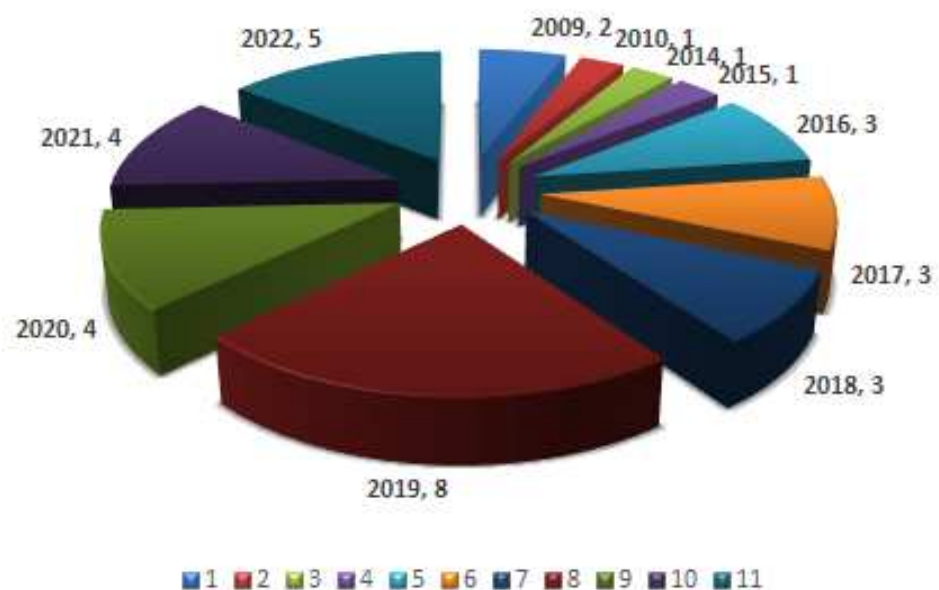

Figure 3. Number of papers collected per year.

All identified articles were scrutinized by their title and abstract. The current review paper provides a thorough picture on the metaheuristics used for hybridization and also presents a range of various classifiers and datasets, the application fields of the corresponding techniques, their objective/fitness functions and assessment metrics, and the application fields of various hybridised approaches, in contrast to individual methods.

In Table 2, a brief introduction is given about each one of the collected papers in the relevant literature.

Table 2. Introduction to the collected papers.

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P1 [64]	2022	The intention of this article is to design a simplified and functional hybrid algorithm for FS by considering the simplicity of the WOA and the stochastic nature of HHO.	Experimental findings on 18 benchmark datasets reveal that the proposed hybrid method is capable of enhancing the achievement of the conventional WOA concerning ACC, selected feature count, and execution time.	CA, MF, BF, WF, NSF, and RT.

Table 2. *Cont.*

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P2 [65]	2020	This study attempts to combine DE and Chaotic Dynamic Weight Particle Swarm Optimization (CHPSO) in an effort to enhance CHPSO.	According to the simulation outputs, CHPSO-DE performs better than other solutions at solving practically the FS challenge.	Average NFE.
P3 [66]	2020	This study presented a novel hybrid FS model called RTHS, which is based on the HS metaheuristic and RTEA.	The RTHS approach was applied on 18 standard datasets from UCI Machine Learning Repository and it was contrasted to 10 popular evolutionary FS methods. The findings indicated that the RTHS approach appears to be more effective than the considered approaches.	CA, PR, RE, FSc, and AUROC.
P4 [67]	2022	The aim of this research is to use a unique hybridised wrapper-based brain storm optimization-firefly algorithm (BSO-FA) strategy in order to enhance the FS technique and produce improved classification outcomes on standard UCI datasets, including a publicly available dataset with COVID-19 patient health data.	On 21 UCI datasets, the suggested approach is assessed and contrasted with 11 optimization techniques. The proposed approach is also used for a dataset related to coronavirus disease. The observed experimental findings support the robustness of the suggested hybrid model. In comparison to other methods in the literature, it effectively decreases and chooses the features and also produces better CA.	CA and MF.
P5 [68]	2022	The objective of this study is to apply a hybridised grey wolf optimization-heap-based optimizer (GWO-HBO) methodology as a wrapper for the FS phase of a fault-diagnosis system.	The suggested approach is validated on four separate datasets to ensure its efficiency. The proposed approach is compared to three methods, namely BGWO, BPSO, and GA, and the test results may attest to its predictability.	CA.
P6 [69]	2022	This work aims at designing a hybrid optimization technique based on the CSA and the HHO for selecting features and mass categorization in digital mammograms.	This strategy was tested by using 651 mammograms. When compared with respect to conventional CSA and HHO methods employing experimental data, the new CSAHHO method was found to perform better.	CA, SN, SP, FPR, FNR, FSc, and Kappa coefficient.

Table 2. *Cont.*

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P7 [70]	2020	The purpose of this article is to acquire the balance between exploitation and exploration by mixing the advantages of GWO and PSO.	Seventeen UCI datasets are used to measure the suggested optimizer's consistency, dependability, and stability.	Average error, average NSF, MF, BF, WF, STD of fitness, and RT.
P8 [71]	2019	In order to determine the appropriate trait subgroup and resolve FS issues, this article suggests a hybrid PSO and GWO.	The study's results highlight that the BGWOPSO framework is superior in computation time, PR, and FS. The findings of the BGWOPSO procedure have shown that it is easier than other approaches to monitor the compromise between exploratory and exploitative behaviours.	Average CA, NSF, MF, BF, WF, and RT.
P9 [72]	2019	This study provides three hybrid structures for the FS task based on TEO and SOA.	The simulation outcomes have demonstrated that the suggested hybrid model enhances classification efficiency, guarantees the choice of hybrid SOA-algorithms, decreases CPU time, and selects the salient factor.	RT, average NSF, and CA.
P10 [73]	2019	This study provides two separate hybrid versions of the spotted hyena optimization (SHO) for FS problems. In the first version (SHOSA-1), the SA is embedded in the SHO algorithm. In the second version (SHOSA-2), the SA is used to enhance the ultimate solution obtained by the SHO algorithm.	The findings of the tests revealed that the SHOSA-1 strategy improves recognition rate and reduces the number of chosen features in relation to other wrapper methods. Experiments also demonstrated that SHOSA-1 has excellent success (compared to SHOSA-2) in the spatial search and choice of feature characteristics.	Average CA, NSF, MF, STD, RT, SN, and SP.
P11 [74]	2019	In this paper, the OBL concept is integrated with a DE technique and an MFO approach in order to boost the capacity of the MFO algorithm for generating an optimal attribute array.	The findings clearly reveal that the presented algorithm is better in terms of efficiency and the methodology suggested with a limited range of selected features and a minimal CPU time in comparison to other benchmark evolutionary approaches.	MF, STD, average RT, selection ratio (SR), CA, FSc, PR, RE, and WRS.

Table 2. *Cont.*

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P12 [75]	2019	This article presents a hybrid GWOCSA that efficiently blends the strengths of both GWO and the crow search optimizer (CSO) to provide optimal solutions for the most efficient global operation.	The experimental findings suggest that the GWOCSA has improved fitness optimization and performed at a higher convergence speed compared to the other FS methodologies to solve the FS problem and achieved more satisfactory optimization results in fewer iterations. This demonstrates the potential of the model to solve difficult issues in real-world large datasets.	CA, average NSF, MF, STD, and WRS.
P13 [76]	2018	This paper proposes a unique hybrid method for the FS problems known as the HBBEPSO.	The outcomes from testing the HBBEPSO demonstrate the possibility of using the recommended hybrid strategy to determine the ideal variable combination.	MF, BF, WF, STD, average SR, and average FSc.
P14 [77]	2018	This article utilizes a hybrid method of discrete PSO and the SFLA to reduce the feature dimension and choose optimal parameter subsets.	The simulation outputs indicate that the suggested hybrid approach is good enough to provide an optimized attribute subset and obtain a high CA values.	CA, PR, and RE.
P15 [78]	2017	The DE operators are used as local search techniques in this work to address the difficulties in SCA.	The outcomes of the execution conclude that the new technique will function better than the alternatives on the basis of success metrics and predictive analysis.	CE, FSc, MF, BF, WF, STD, SR, and RT.
P16 [79]	2017	In this paper, hybridized frameworks are introduced to construct FS models based on the WOA.	The proposed hybrid models combine SA with WOA. The derived experimentation results have shown that the performance and the capacity of the hybrid WOA approach for choosing the most informational features and for searching the feature space are improved compared with individual wrapper approaches.	CA, NSF, MF, BF, and WF.
P17 [6]	2016	This article presents a hybrid evolutionary optimization technique called the artificial bee colony-particle swarm optimization (ABC-PSO) for optimum selection and retrieval of features.	The findings show that the overall efficiency of the method is very good and the suggested hybrid method is better suited for voice recognition upon implementing it in the MATLAB working platform.	CA, SN, SP, PPV, NPV, FPR, FDR, and MCC.

Table 2. *Cont.*

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P18 [80]	2016	This article suggests nature-based hybrid techniques for FS. The techniques are based on two strategies for swarm intelligence: ACO and PSO.	The experimental findings conclude that the proposed approaches have better efficiency for reducing the NSF and also in terms of CA.	CA and NSF.
P19 [81]	2015	This paper proposes the use of a hybrid MA for FS combined with the KHA.	The test results reveal that the proposed MAKHA technique can easily find an optimal or almost an optimal set of combination of attributes by minimizing the objective function and achieving sufficient efficiency to increase the accuracy of feature classification.	BF, MF, WF, and CE.
P20 [9]	2010	This research suggests a hybrid approach for FS based on GA and SA.	The FS results were improved by correcting the SA in the creation of the next generation (by considering two maximum and minimum thresholds).	CA and NSF.
P21 [82]	2009	This article suggests a new FS framework which combines GA with ACO to increase and improve search capabilities in protein structure forecasting.	The testing results show the superiority of the proposed hybrid method (compared to ACO ans GA) and also present the low computational complexity of the suggested hybrid approach.	Predictive accuracy (PA).
P22 [83]	2009	In this paper, TS is combined with binary PSO to select an optimal feature vector in FS.	Testing results from applying the method on 11 classification problems taken from the literature show that this approach simplifies features effectively. This method has the ability to obtain higher CA and to use fewer features compared to other FS methods.	CA.
P23 [84]	2019	This study suggests the combination of BPSO and SCA. The aim of the approach is to perform FS and cluster analysis by employing a cross-breed approach of SCA to BPSO.	The experimental findings (on 10 benchmark test functions and seven datasets taken from the UCI repository) show that the suggested HBPSOSCA approach generally performs better than other FS approaches.	MF, BF, WF, Average NSF, SI, DI, and DBI.

Table 2. *Cont.*

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P24 [85]	2019	This research presents a hybrid filter-wrapper approach for the collection of attribute subsets, focused on the hybridization of GA and PSO. The method utilizes an ANN in the fitness/objective function.	The experimental findings on five datasets showed that the suggested hybrid approach achieves a higher PR of classification in comparison to other competitor techniques.	Average NSF, average CA, best ACC, and average RT.
P25 [86]	2018	This paper presents a hybrid of two methods, ALO and GWO, that provides the strength of having a good understanding from fewer instances and the decent collection of characteristics from a very wide range, thus maintaining a high PR in the classification results.	Datasets with around 50,000 characteristics and fewer than 200 examples were utilized to measure the accuracy of the system. The test findings are positive with respect to GA and PSO.	MF, BF, WF, STD, CMSE, average NSF, average Fisher score, and WRS.
P26 [87]	2019	This paper proposes a new hybrid ALO with elitism-based DE to tackle task-scheduling problems in cloud environments.	The experimental results showed that for larger search spaces, the modified-ALO (MALO) approach converged faster, proving it ideal for massive task scheduling jobs. The statistical t-tests were used to analyse the data, indicating that MALO significantly improved the results.	Degree of imbalance, size of tasks, makespan, and RT.
P27 [88]	2021	The purpose of this research is to perform FS by fusing an improved CSA method with PSO.	With the use of 15 datasets from the UCI, the presented technique is compared to four well-known optimization techniques, namely PSO, binary PSO, CSA, and chaotic CSA. Distinct performance indicators were applied in the tests by using KNN as classifier. This hybrid approach was found to perform better than cutting-edge techniques.	MF, BF, WF, and STD of fitness.

Table 2. *Cont.*

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P28 [89]	2021	The main goal of this approach is to shorten the size of the selected feature vector by combining TLBO and SSA techniques, which can also increase the classifier's predictability.	A total of 651 breast cancer screenings were produced by the hybrid approach, and the outputs demonstrate that TLBO-SSA performs better than the TLBO. Once more, the strength of this metaheuristic approach was evaluated by taking a UCI dataset. The TLBO-SSA result demonstrated its superiority when compared to GA.	SE, SP, CA, FSc, Kappa coeff, FPR, and FNR.
P29 [90]	2020	In order to increase the initial HHO's effectiveness for collecting chemical descriptors and chemical composites, this work combined HHO, CS, and chaotic maps.	Some UCI datasets and two chemical datasets are considered to validate the presented solution. Comprehensive experimental and computational analysis showed that the proposed approach has achieved many desired solutions over other competing solutions.	CA, SE, SP, RE, PR, and FSc.
P30 [91]	2021	This article proposes a hybrid optimal strategy that includes SCA in HHO. By adjusting the candidate solutions in a complex manner, SCA attempts to tackle ineffective HHO identification and to prevent stagnation situations in HHO.	With 16 datasets including more than 15,000 in attributes and the CEC'17 computational optimising trials, this recommended approach was evaluated and contrasted with SCA, HHO, and other existing methods. The detailed evaluations of experiments and statistics showed that the suggested HHO hybrid variant produced effective results without extra computational cost.	Average CA, MF, average NSF, SR, average RT, and STD.
P31 [92]	2021	This article suggests a hybrid GWO-HHO-based FS technique.	In comparison to GWO, PSO, HHO, and GA, the accuracy of the suggested hybrid approach was tested and evaluated on 18 UCI datasets. The approach performed better than the GWO.	Average CA, MF, BF, WF, average NSF, and average RT.
P32 [93]	2017	This research suggests a new AC-ABC hybrid technique that incorporates the FS characteristics of ACO and ABC. By employing hybridization, the stagnation behavior of the ants is removed, and lengthy global searches for the original solutions by the employed bees.	The suggested method was evaluated by 13 UCI datasets. Experimental findings revealed the positive characteristics of the proposed technique which was found to achieve a high accuracy rate and optimum selection of features.	NSF, CA, and RT.

Table 2. *Cont.*

Paper	Year	Aim	Experimental Evaluation	Assessment Metrics
P33 [94]	2016	In this paper, a hybrid approach that merges the ABC optimizer with DE is recommended for FS in classification problems.	The approach was tested by using 15 UCI datasets and was compared with ABC and DE based FS, and also with gain, chi-square and correlation based FS. The empirical outputs of this study indicate that the new technique selects informative features for classification that increase the classifier's efficiency and accuracy.	FSc, NSF, and RT.
P34 [95]	2014	In this paper, a novel hybrid evolutionary technique called ant–cuckoo-produced by the fusion of ACO and CS methods-is introduced for performing FS in digital mammogram.	The tests are carried out on the miniMIAS database of mammograms. Compared with ACO and PSO algorithms, the efficiency of the ant–cuckoo method was analyzed. The findings indicated that FS optimization for the hybrid ant–cuckoo method was more accurate than the one achieved by the individual FS approaches.	SN, SP, CA, and AUROC.
P35 [17]	2022	By using a MOChOA-based FS technique, this method seeks to identify pertinent parameters for forecasting the health status of COVID-19 patients.	By contrasting this strategy with five other existing FS procedures on nine distinct datasets, its efficacy is demonstrated.	Average NSF, average CA, average RT, and IGD.

The search methods that have been fused together in each metaheuristic approach, the details of the corresponding fitness/objective function along with the respective means of hybridization which have been used in each approach are given in Table 3.

Table 3. Search methods, their fitness function details, and means of hybridization.

Paper	Search Method	Fitness Function	Means of Hybridization
P1 [64]	WOA, HHO	$Fitness = \alpha(ER) + (1 - \alpha)\frac{\lvert S_f \rvert}{\lvert T_f \rvert}$, where $\alpha \in [0,1]$, ER: error, S_f: #picked factors and T_f: #actual factors.	The exploration technique of HHO is immersed in the WOA to rise the randomness of the optimum solution search, based on the humpback whale's exploitative manner.
P2 [65]	Chaotic PSO (CPSO), DE.	NFE.	In order to prevent the decay that is normally discovered by the CPSO, the DE approach is combined with CPSO. As the swarm begins to deteriorate, the DE is used to provide the required momentum for the particles to travel through the search area and thereby flee from the local optima.

Table 3. *Cont.*

Paper	Search Method	Fitness Function	Means of Hybridization
P3 [66]	HS, RTEA	$Fitness = \omega\zeta(F') + (1 - \omega)^{\times}\frac{|F'|}{|F|}$, where F': array of chosen attributes, $\zeta(F')$: error rate with reduced feature string, F: original feature array and $\omega \in [0,1]$	HS and RTEA have been hybridised by following the pipeline model.
P4 [67]	BSO and FA	$Fitness = \alpha E_R(D) + \beta\frac{|R|}{|C|}$, where $E_R(D)$: classifier's error rate, $|R|$: size of the chosen attribute string, $|C|$: count of total features and α lies in between 0 and 1 and $\beta = (1 - \alpha)$.	To reduce the drawbacks of the conventional BSO, this new architecture combines the best elements of BSO's great exploration and FA's exceptional exploitation where if the cycle counter is odd, the FA search mechanism is used for a location change; otherwise, the original BSO is used for solution improvement.
P5 [68]	GWO and HBO	$Fitness = \frac{|N_T|}{|N_T+N_f|} \times 100\%$, where NT: #truly predicted instances, NF: #instances that are falsely predicted.	The best solution obtained from GWO is stored as a record. If the new solution generated by the HBO is more than 90% similar to the above record, then crossover is used. After crossover, if the new solution is the same as the record then mutation is performed.
P6 [69]	CSA and HHO	$cost(x_i(t)) = (E_{x_i(t)} \times (1 + 0.5 \times \frac{FS}{N}))^2$, where $cost(x_i(t))$: fitness value of $x_i(t)$, $E_{x_i(t)}$: performance of classifier. FS/N: #features selected/#total features.	The probability $P_i = \frac{cost_i}{\sum_{j=1}^{pop\_size} cost_i}$ opts for either solutions to be updated by CSA or HHO.
P7 [70]	GWO and PSO	$Fitness = h_1 E(D) + h_2\frac{|s|}{|f|}$, where E(D): CE, s: #selected features, f: #features and $h \in [0,1]$, $h_2 = 1 - h_1$ are constants.	Starting with an arbitrary selection of solutions, the optimization process begins. After determining the fitness function for each individual for each iteration, the first three leaders are given the names alpha, beta, and delta. After that, the population is equally split into two classes, with the first class following GWO operations and the second class following PSO processes. In this manner, the search space is thoroughly examined for potential points, and these points are then utilised by the potent PSO and GWO.
P8 [71]	GWO-PSO	$Fitness = \alpha\rho_R(D) + \beta\frac{|S|}{|T|}$, where $\alpha = [0,1]$ and $\beta = (1 - \alpha)$, $\rho_R(D)$: KNN's error rate, $|S|$: length of chosen feature vector and $|T|$: length of actual feature vector.	The basic principle of PSOGWO is to enhance the potential of the system to exploit PSO to explore GWO in order to accomplish both optimizer powers, where the location of the first three agents is modified, rather than with the normal calculation, exploitation, and discovery of the grey wolf.

Table 3. *Cont.*

Paper	Search Method	Fitness Function	Means of Hybridization
P9 [72]	SOA, TEO	$Fitness = \alpha\gamma_R(D) + \beta \times \left(\frac{|R|}{|N|}\right)$, where $\gamma_R(D)$: CE, $|R|$: length of the chosen substring, $|N|$: length of whole feature set, $\alpha \in [0,1]$ and $\beta = (1-\alpha)$.	Three hybrid ways to manage FS tasks based on SOA and TEO are proposed in this paper. Either of the two algorithms is selected for updating the position in the first method based on the roulette wheel. The second approach is followed by SOA optimization by TEO. The final technique uses the TEO formulation for heat exchange to boost the SOA style of attack.
P10 [73]	SA, SHO algorithm	$Fitness = \alpha\gamma_R(D) + \beta \times \left(\frac{|R|}{|N|}\right)$, where $\gamma_R(D)$: CE, $|R|$: length of the chosen substring, $|N|$: length of whole feature set, $\alpha \in [0,1]$ and $\beta = (1-\alpha)$.	Two hybrid systems to enhance the use of the SHO model are presented in this article. SA is used as part of SHO in the first hybrid model and SHO and SA techniques are executed once for every iteration. In the case of second architecture, the first SHO model is applied to seek out the optimum solution, followed by the SA to find the new best solution.
P11 [74]	MFO, DE	$f(x_i) = \xi \times \text{Err}_{x_i} + (1 - \xi) \times \left(\frac{|x_i|}{Dim}\right)$, where Err_{x_i}: error of the classifier, $|x_i|$: count of chosen attributes, Dim: whole features count and ξ: any random between 0 and 1.	The suggested model utilizes the OBL principle to generate initial solutions, and the DE operators to boost the operational capabilities of MFO.
P12 [75]	CSA, GWO	$Fitness = \alpha\gamma_R(D) + \beta \times \left(\frac{|R|}{|N|}\right)$, where $\gamma_R(D)$: CE, $|R|$: length of the chosen substring, $|N|$: length of whole feature set, $\alpha \in [0,1]$ and $\beta = (1-\alpha)$.	In particular, in its location change equation, the CSA integrates a control parameter. This parameter plays a key part in achieving the global optimum as the big value of this factor results in global discovery and a small figure results in a local search. In the suggested GWOCSA, a greater value of the parameter is used to make use of the CSA's outstanding discovery quality.
P13 [76]	PSO, bat algorithm	$Fitness = \alpha E_R(D) + \beta\frac{|R|}{|C|}$, where $E_R(D)$: classifier's error rate, $|R|$: size of the chosen attribute string, $|C|$: count of total features and α lies in between 0 and 1 and $\beta = (1-\alpha)$.	The separation of the speed vectors of the bats and the particles calls for a new design of the suggested method. This is because the personal and global solutions are not modified after the BBA, but only after the full round of the PSO.
P14 [77]	Binary PSO (BPSO), frog leaping algorithm (FLA)	$Fitness(x) = accuracy(x)$, where $Accuracy(x)$: NB's accuracy and x: feature subset.	The population, which includes an optimized BPSO extracted feature, is given as an input for the FLA under the presented hybrid system.

Table 3. *Cont.*

Paper	Search Method	Fitness Function	Means of Hybridization
P15 [78]	SCA-DE	$f(x_i) = \xi \times Err_{x_i} + (1 - \xi) \times \left(1 - \frac{|S|}{D}\right)$, where Err_{x_i}: LR's error rate, $|S|$: count of picked features, D: total feature count and $\xi \in [0, 1]$.	The DE operators are applied as a local search strategy to help the SCA to avoid the local spot.
P16 [79]	SA, WOA	$Fitness = \alpha \gamma_R(D) + \beta \times \left(\frac{|R|}{|N|}\right)$, where $\gamma_R(D)$: CE, $|R|$: length of the chosen substring, $|N|$: length of whole feature set, $\alpha \in [0, 1]$ and $\beta = (1 - \alpha)$.	Two methods to address the FS problem are employed in this article. The SA algorithm in WOASA-1 operates as a WOA algorithm operator. WOASA-2 enhances the optimal solution discovered by WOA.
P17 [6]	ABC, PSO	$Fitness = \alpha \cdot \psi_s + \beta \frac{|\mathbb{N}| - |S|}{|N|}$, where ψ_s: classifier performance with the subset S, N: total attribute count, β: size of the attribute subset and α: standard of the classification.	Each employed bee produces a new food source and exploits a good source. Each onlooker bee selects a supply based on the amount of its solution, creates a new food source and exploits the better one. PSO is used instead of scouting bees for the hunt for new sources after deciding the source to be left and assigning the employed bee as scout.
P18 [80]	ACO, PSO	$Fitness_p^t = \frac{1}{BER_p^t}$, where BER_p^t: CE with the attribute vector chosen by particle p at the tth iteration.	Here, both ACO and PSO are executed simultaneously by each individual. Estimate the value of the chosen subset of each ant and particles by the classifier and pick the best one for the next generation.
P19 [81]	MA with KHA	$f_\theta = \omega \times E + (1 - \omega)\frac{\Sigma \theta_i}{N}$, where $f_\theta-$: fitness function with θ number of features, N: total feature count, E: CE, ω: constant.	The suggested MAKHA hybrid technique employs foraging operation and physical random diffusion with crossover and mutation and uses somersault process and watch–jump process from the MA.
P20 [9]	GA, SA	Percentage of recognition of the Bayesian classifier.	SA is used to select the chromosomes for the next generation.
P21 [82]	GA-ACO	CA	In the proposed hybrid method, ACO utilizes the GA's crossover and mutation strategy. This results in the exploration of ants near the optimum solution. The mechanism is again iterated after pheromone upgrading.
P22 [83]	TS-BPSO	The KNN with LOOCV and SVM with OVR serves as estimators of the TS and BPSO objective functions.	In the suggested hybrid architecture, BPSO acts as a local optimization technique for the Tabu search method.

Table 3. *Cont.*

Paper	Search Method	Fitness Function	Means of Hybridization
P23 [84]	SCA, Binary PSO (BPSO)	SI: $\mathrm{Sil}(x_i) = \frac{f(x_i)-p(x_i)}{\max\{f(x_i),p(x_i)\}}$.	The methodology is used in this article to increase the search capabilities and find a close-to-optimum global solution by combining the BPSO with the SCA. In this context, SCA improves the movement of a particle in the BPSO.
P24 [85]	GA, PSO	ANN	The procedure is applied three times at a time and continued until the given number of generations is reached after producing the initial population and determining the cost of each solution. The GA and PSO follow these three moves. Two steps are taken concurrently in the GA, while the PSO takes only a single move.
P25 [86]	ALO, GWO	$f_\theta = \alpha \cdot E + (1-\alpha)\frac{\sum_\theta}{N}$, where f_θ: fitness function considering θ number of features, N: total feature count, E: CE, α: constant.	The suggested hybrid approach updates the ants applying the essence of ALO, and the ant lions, with the help of the GWO concept, that deserve to be converged more quickly.
P26 [87]	ALO, DE	$F = (\max\{ECT_{ik}\}\,\&\min\{Ru_k\})$ $\forall \in [1,N_{tsk}]$ mapped to $k_{\mathrm{th}}\,VM$, where $k = 1,2,3,\dots,N_{vm}$ $i = 1,2,3,\dots\dots,N_{tsk}$ ECT_{ik}: required RT, N_{vm}: #virtual machines, N_{tsk}: #tasks.	In each iteration, the ant lions are updated by using DE operators.
P27 [88]	PSO and CSA	$Fitness = \alpha\Delta_R(D) + \beta\frac{|Y|}{|T|}$, where $\alpha\Delta_R(D)$: classifier's error rate, $|Y|$: size of the subset, and $|T|$: #total features.	This strategy merely targets a few chosen crows with the greatest feeds during the hybridization process to improve the effectiveness of randomly following every crow in the original CSA. The next step is to apply the OBL approach to create the crows' opposite location and update their locations in the PSO. This is done so that the result generated by each method can explore the search space, in turn, without interfering with one another.
P28 [89]	SSA, TLBO	$f(X_i) = (E_{x_i} \times (1+0.5\times \frac{S}{N}))^2$, where $f(X_i)$: fitness value of x_i. E_{x_i}: performance of the classifier, S/N: #features selected/#total features.	During the teaching and learning phase, the population change is accomplished by using the TLBO methodology or the SSA.

Table 3. *Cont.*

Paper	Search Method	Fitness Function	Means of Hybridization
P29 [90]	HHO-CS	$Fitness = \alpha + \beta \frac{|R|}{|C|} - G$, where R: CE, C: total attribute count, β: size of subset, α: performance of the classification.	The merits of the CS approach for controlling HHO vectors in place is taken in the CHHO–CS algorithm. CS attempts to determine the optimal solution after each iteration T. As a result, if the fitness of the current solution is greater than that of the new solution derived from HHO, the new solutions will be determined; otherwise, the older one will stay intact.
P30 [91]	SCA-HHO	$f_i = w_1 \times \epsilon_i + w_2 \times \frac{d_i}{D}, w_1 = 0.99, w_2 = 1 - w_1$, where ϵ_i: error by KNN, d_i: count of attributes picked, D: actual #feature.	SCA and HHO are paired to execute their discovery task by SCA and exploitation by HHO.
P31 [92]	GWO-HHO	$Fitness = \alpha(ER) + (1 - \alpha)\frac{|S_f|}{|T_f|}$, where α lies between 0 and 1, ER: error, S_f: #picked factors and T_f: # actual factors.	Exploration is carried out by HHO, while exploitation is done by GWO.
P32 [93]	ACO-ABC	$fit_j = \frac{1}{1+f_j}$, where f_j: value of the objective for the corresponding attribute set.	The ants use the bees' exploitation to decide the best ant and optimal attribute substring; bees incorporate the attribute substring that the ants create as a supply of food.
P33 [94]	DE-ABC	Weighted average F-measure from J48.	If the fitness probability is > *rnd*, DE mutation is performed; otherwise, ABC neighbourhood solution creation procedure is followed.
P34 [95]	ACO-CS	Mean square error (MSE) of SVM.	ACO is an excellent evolutionary strategy. The disadvantage of this strategy is that the ant moves in the direction where the pheromone density is high, slowing down the operation. CS is therefore used to perform the local ACO scan.
P35 [17]	ChOA and HHO	$Fitness1 = \alpha \times classification\_error + (1 - \alpha) \times \frac{LS}{L}$ and $Fitness2 = \frac{1}{LS}\sum MI(f_i, class) \times \frac{1}{LS}\sum PCC(f_i, class)$, MI: mutual information and PCC: Pearson correlation coefficient.	Hybrid solutions are created based on ChOA and HHO solutions. Then, the best solutions among the ChOA, HHO, and hybrid solutions are treated as the current solution. Then ChOA is used to update the position.

Table 4 gives the details of the classifiers used in the fitness-assessment process, datasets taken for the experiment, and the applications of the mentioned research.

Table 4. Classifiers, datasets used, and application.

Paper	Classifier	Dataset	Application Domain
P1 [64]	KNN	Breast Cancer, Breast EW, Congress EW, Exactly, Exactly 2, Heart EW, Ionosphere EW, Krvskp EW, Lymphography, M of N, Penglung EW, Sonar EW, Spect EW, Tic-tac-toe, Vote, Waveform EW, Wine EW, Zoo.	For FS (Miscellaneous).
P2 [65]		Eight benchmark functions.	For FS (Miscellaneous).
P3 [66]	KNN, NB, RF	UCI (Zoo, Breast Cancer, Breast EW, Congress EW, Exactly, Exactly 2, Heart EW, Ionosphere, KrvskpEW, Lymphography EW, M of N, Penglung EW, Sonar EW, Spect EW, Tic-tac-toe, Vote, Waveform EW, Wine EW).	It focuses on FS (Biology, Politics, Game, Physics, Chemistry, Electromagnetic).
P4 [67]	KNN	Breast Cancer, Tic-tac-toe, Zoo, Wine EW, Spect EW, Sonar EW, Ionosphere EW, Heart EW, Congress EW, Krvskp EW, Waveform EW, Exactly, Exactly 2, M of N, Vote, Breast EW, Semeion, Clean 1, Clean 2, Lymphography, Penghung EW.	FS for COVID-19 classification (Medical).
P5 [68]	KNN	BreastEW, Ionosphere, PenglungEW, Segmentation, Sonar, Vehicle, Bearing dataset, CWRU, and MFPT benchmark dataset.	FS for Fault Diagnosis (Engineering).
P6 [69]	ANN	651 mammograms obtained from the Digital Database for Screening Mammography (DDSM).	FS and classification in mammography (Medical).
P7 [70]	KNN	Hepatitis, Ionosphere, Vertebral, Seeds, Parkinson, Australian, Blood, Breast Cancer, Diabetes, Lymphography, Parkinson, Ring, Titanic, Towonorm, WaveformEW, Tic-Tac-Toe, M of N.	For enhancing FS (Miscellaneous).
P8 [71]	KNN	UCI (Zoo, Breast Cancer, Breast EW, Congress EW, Exactly, Exactly 2, Heart EW, Ionosphere, KrvskpEW, Lymphography EW, M of N, Penglung EW, Sonar EW, Spect EW, Tic-tac-toe, Vote, Waveform EW, Wine EW).	For handling FS tasks (Miscellaneous).
P9 [72]	KNN	UCI (Iris, Wine, Glass, Diabetes, Heartstatlog, Ionosphere, Sonar, Vehicle, Balance Scale, CMC, Cancer, Seed, Blood, Aggregation, Vowel, WBC, Bupa, Jain, Thyroid, WDBC).	For FS (miscellaneous).
P10 [73]	KNN	UCI (BreastCW, Congressional, Connect-Bench, Dermatology, Drug_consumption, Glass, Heart, Hepatitis, Horse-colic, ILPD, Ionosphere, Primary-tumor, Seeds, Soybean, Spambase, SPECT Heart, SteelPF, Thoracic Surgery, Tic-tac-toe, Zoo).	To solve FS issues (Miscellaneous).

Table 4. *Cont.*

Paper	Classifier	Dataset	Application Domain
P11 [74]	KNN	UCI (WBDC, Hepatitis, Heart, Sonar, Lymphography, Clean 1, Breastcancer, Clean 2, Waveform, Ionosphere).	To enhance the FS (Galaxies Classification).
P12 [75]	KNN	UCI (Zoo, Breast Cancer, Congress EW, Exactly, Ionosphere, M of N, Penglung EW, Sonar EW, Vote, Wine EW, Exactly 2, Heart EW, Tic-tac-toe, Waveform EW, Krvskp EW, Lymphography EW, Spect EW, Clean 1, Clean 2, Semeion)	For producing optimistic nominee solutions to obtain global optima efficiently which can be used in solving real-world complex problems and FS (miscellaneous).
P13 [76]	KNN	UCI (Zoo, Breast Cancer, Breast EW, Congress, Exactly, Ionosphere, M of N, Sonar EW, Wine EW, Exactly 2, Heart EW, Tic-tac-toe, Waveform EW, Lymphography EW, Spect EW, Dermatology, Krvskp EW, Echocardiogram, Hepatitis, Lung Cancer).	To solve FS problems (Miscellaneous).
P14 [77]	NB	The dataset consists of 1600 reviews of the 20 well known Chicago hotels that are organized as: 800 positive reviews (400-truthful, 400-deceptive), and 800 negative reviews (400-truthful, 400-deceptive).	It helps with the identification of fake reviews and also to discard irrelevant reviews. It is able to classify efficiently the reviews into spam and ham reviews.
P15 [78]	LR	UCI (Breast, SPECT, Ionosphere, Wine, Congress, Sensor, Clean 1, Clean 2)	To solve FS problems (Miscellaneous).
P16 [79]	KNN	UCI (Zoo, Breast Cancer, Breast EW, Congress EW, Exactly, Exactly 2, Heart EW, Ionosphere, KrvskpEW, Lymphography EW, M of N, Penglung EW, Sonar EW, Spect EW, Tic-tac-toe, Vote, Waveform EW, Wine EW).	To design different FS techniques (Miscellaneous).
P17 [6]	SVM	Three types of datasets are used: 1. 100 recorded speech signals of fruits type, 2. 80 recorded speech signals of animals, and 3. 120 recorded combined speech signals.	Feature selection for automatic speech recognition.
P18 [80]	NB	UCI Machine Learning Repository (SpamBase, BreastCancer, German, Hepatitis, Liver, Musk).	To improve the CA and enhance the FS (Miscellaneous).
P19 [81]	KNN	UCI (Zoo, Breast Cancer, Breast EW, Congress EW, Exactly, Exactly 2, Heart EW, Ionosphere, KrvskpEW, Lymphography EW, M of N, Penglung EW, Sonar EW, Spect EW, Tic-tac-toe, Vote, Waveform EW, Wine EW).	To enhance the FS and to increase the classification performance (Miscellaneous).
P20 [9]	Bayesian	Handwritten Farsi characters having 100 samples for each 33 characters.	For the identification of handprinted Farsi characters.

Table 4. *Cont.*

Paper	Classifier	Dataset	Application Domain
P21 [82]	KNN	GPCR-PROSITE dataset, ENZYME-PROSITE dataset.	For FS in protein function prediction.
P22 [83]	KNN, SVM	Tumors 9, Tumors 11, Brain Tumor 1, Tumors 14, Brain Tumor 2, Leukemia 1, Leukemia 2, Lung Cancer, SRBCT, Prostate Tumor and diffuse large B-cell lymphoma datasets.	To improve gene selection in medical diagnosis.
P23 [84]		Ionosphere, Breast Cancer Wisconsin, Connectionist Bench, Statlog, Parkinson, 9_Tumors, Leukemia2.	To solve FS problems (Miscellaneous).
P24 [85]	Five nearest neighbors (5-NN)	Hill-Valley, Gas 6, Musk 1, Madelon, Isolet 5, Lung.	To solve the FS problem in high-dimensional datasets (Miscellaneous).
P25 [86]	KNN	UCI (Zoo, Breast Cancer, Breast EW, Congress EW, Exactly, Exactly 2, Heart EW, Ionosphere, KrvskpEW, Lymphography EW, M of N, Penglung EW, Sonar EW, Spect EW, Tic-tac-toe, Vote, Waveform EW, Wine EW).	To select significant features from datasets (Bioinformatics).
P26 [87]		Synthetic and real trace datasets.	To solve task scheduling problems in cloud computing environments.
P27 [88]	KNN	Wine, Dermatology, Heart, Ionosphere, Lung cancer, Thoracic surgery, Hepatitis, Parkinson, Phishing website, Qsar biodegradation, Absenteeism at work, Divorce, Wpdc, Risk factor cervical cancer, Wdpc.	For the FS task (Miscellaneous).
P28 [89]	ANN	Digital Database for Screening Mammography (DDSM), Breast Cancer Wisconsin (WBC) dataset.	To solve the FS (Medical).
P29 [90]	SVM	BreastCancer, KCL, WineEW, WDBC, Lung-Cancer, Diabetic, Stock, Scene, Lymphography, and Parkinson.	For chemical descriptor selection and chemical compound activities (Chemical Engineering).
P30 [91]	KNN	Exactly, Exactly 2, Lymphography, Spect EW, Congress EW, Ionosphere EW, Vote, Wine EW, Breast EW, Brain Tumors 1, Tumors 11, Leukemia 2, SRBCT, DLBCL, Prostate Tumors and Tumors 14.	To boost the FS process (Miscellaneous).
P31 [92]	KNN	Breast Cancer, Breast EW, Congress EW, Exactly, Exactly 2, Heart EW, Ionosphere EW, Krvskp EW, Lymphography, M of N, Penglung EW, Sonar EW, Spect EW, Tic-tac-toe, Vote, Waveform EW, Wine EW and Zoo.	the FS task (Miscellaneous).
P32 [93]	DT (J48)	Heart-Cleveland, Dermatology, Hepatitis, Lung Cancer, Lymphography, Pima Indian Diabetes, Iris, Breast Cancer W, Diabetes, Heart-Stalog, Thyroid, Sonar, Gene.	For FS in classification (Miscellaneous).

Table 4. *Cont.*

Paper	Classifier	Dataset	Application Domain
P33 [94]	J48	Autos, Breast-w, Car, Glass, Heart-C, Dermatology, Hepatitis, Thoraric-Surgery, Lymph, Credit-g, Sonar, Ionosphere, Liver-Disorders, Vote, Zoo.	For the FS tasks in classification (Miscellaneous).
P34 [95]	SVM with RBF kernel	miniMIAS data with 100 mammograms (50-normal, 50-abnormal).	For FS in digital mammogram (Medical).
P35 [17]	KNN	Lymphography, Diabetic, Cardiotocography, Cervical Cancer, Lung Cancer, Arrhythmia, Parkinson, Colon Tumor, Leukemia and three COVID-19 datasets.	To enhance FS for COVID-19 prediction (Medical).

Finally, the descriptions of the classifiers used by the aforementioned articles as wrappers are given in Table 5.

Table 5. Summary of classifiers used.

Paper	Classifier	Description
P3 [66]	RF (Random Forest)	RF is a selection algorithm made up of several decision trees. It builds each particular tree by using bagging and features variability and attempts to generate a nonoverlapping forest of trees whose forecast is more reliable than that of any individual.
P1 [64], P3 [66], P4 [67], P5 [68], P7 [70], P8 [71], P9 [72], P10 [73], P11 [74], P12 [75], P13 [76], P16 [79], P19 [81], P21 [82], P22 [83], P24 [85], P25 [86], P27 [88], P30 [91], P31 [92], P35 [17]	KNN	KNN is a straightforward classifier that records all available samples and categorizes new samples focusing on a similarity metric. It is usually used to categorize a piece of data depending on how its neighbors are graded.
P3 [66], P14 [77], P18 [80]	NB	An NB model believes that the existence of one attribute in a class has no influence on the existence of any other attribute.
P17 [6], P22 [83], P29 [90], P34 [95]	SVM	SVM is a supervised ML model for two-group classification tasks. They will identify a new instance after providing an SVM with sets of named training examples for each type.
P15 [78]	LR	The logistic sigmoid is utilised to convert the outcome of the LR classification method, which assigns samples to a discrete set of classes and then returns a probability value.

Table 5. *Cont.*

Paper	Classifier	Description
P20 [9]	Bayesian	The Bayesian classification model forecasts membership estimates for every group, such as the likelihood that a certain record contributes to a certain class. The category with the greatest likelihood is thought to be the most probable.
P32 [93], P33 [94]	DT, J48 classifier	Ross Quinlan's DT is generated by using C4.5 (J48). C4.5 is an extension of the previous ID3 algorithm of Quinlan. For classification, the DTs generated by C4.5 can be used and, hence, C4.5 is also called a statistical classifier.
P6 [69], P28 [89]	ANN	To provide computer programs with the ability to process information and make judgments similarly to those of humans, ANN models are used in AI to simulate the networks of neurons that make up the human brain. In order to develop an ANN, an artificial model is designed and programmed to operate analogously to a network of interconnected neurons and brain cells.

4. Analysis and Discussion

According to the analysis of the mentioned articles, the majority of studies employed the wrapper strategy mainly due to its supremacy in terms of higher accuracy compared to filter techniques, which have consistently been shown through experimentation that perform inaccurate filtration. In an effort to utilize all the advantages from both approaches, numerous researchers have tried to integrate and hybridize filter and wrapper methods. Figure 4 displays the number of papers broken down by the evolutionary methods that were employed in the corresponding studies.
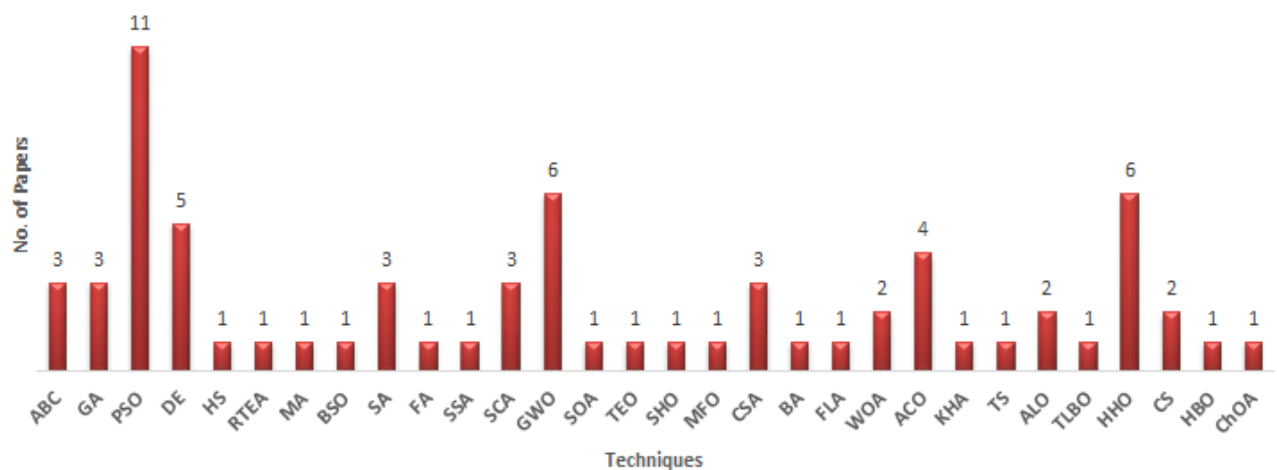

Figure 4. Number of papers per technique.

These results demonstrate unequivocally that PSO is utilised for fusion in the most number of research articles (11). This is most likely caused by the PSO's lack of derivatives and its simpler concept and coding technology compared to other evolutionary methods. In particular, in contrast to the other competing evolutionary methods, PSO uses two acceleration coefficients (i.e., the cognitive and social parameters, respectively) and an inertia weight, and thus in PSO there are few parameters required to be adjusted. Furthermore, on average the convergence rate of PSO is faster than other stochastic algorithms [96].

Feature-selection techniques can be applicable to any area where there is a chance of facing the "curse of dimensionality problem". However, after studying the presented works (Figure 5), we found that most of the hybrid FS techniques (54%) have verified

their performance by considering some benchmark datasets. Only 22% of the total articles have applied their technique to the biomedical area (microarray gene selection, disease diagnosis etc.).
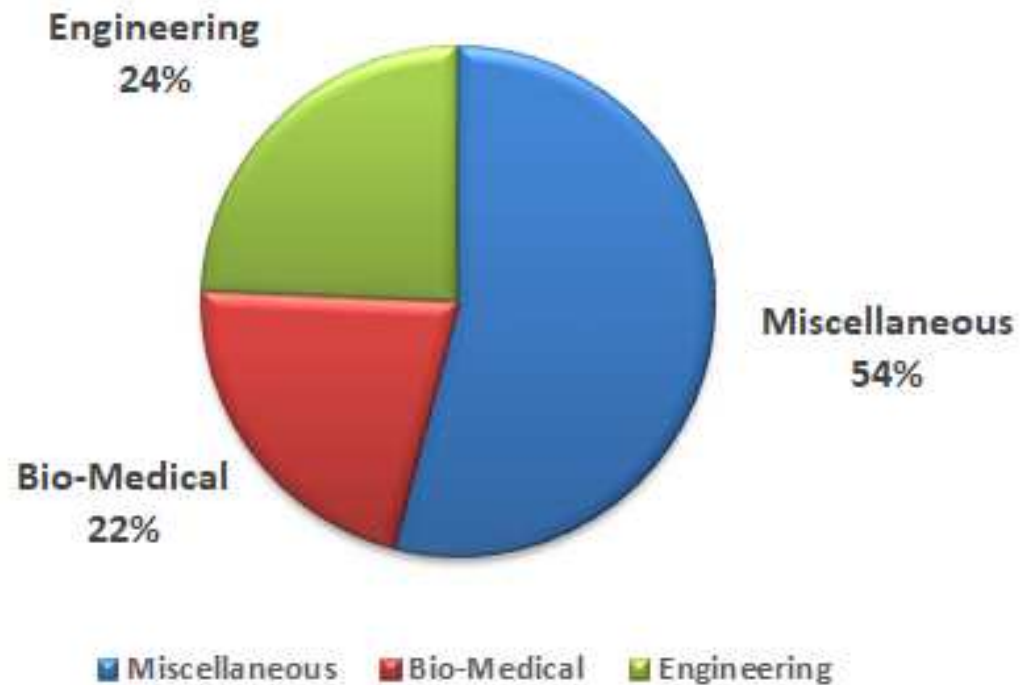

Figure 5. Distribution of papers according to application area.

Additionally, Figure 6 displays the quantity of papers by using different standard classifiers as wrappers. Because it is simple to grasp and requires less calculation time, 21 out of 35 research employed KNN as a wrapper in their fitness computation procedure. KNN's training procedure is also incredibly fast because it makes decisions without using any training data.

Figure 6. Number of papers vs. classifier.

Additionally, the bulk of FS researchers using hybrid evolutionary techniques have the goal of reducing both the number of features and the error rate. This is a challenging task,

despite the inadequacies of many ECs. The recommended hybrid approaches, however, surpass the current FS methodology after hybridization, which looks for comparable alternatives to produce the best results when tackling optimization tasks, according to tests performed on various datasets. By combining and merging the exploration and exploitation processes, this is achieved. Overall, the previous studies have led to several enhancements and alterations, and, in short, each specific research approach requires the employment of unique approaches in order to provide the required outcomes. The solution model may change over time because there is no one technique that can be used to solve every problem.

After investigating the abovementioned works on hybrid evolutionary algorithms for FS in classification, we are able to list out the following advantages of hybridization.

- Efficiency of the base algorithm can be improved (P1 [64], P4 [67], P8 [71], P11 [74], P14 [77], P20 [9], P22 [83], P25 [86], P26 [87], P28 [89], P35 [17]).
- Premature convergence and the local optimum trap issue can be addressed (P2 [65], P5 [68], P6 [69], P11 [74], P15 [78], P31 [92], P32 [93], P33 [94], P34 [95], P35 [17]).
- Balance between both exploration and exploitation can be maintained (P3 [66], P7 [70], P8 [71], P19 [81], P29 [90]).
- The poor exploitation capability of some of the base methods can be improved (P9 [72], P21 [82], P23 [84], P30 [91], P35).
- The optimal solution identified in each iteration can be enhanced (P10 [73], P12 [75], P16 [79], P18 [80]).
- The searching procedure can converge to the best global solution (P12 [75], P13 [76], P17 [6], P24 [85], P27 [88]).

Although the presented articles are able to improve the performance of the FS techniques, they still have some limitations that point in the direction of future research.

- They are not verified with real-world applications like the biomedical domain (P1 [64], P3 [66], P8 [71], P10 [73], P12 [75], P13 [76], P15 [78], P16 [79], P31 [92], P32 [93]).
- They are not tested with high-dimensional datasets (P1 [64], P3 [66], P4 [67], P6 [69], P7 [70], P8 [71], P9 [72], P10 [73], P11 [74], P12 [75], P13 [76], P15 [78], P16 [79], P18 [80], P19 [81], P24 [85], P27 [88], P29 [90], P32 [93], P33 [94]).
- In some cases, the proposed algorithm is unable to find the global optimum (P3 [66], P14 [77], P27 [88]).
- The fitness value focused only on the error rate and not on the number of features (P5 [68]).
- They take longer to execute (P35 [17]).
- The performance of the proposed approach is not compared with other existing hybrid approaches (P6 [69], P7 [70], P10 [73], P11 [74], P12 [75], P15 [78], P16 [79], P17 [6], P18 [80], P19 [81], P20 [9], P21 [82], P22 [83], P23 [84], P25 [86], P26 [87], P27 [88], P28 [89], P30 [91], P34 [95]).
- They are verified with a few datasets (P14 [77], P17 [6], P20 [9], P21 [82]).

As hybrid models for FS are becoming more effective and more efficient solutions, the following concerns have to be analysed by performing further enhancement.

- The capability of newly developed methods has not been thoroughly explored, particularly in terms of their scalability, and therefore additional research is suggested for FS in high-dimensional real-world applications.
- Since computation complexity is one of the key issues in most hybrid approaches for FS, it is recommended that more appropriate measures to reduce computational complexity should be proposed. Two key considerations must be weighed in order to do so: (1) more efficient ways to perform searching in the large solution spaces and (2) faster evaluation tools.
- The FS priorities, such as computational burden and space complexity, can indeed be viewed in combination with the two main objectives of the hybrid FS problem (i.e., exploration and exploitation).

- Proposing new methodologies that soften the fitness landscape will significantly reduce the problem's complexities and motivate the development of more effective search strategies.
- Most of the existing studies in the literature used only one fitness function. However, FS can be viewed as an MO problem and thus, the application of hybridization in multiobjective FS tasks is an open research domain for researchers.
- As hybrid FS techniques are time-consuming as compared to the others, employing parallel processing during the FS phase is also an area of research to be explored.
- Most of the abovementioned articles are wrapper-based; however, the optimal solutions generated by wrapper approaches are less generic. Therefore, a hybrid–hybrid approach (i.e., hybridising filter and wrapper criteria while mixing evolutionary techniques) for FS is a challenging research domain.
- Feature selection plays a vital role in the biomedical area due to the high dimensionality of the data. However, very few works (22%) explored their techniques in this field. Therefore, the application of hybrid FS techniques to biomedical data is a very good research area for the future.

5. Conclusions and Future Work

Over the years, academics conducting knowledge extraction and elicitation research have emphasized hybrid metaheuristic approaches for optimal feature identification and selection. The "No Free Lunch" (NFL) theorem states that there has never been and will never be an optimization method that can adequately handle all problems. Therefore, in this paper we tried a systematized analysis of the literature, taking into account research works released from 2009 to 2022, to point out the key difficulties and strategies for hybrid FS and provide a comprehensive investigation of the metaheuristic approaches employed in the development of hybridized FS techniques. According to the survey's findings, substantial efforts have been made to improve metaheuristic wrapper FS methods' performance through hybridization in terms of the precision and the size of the considered feature subsets, paving the path for potential advancements. Finally, since there is still room for further development, any hybrid evolutionary FS technique should be extended into a variety of hybridization strategies and variations based on the needs of the specific problems under consideration. As a result, researchers studying hybrid evolutionary methods for addressing FS tasks could use the results of this review study to further investigate more effective and efficient techniques for solving the latest challenges in FS.

Author Contributions: J.P., P.M., R.D., B.A., V.C.G. and A.K.; methodology, J.P., P.M., R.D., B.A., V.C.G. and A.K.; validation, J.P., P.M., R.D., B.A., V.C.G. and A.K.; formal analysis, J.P., P.M., R.D., B.A., V.C.G. and A.K.; investigation, J.P., P.M., R.D., B.A., V.C.G. and A.K.; writing—original draft preparation, J.P., P.M., R.D., B.A., V.C.G. and A.K.; writing—review and editing, J.P., P.M., R.D., B.A., V.C.G. and A.K. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Piri, J.; Mohapatra, P.; Dey, R. Fetal Health Status Classification Using MOGA—CD Based Feature Selection Approach. In Proceedings of the IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2–4 July 2020; pp. 1–6.
2. Bhattacharyya, T.; Chatterjee, B.; Singh, P.K.; Yoon, J.H.; Geem, Z.W.; Sarkar, R. Mayfly in Harmony: A New Hybrid Meta-Heuristic Feature Selection Algorithm. *IEEE Access* **2020**, *8*, 195929–195945. [CrossRef]
3. Piri, J.; Mohapatra, P. Exploring Fetal Health Status Using an Association Based Classification Approach. In Proceedings of the IEEE International Conference on Information Technology (ICIT), Bhubaneswar, India, 19–21 December 2019; pp. 166–171.
4. Piri, J.; Mohapatra, P.; Acharya, B.; Gharehchopogh, F.S.; Gerogiannis, V.C.; Kanavos, A.; Manika, S. Feature Selection Using Artificial Gorilla Troop Optimization for Biomedical Data: A Case Analysis with COVID-19 Data. *Mathematics* **2022**, *10*, 2742. [CrossRef]

5. Jain, D.; Singh, V. Diagnosis of Breast Cancer and Diabetes using Hybrid Feature Selection Method. In Proceedings of the 5th International Conference on Parallel, Distributed and Grid Computing (PDGC), Solan, India, 20–22 December 2018; pp. 64–69.

6. Mendiratta, S.; Turk, N.; Bansal, D. Automatic Speech Recognition using Optimal Selection of Features based on Hybrid ABC-PSO. In Proceedings of the IEEE International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–27 August 2016; Volume 2, pp. 1–7.

7. Naik, A.; Kuppili, V.; Edla, D.R. Binary Dragonfly Algorithm and Fisher Score Based Hybrid Feature Selection Adopting a Novel Fitness Function Applied to Microarray Data. In Proceedings of the International IEEE Conference on Applied Machine Learning (ICAML), Bhubaneswar, India, 27–28 September 2019; pp. 40–43.

8. Monica, K.M.; Parvathi, R. Hybrid FOW—A Novel Whale Optimized Firefly Feature Selector for Gait Analysis. *Pers. Ubiquitous Comput.* **2021**, 1–13. [CrossRef]

9. Azmi, R.; Pishgoo, B.; Norozi, N.; Koohzadi, M.; Baesi, F. A Hybrid GA and SA Algorithms for Feature Selection in Recognition of Hand-printed Farsi Characters. In Proceedings of the IEEE International Conference on Intelligent Computing and Intelligent Systems, Xiamen, China, 29–31 October 2010; Volume 3, pp. 384–387.

10. Al-Tashi, Q.; Abdulkadir, S.J.; Rais, H.M.; Mirjalili, S.; Alhussian, H. Approaches to Multi-Objective Feature Selection: A Systematic Literature Review. *IEEE Access* **2020**, *8*, 125076–125096. [CrossRef]

11. Brezočnik, L.; Fister, I.; Podgorelec, V. Swarm Intelligence Algorithms for Feature Selection: A Review. *Appl. Sci.* **2018**, *8*, 1521. [CrossRef]

12. Venkatesh, B.; Anuradha, J. A Review of Feature Selection and Its Methods. *Cybern. Inf. Technol.* **2019**, *19*, 3–26. [CrossRef]

13. Abd-Alsabour, N. A Review on Evolutionary Feature Selection. In Proceedings of the IEEE European Modelling Symposium, Pisa, Italy, 21–23 October 2014; pp. 20–26.

14. Wolpert, D.H.; Macready, W.G. No Free Lunch Theorems for Optimization. *IEEE Trans. Evol. Comput.* **1997**, *1*, 67–82. [CrossRef]

15. Cheng, M.Y.; Prayogo, D. Symbiotic Organisms Search: A new Metaheuristic Optimization Algorithm. *Comput. Struct.* **2014**, *139*, 98–112. [CrossRef]

16. Singh, N.; Son, L.H.; Chiclana, F.; Magnot, J.P. A new Fusion of Salp Swarm with Sine Cosine for Optimization of Non-Linear Functions. *Eng. Comput.* **2020**, *36*, 185–212. [CrossRef]

17. Piri, J.; Mohapatra, P.; Singh, H.K.R.; Acharya, B.; Patra, T.K. An Enhanced Binary Multiobjective Hybrid Filter-Wrapper Chimp Optimization Based Feature Selection Method for COVID-19 Patient Health Prediction. *IEEE Access* **2022**, *10*, 100376–100396. [CrossRef]

18. Piri, J.; Mohapatra, P.; Dey, R.; Panda, N. Role of Hybrid Evolutionary Approaches for Feature Selection in Classification: A Review. In Proceedings of the International Conference on Metaheuristics in Software Engineering and its Application, Marrakech, Morocco, 27–30 October 2022; pp. 92–103.

19. Blum, A.; Langley, P. Selection of Relevant Features and Examples in Machine Learning. *Artif. Intell.* **1997**, *97*, 245–271. [CrossRef]

20. Liu, H.; Motoda, H. *Feature Selection for Knowledge Discovery and Data Mining*; The Springer International Series in Engineering and Computer Science; Springer: Berlin/Heidelberg, Germany, 1998; Volume 454.

21. Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.

22. Pudil, P.; Novovicová, J.; Kittler, J. Floating Search Methods in Feature Selection. *Pattern Recognit. Lett.* **1994**, *15*, 1119–1125. [CrossRef]

23. Mao, Q.; Tsang, I.W. A Feature Selection Method for Multivariate Performance Measures. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2051–2063. [CrossRef]

24. Min, F.; Hu, Q.; Zhu, W. Feature Selection with Test Cost Constraint. *Int. J. Approx. Reason.* **2014**, *55*, 167–179. [CrossRef]

25. Vivekanandan, T.; Iyengar, N.C.S.N. Optimal Feature Selection using a Modified Differential Evolution Algorithm and its Effectiveness for Prediction of Heart Disease. *Comput. Biol. Med.* **2017**, *90*, 125–136. [CrossRef]

26. Sahebi, G.; Movahedi, P.; Ebrahimi, M.; Pahikkala, T.; Plosila, J.; Tenhunen, H. GeFeS: A Generalized Wrapper Feature Selection Approach for Optimizing Classification Performance. *Comput. Biol. Med.* **2020**, *125*, 103974. [CrossRef]

27. Al-Tashi, Q.; Rais, H.; Jadid, S. Feature Selection Method Based on Grey Wolf Optimization for Coronary Artery Disease Classification. In Proceedings of the International Conference of Reliable Information and Communication Technology, Kuala Lumpur, Malaysia, 23–24 July 2018; pp. 257–266.

28. Too, J.; Abdullah, A.R. Opposition based Competitive Grey Wolf Optimizer for EMG Feature Selection. *Evol. Intell.* **2021**, *14*, 1691–1705. [CrossRef]

29. Aghdam, M.H.; Ghasem-Aghaee, N.; Basiri, M.E. Text Feature Selection using Ant Colony Optimization. *Expert Syst. Appl.* **2009**, *36*, 6843–6853. [CrossRef]

30. Erguzel, T.T.; Tas, C.; Cebi, M. A Wrapper-based Approach for Feature Selection and Classification of Major Depressive Disorder-Bipolar Disorders. *Comput. Biol. Med.* **2015**, *64*, 127–137. [CrossRef]

31. Huang, H.; Xie, H.; Guo, J.; Chen, H. Ant Colony Optimization-based Feature Selection Method for Surface Electromyography Signals Classification. *Comput. Biol. Med.* **2012**, *42*, 30–38. [CrossRef]

32. Piri, J.; Mohapatra, P. An Analytical Study of Modified Multi-objective Harris Hawk Optimizer towards Medical Data Feature Selection. *Comput. Biol. Med.* **2021**, *135*, 104558. [CrossRef] [PubMed]

33. Too, J.; Abdullah, A.R.; Saad, N.M. A New Quadratic Binary Harris Hawk Optimization for Feature Selection. *Electronics* **2019**, *8*, 1130. [CrossRef]

34. Zhang, Y.; Liu, R.; Wang, X.; Chen, H.; Li, C. Boosted Binary Harris Hawks Optimizer and Feature Selection. *Eng. Comput.* **2021**, *37*, 3741–3770. [CrossRef]

35. Emary, E.; Zawbaa, H.M.; Hassanien, A.E. Binary Ant Lion Approaches for Feature Selection. *Neurocomputing* **2016**, *213*, 54–65. [CrossRef]

36. Piri, J.; Mohapatra, P.; Dey, R. Multi-objective Ant Lion Optimization Based Feature Retrieval Methodology for Investigation of Fetal Wellbeing. In Proceedings of the 3rd IEEE International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 21–23 September 2021; pp. 1732–1737.

37. Hegazy, A.E.; Makhlouf, M.A.A.; El-Tawel, G.S. Improved Salp Swarm Algorithm for Feature Selection. *J. King Saud Univ.-Comput. Inf. Sci.* **2020**, *32*, 335–344. [CrossRef]

38. Mafarja, M.M.; Aljarah, I.; Heidari, A.A.; Faris, H.; Fournier-Viger, P.; Li, X.; Mirjalili, S. Binary Dragonfly Optimization for Feature Selection using Time-varying Transfer Functions. *Knowl. Based Syst.* **2018**, *161*, 185–204. [CrossRef]

39. Sreejith, S.; Nehemiah, H.K.; Kannan, A. Clinical Data Classification using an Enhanced SMOTE and Chaotic Evolutionary Feature Selection. *Comput. Biol. Med.* **2020**, *126*, 103991. [CrossRef]

40. Das, H.; Naik, B.; Behera, H.S. A Jaya Algorithm based Wrapper Method for Optimal Feature Selection in Supervised Classification. *J. King Saud Univ.-Comput. Inf. Sci.* **2020**, *34*, 3851–3863. [CrossRef]

41. Tiwari, V.; Jain, S.C. An Optimal Feature Selection Method for Histopathology Tissue Image Classification using Adaptive Jaya Algorithm. *Evol. Intell.* **2021**, *14*, 1279–1292. [CrossRef]

42. Haouassi, H.; Merah, E.; Rafik, M.; Messaoud, M.T.; Chouhal, O. A new Binary Grasshopper Optimization Algorithm for Feature Selection Problem. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 316–328.

43. Mohan, A.; Nandhini, M. Optimal Feature Selection using Binary Teaching Learning based Optimization Algorithm. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 329–341.

44. Dash, R. An Adaptive Harmony Search Approach for Gene Selection and Classification of High Dimensional Medical Data. *J. King Saud Univ.-Comput. Inf. Sci.* **2021**, *33*, 195–207. [CrossRef]

45. Gharehchopogh, F.S.; Maleki, I.; Dizaji, Z.A. Chaotic Vortex Search Algorithm: Metaheuristic Algorithm for Feature Selection. *Evol. Intell.* **2022**, *15*, 1777–1808. [CrossRef]

46. Mitchell, M. *An Introduction to Genetic Algorithms*; MIT Press: Cambridge, MA, USA, 1998.

47. Kennedy, J.; Eberhart, R. Particle Swarm Optimization. In Proceedings of the IEEE International Conference on Neural Networks (ICNN), Perth, WA, Australia, 27 November–1 December 1995; pp. 1942–1948.

48. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]

49. Heidari, A.A.; Mirjalili, S.; Faris, H.; Aljarah, I.; Mafarja, M.M.; Chen, H. Harris Hawks Optimization: Algorithm and Applications. *Future Gener. Comput. Syst.* **2019**, *97*, 849–872. [CrossRef]

50. Liu, H.; Zhao, Z. Manipulating Data and Dimension Reduction Methods: Feature Selection. In *Encyclopedia of Complexity and Systems Science*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 5348–5359.

51. Liu, H.; Motoda, H.; Setiono, R.; Zhao, Z. Feature Selection: An Ever Evolving Frontier in Data Mining. In Proceedings of the 4th International Workshop on Feature Selection in Data Mining (FSDM), Hyderabad, India, 21 June 2010; Volume 10, pp. 4–13.

52. Xue, B.; Zhang, M.; Browne, W.N. Particle Swarm Optimization for Feature Selection in Classification: A Multi-Objective Approach. *IEEE Trans. Cybern.* **2013**, *43*, 1656–1671. [CrossRef]

53. Dash, M.; Liu, H. Feature Selection for Classification. *Intell. Data Anal.* **1997**, *1*, 131–156. [CrossRef]

54. Kira, K.; Rendell, L.A. A Practical Approach to Feature Selection. In Proceedings of the 9th International Workshop on Machine Learning (ML), San Francisco, CA, USA, 1–3 July 1992; Morgan Kaufmann: Burlington, MA, USA, 1992; pp. 249–256.

55. Wang, S.; Pedrycz, W.; Zhu, Q.; Zhu, W. Subspace learning for unsupervised feature selection via matrix factorization. *Pattern Recognit.* **2015**, *48*, 10–19. [CrossRef]

56. Peng, H.; Long, F.; Ding, C.H.Q. Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [CrossRef]

57. Cervante, L.; Xue, B.; Zhang, M.; Shang, L. Binary Particle Swarm Optimisation for Feature Selection: A Filter based Approach. In Proceedings of the IEEE Congress on Evolutionary Computation (CEC), Brisbane, Australia, 10–15 June 2012; pp. 1–8.

58. Ünler, A.; Murat, A.E.; Chinnam, R.B. mr^{2}PSO: A Maximum Relevance Minimum Redundancy Feature Selection Method based on Swarm Intelligence for Support Vector Machine Classification. *Inf. Sci.* **2011**, *181*, 4625–4641. [CrossRef]

59. Tan, N.C.; Fisher, W.G.; Rosenblatt, K.P.; Garner, H.R. Application of Multiple Statistical Tests to Enhance Mass Spectrometry-based Biomarker Discovery. *BMC Bioinform.* **2009**, *10*, 144. [CrossRef]

60. Tan, M.; Tsang, I.W.; Wang, L. Minimax Sparse Logistic Regression for Very High-Dimensional Feature Selection. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 1609–1622. [CrossRef] [PubMed]

61. Zhai, Y.; Ong, Y.; Tsang, I.W. The Emerging "Big Dimensionality". *IEEE Comput. Intell. Mag.* **2014**, *9*, 14–26. [CrossRef]

62. Thiele, L.; Miettinen, K.; Korhonen, P.J.; Luque, J.M. A Preference-Based Evolutionary Algorithm for Multi-Objective Optimization. *Evol. Comput.* **2009**, *17*, 411–436. [CrossRef] [PubMed]

63. Bui, L.T.; Alam, S. *Multi-Objective Optimization in Computational Intelligence: Theory and Practice*; IGI Global: Hershey, PA, USA, 2008.

64. Al-Wajih, R.; Abdulkadir, S.J.; Alhussian, H.; Aziz, N.; Al-Tashi, Q.; Mirjalili, S.; Alqushaibi, A. Hybrid Binary Whale with Harris Hawks for Feature Selection. *Neural Comput. Appl.* **2022**, *34*, 19377–19395. [CrossRef]

65. Ajibade, S.S.M.; Ahmad, N.B.B.; Zainal, A. A Hybrid Chaotic Particle Swarm Optimization with Differential Evolution for Feature Selection. In Proceedings of the IEEE Symposium on Industrial Electronics & Applications (ISIEA), Kristiansand, Norway, 9–13 November 2020; pp. 1–6.

66. Ahmed, S.; Ghosh, K.K.; Singh, P.K.; Geem, Z.W.; Sarkar, R. Hybrid of Harmony Search Algorithm and Ring Theory-Based Evolutionary Algorithm for Feature Selection. *IEEE Access* **2020**, *8*, 102629–102645. [CrossRef]

67. Bezdan, T.; Zivkovic, M.; Bacanin, N.; Chhabra, A.; Suresh, M. Feature Selection by Hybrid Brain Storm Optimization Algorithm for COVID-19 Classification. *J. Comput. Biol.* **2022**, *29*, 515–529. [CrossRef]

68. Lee, C.; Le, T.; Lin, Y. A Feature Selection Approach Hybrid Grey Wolf and Heap-Based Optimizer Applied in Bearing Fault Diagnosis. *IEEE Access* **2022**, *10*, 56691–56705. [CrossRef]

69. Thawkar, S. Feature Selection and Classification in Mammography using Hybrid Crow Search Algorithm with Harris Hawks Optimization. *Biocybern. Biomed. Eng.* **2022**, *42*, 1094–1111. [CrossRef]

70. El-Kenawy, E.S.; Eid, M. Hybrid Gray Wolf and Particle Swarm Optimization for Feature Selection. *Int. J. Innov. Comput. Inf. Control* **2020**, *16*, 831–844.

71. Al-Tashi, Q.; Abdulkadir, S.J.; Rais, H.M.; Mirjalili, S.; Alhussian, H. Binary Optimization Using Hybrid Grey Wolf Optimization for Feature Selection. *IEEE Access* **2019**, *7*, 39496–39508. [CrossRef]

72. Jia, H.; Xing, Z.; Song, W. A New Hybrid Seagull Optimization Algorithm for Feature Selection. *IEEE Access* **2019**, *7*, 49614–49631. [CrossRef]

73. Jia, H.; Li, J.; Song, W.; Peng, X.; Lang, C.; Li, Y. Spotted Hyena Optimization Algorithm with Simulated Annealing for Feature Selection. *IEEE Access* **2019**, *7*, 71943–71962. [CrossRef]

74. Aziz, M.A.E.; Ewees, A.A.; Ibrahim, R.A.; Lu, S. Opposition-based Moth-flame Optimization Improved by Differential Evolution for Feature Selection. *Math. Comput. Simul.* **2020**, *168*, 48–75.

75. Arora, S.; Singh, H.; Sharma, M.; Sharma, S.; Anand, P. A New Hybrid Algorithm Based on Grey Wolf Optimization and Crow Search Algorithm for Unconstrained Function Optimization and Feature Selection. *IEEE Access* **2019**, *7*, 26343–26361. [CrossRef]

76. Tawhid, M.A.; Dsouza, K.B. Hybrid Binary Bat Enhanced Particle Swarm Optimization Algorithm for Solving Feature Selection Problems. *Appl. Comput. Inform.* **2018**, *16*, 117–136. [CrossRef]

77. Rajamohana, S.P.; Umamaheswari, K. Hybrid Approach of Improved Binary Particle Swarm Optimization and Shuffled Frog Leaping for Feature Selection. *Comput. Electr. Eng.* **2018**, *67*, 497–508. [CrossRef]

78. Elaziz, M.E.A.; Ewees, A.A.; Oliva, D.; Duan, P.; Xiong, S. A Hybrid Method of Sine Cosine Algorithm and Differential Evolution for Feature Selection. In Proceedings of the 24th International Conference on Neural Information Processing (ICONIP), Guangzhou, China, 14–18 November 2017; Volume 10638, pp. 145–155.

79. Mafarja, M.M.; Mirjalili, S. Hybrid Whale Optimization Algorithm with Simulated Annealing for Feature Selection. *Neurocomputing* **2017**, *260*, 302–312. [CrossRef]

80. Menghour, K.; Souici-Meslati, L. Hybrid ACO-PSO Based Approaches for Feature Selection. *Int. J. Intell. Eng. Syst.* **2016**, *9*, 65–79. [CrossRef]

81. Hafez, A.I.; Hassanien, A.E.; Zawbaa, H.M.; Emary, E. Hybrid Monkey Algorithm with Krill Herd Algorithm optimization for Feature Selection. In Proceedings of the 11th IEEE International Computer Engineering Conference (ICENCO), Cairo, Egypt, 29–30 December 2015; pp. 273–277.

82. Nemati, S.; Basiri, M.E.; Ghasem-Aghaee, N.; Aghdam, M.H. A Novel ACO-GA Hybrid Algorithm for Feature Selection in Protein Function Prediction. *Expert Syst. Appl.* **2009**, *36*, 12086–12094. [CrossRef]

83. Chuang, L.; Yang, C.; Yang, C. Tabu Search and Binary Particle Swarm Optimization for Feature Selection Using Microarray Data. *J. Comput. Biol.* **2009**, *16*, 1689–1703. [CrossRef] [PubMed]

84. Kumar, L.; Bharti, K.K. A Novel Hybrid BPSO-SCA Approach for Feature Selection. *Nat. Comput.* **2021**, *20*, 39–61. [CrossRef]

85. Moslehi, F.; Haeri, A. A Novel Hybrid Wrapper-filter Approach based on Genetic Algorithm, Particle Swarm Optimization for Feature Subset Selection. *J. Ambient Intell. Humaniz. Comput.* **2020**, *11*, 1105–1127. [CrossRef]

86. Zawbaa, H.M.; Emary, E.; Grosan, C.; Snásel, V. Large-dimensionality Small-instance Set Feature Selection: A Hybrid Bio-inspired Heuristic Approach. *Swarm Evol. Comput.* **2018**, *42*, 29–42. [CrossRef]

87. Abualigah, L.M.; Diabat, A. A Novel Hybrid Antlion Optimization Algorithm for Multi-objective Task Scheduling Problems in Cloud Computing Environments. *Clust. Comput.* **2021**, *24*, 205–223. [CrossRef]

88. Adamu, A.; Abdullahi, M.; Junaidu, S.B.; Hassan, I.H. An Hybrid Particle Swarm Optimization with Crow Search Algorithm for Feature Selection. *Mach. Learn. Appl.* **2021**, *6*, 100108. [CrossRef]

89. Thawkar, S. A Hybrid Model using Teaching-learning-based Optimization and Salp Swarm Algorithm for Feature Selection and Classification in Digital Mammography. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 8793–8808. [CrossRef]

90. Houssein, E.H.; Hosney, M.E.; Elhoseny, M.; Oliva, D.; Mohamed, W.M.; Hassaballah, M. Hybrid Harris Hawks Optimization with Cuckoo Search for Drug Design and Discovery in Chemoinformatics. *Sci. Rep.* **2020**, *10*, 1–22. [CrossRef]

91. Hussain, K.; Neggaz, N.; Zhu, W.; Houssein, E.H. An Efficient Hybrid Sine-cosine Harris Hawks Optimization for Low and High-dimensional Feature Selection. *Expert Syst. Appl.* **2021**, *176*, 114778. [CrossRef]

92. Al-Wajih, R.; Abdulkadir, S.J.; Aziz, N.; Al-Tashi, Q.; Talpur, N. Hybrid Binary Grey Wolf With Harris Hawks Optimizer for Feature Selection. *IEEE Access* **2021**, *9*, 31662–31677. [CrossRef]

93. Shunmugapriya, P.; Kanmani, S. A Hybrid Algorithm using Ant and Bee Colony Optimization for Feature Selection and Classification (AC-ABC Hybrid). *Swarm Evol. Comput.* **2017**, *36*, 27–36. [CrossRef]
94. Zorarpaci, E.; Özel, S.A. A Hybrid Approach of Differential Evolution and Artificial Bee Colony for Feature Selection. *Expert Syst. Appl.* **2016**, *62*, 91–103. [CrossRef]
95. Jona, J.B.; Nagaveni, N. Ant-cuckoo Colony Optimization for Feature Selection in Digital Mammogram. *Pak. J. Biol. Sci. PJBS* **2014**, *17*, 266–271. [CrossRef] [PubMed]
96. Abdmouleh, Z.; Gastli, A.; Ben-Brahim, L.; Haouari, M.; Al-Emadi, N.A. Review of Optimization Techniques applied for the Integration of Distributed Generation from Renewable Energy Sources. *Renew. Energy* **2017**, *113*, 266–280. [CrossRef]