Review on Hybrid Swarm Algorithms for Feature Selection

Abubakr S. Issa[1], Yossra H. Ali[2*], Tarik A. Rashid[1]
[1]Computer Science Department, University of Technology, Baghdad , Iraq
[2]Computer Science & Engineering Department, University of kurdidtan Hewler, KRG, Iraq

Abstract
   Feature selection represents one of the critical processes in machine learning (ML). The fundamental aim of the problem of feature selection is to maintain performance accuracy while reducing the dimension of feature selection. Different approaches were created for classifying the datasets. In a range of optimization problems, swarming techniques produced better outcomes. At the same time, hybrid algorithms have gotten a lot of attention recently when it comes to solving optimization problems. As a result, this study provides a thorough assessment of the literature on feature selection problems using hybrid swarm algorithms that have been developed over time (2018-2021). Lastly, when compared with current feature selection procedures, the majority of hybrid algorithms enhance classification accuracy.

Keywords: Feature selection, Hybrid Swarm intelligence, literature review, classification.

مراجعة عن الخوارزميات الهجينة المستعملة في عملية اختيار المميزات

ابوبكر صائب عيسى[1], يسرى حسين علي[1*], طارق احمد رشيد[2]
[1]قسم علوم الحاسوب , الجامعة التكنولوجية , بغداد , العراق
[2]قسم هندسة و علوم الحاسوب , جامعة كردستان , اربيل , العراق

الخلاصة
   اختيار المميزات تمثل واحدة من اكثر العمليات الحاسمة في تعلم الماكنة . الهدف الأساسي لمشكلة اختيار المميزات هو الحفاظ على دقة الأداء ، في حين تقليل ابعاد المميزات المختارة . تم إنشاء طرق مختلفة لتصنيف مجموعات البيانات. في نطاق مشاكل التحسين ، أنتجت تقنيات السرب نتائج أفضل. في الوقت نفسه ، حصلت الخوارزميات الهجينة على الكثير من الاهتمام مؤخرًا عندما يتعلق الأمر بحل مشاكل التحسين. ونتيجة لذلك ، توفر هذه الدراسة مراجعة ادبية حول مشاكل اختيار الميزات باستعمال خوارزميات سرب هجينة التي تم تطويرها بين الاعوام (2018−2021). أخيرًا ، بالمقارنة مع إجراءات اختيار المميزات الحالية ، فإن غالبية الخوارزميات الهجينة تعزز دقة التصنيف.

1. Introduction

_____

*Email: Yossra.H.Ali@uotechnology.edu.iq

Lately, data has increased in numerous industries, including blogs, social media, scientific research, business, and medical applications. This dataset has many features that might be extracted, yet not consistently; each feature is essential for extracting useful data from datasets. Some features could be redundant or unnecessary, lowering the model's performance. As a result, reducing such features is critical for improving model accuracy while minimizing the computational cost [1].

The paper is structured as follows: Section 2 goes over feature extraction and feature selection. Section 3 describes metaheuristic algorithms. Section 4 illustrates swarm intelligence. Hybrid swarm algorithms will be discussed in Section 5. Section 6 explains the classification. Section 7 shows a literature review. Section 8 will discuss comparative studies. Finally, there is a conclusion in Section 9.

2. Feature extraction and feature selection
Feature extraction uses two types of transformations, non-linear and linear, for transforming the original space of features into a new one with dimensionality reduction. In contrast, feature selection decreases the number of original features through the selection of a subset of the features that consist of the most relevant and important information for classification [2]. Yet, feature selection has been considered a difficult task as a result of the large search space. Assuming there are x features, the total number of feature subsets is 2x, and that value dramatically increases compared with the number of original features [3]. Figure 1 depicts the four main processes in the feature selection algorithm. The general procedure for feature selection starts by generating a candidate feature subset for evaluation. Each candidate subset is evaluated by using an evaluation criterion to measure the quality of the selected features. The process of subset generation and evaluation is repeated until a predefined stopping criterion is satisfied. The feature selection process ends by submitting the selected subset of features to a validation procedure. The stopping criteria can be the number of iterations, the number of features, or the best classification accuracy, depending on the selected subset.
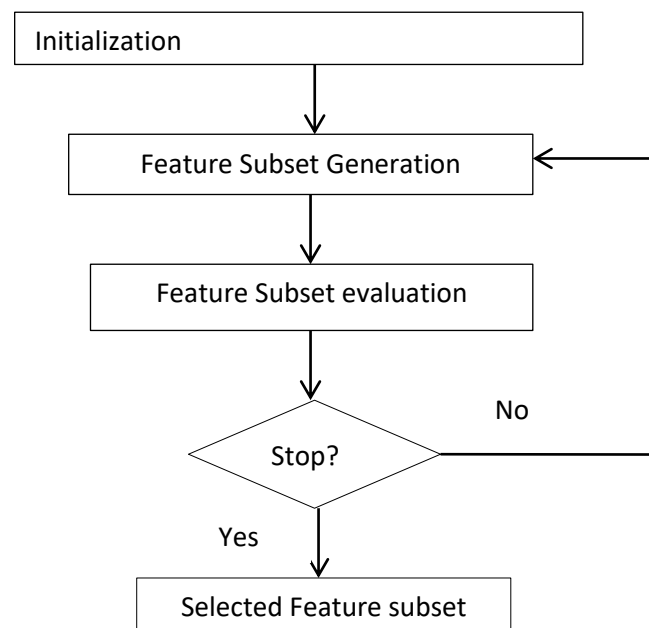
Figure 1: An overall feature selection process [4]

Feature selection could be separated into three primary methods based on evaluation criteria: embedded, wrapper, and filter approaches. Wrapper approaches use a classification algorithm to locate subsets of features, whereas embedded approaches utilize a classification algorithm but select subsets throughout the classifier's training process [3]. On the other hand, filter techniques do not require a classification algorithm to create subsets. Usually, the evaluation depends on the dataset's inherent properties. Table 1 depicts the categories of the feature selection algorithms.

Table 1: Categories of the feature selection algorithms [5]

Category	Category description
Filters	Select features using independent approaches. A score, or an evaluation criterion, is used to choose the set of features by assessing the degree of relevance regarding each of the characteristics.
Wrappers	The wrappers represent feature selection approaches that use a learning algorithm for evaluating a subset of characteristics' classification performance. The evaluation is carried out with the help of a classifier, which evaluates the relevance of a subset of features.
Embedded	Wrapper and filter approaches are combined in embedded methods. Because the filter approaches were faster but not particularly effective, and the wrapper approaches were more effective yet computationally expensive, particularly with large datasets, a solution that included the benefits of both approaches has been required.
Hybrid	Multiple conjunctive primary feature selection methods are applied sequentially in this feature selection approach.

3. Metaheuristic Algorithms

Metaheuristic algorithms represent optimization-based approaches for finding the optimum (or near-optimum) solution to optimization problems. Those algorithms have the benefits of simplicity, flexibility, and the capacity to avoid the local optimum problem [6].

Exploitation and exploration are the two major steps of such algorithms [7]. The algorithms extensively investigate the promising search space in the exploration phase, and the exploitation phase does a local search of the promising field(s) discovered in the exploration phase [1]. Metaheuristic algorithms have two fundamental categories:
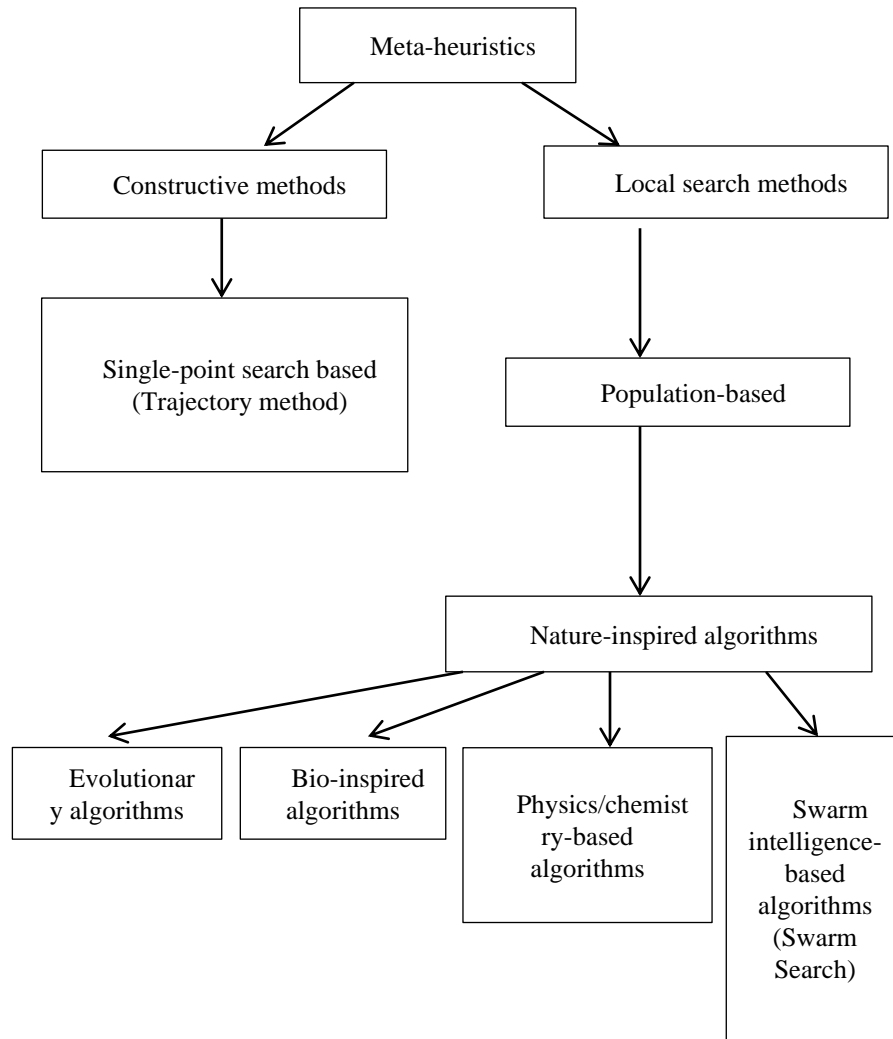
1- Single solution based metaheuristic algorithms
This algorithm begins its optimization process with a single solution, which is subsequently updated via iterations of optimization. However, the disadvantages of this algorithm are that it might enter local optima and doesn't search the entire search space.
2- Population (multiple) solution based metaheuristic algorithms:
This type of algorithm begins its optimization process with a set of solutions referred to as the "population," which are updated via optimization iterations to assist each other in getting out of the local optima problem and exploring the entire search space, making this sort of algorithm recommended for dealing with real-world problems [1].

As illustrated in Figure 2, metaheuristic algorithms might be classified into four groups depending on their behavior: swarm intelligence-based, evolution-based, physics-based, and human-related algorithms [8].

Figure 2: Classification of metaheuristic algorithms [9]

Evolution-based algorithms: This algorithm mimics natural biological evolution by starting with a population of randomly generated solutions. After that, the best solutions are summed to generate new ones. Crossover, mutation, and selection of the best solution are used to create new solutions. The genetic algorithm (GA), which depends on Darwin's evolution approach, is the most widely utilized in this category [10].

Swarm intelligence-based algorithms (SI): Those algorithms have been based on the social activities of animals, birds, creatures, etc. Particle Swarm Optimization (PSO), created by Eberhart and Kennedy, is the most widely utilized approach [11].

Physics-based algorithms: These have been inspired by physics rules in the universe. Examples of this type are simulated annealing [12], harmony search [13], etc.

Human behavior-related algorithms: Human behavior is the inspiration for such approaches. The researchers are motivated to create these algorithms by researching the actions of each individual that influence their performance. The Teaching learning-based optimization

algorithm (TLBO) [14], League Championship algorithm [15], and others are among the most important.

4. Swarm intelligence

Each SI algorithm should follow a set of basic steps. As a result, the SI framework (Figure 3) is as follows:
1. Initialization of the population
2. Definition of the stop condition
3. Evaluation of the fitness function
4. Updating and moving agents
5. Returning global best solution
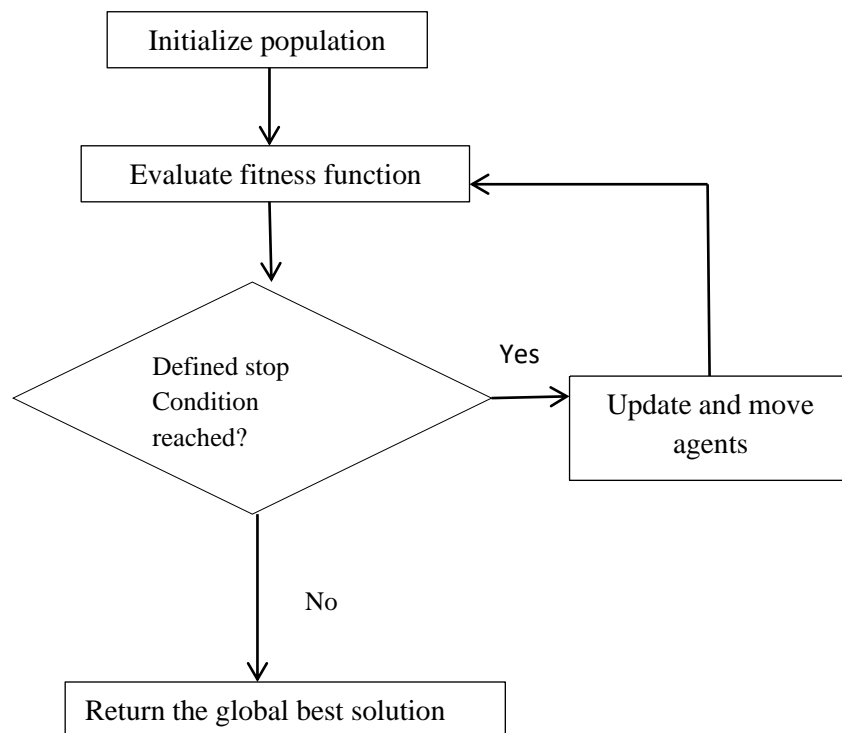
Figure 3: Swarm intelligence framework [16]

Swam search is one of the most flexible metaheuristic searches utilized in feature selection, in which one of the swarm algorithms is combined with one of the classification techniques to formulate the feature selection process. For example, let the original set of features of size k, A={i1,i2,i3,ik}, The swarm algorithm (Swam Search) constructs a classifier with a classification error rate (e) and provides an optimal feature subset S such that e(A)≥e(S), where S is one subset of all the possible subsets that exist in the hyperspace of A.

5. Hybrid Swarm algorithms

Hybrid swarm algorithms are created by combining the best operators from other swarm algorithms or combining a swarm algorithm with another approach for producing the best subset of characteristics. When it comes to solving optimization problems, hybrid algorithms have recently gotten a lot of attention. Various hybrid swarm algorithms were created for the purpose of obtaining the most relevant and optimal feature subset from the original dataset, especially for the feature selection problem. There are many possibilities for improving an

algorithm that finds the optimal solution, even by modifying the algorithm to enhance the quality of solutions or by doing hybridization with other algorithms to overcome problems related to the concerned algorithm. The enhanced algorithms help to avoid local optima without the risk of early convergence, allowing for more effectively and efficiently exploring and exploiting the search space. Also, the improved algorithms accomplish the optimal or near-optimal solution and strike a better balance between the exploitation and exploration qualities of the algorithm. A few algorithms were presented that combined the best aspects of several algorithms to produce a new one.

6. Classification

The process of developing a model that specifies data concepts or classes in order to determine the class of unknown-class-label objects is referred to as classification [17]. There are two basic steps in classification:
1-Training: In order for the class label to stand out, the classification creates a model using the training data in this step.
2-Testing: This step involves putting the model to the test by labeling data objects in the test dataset with class labels [18].
Classifying a certain email into "non-spam" or "spam" or allocating a diagnosis to a patient depending on observable patient features (blood pressure, age, the absence or presence of specific symptoms, etc.) are two classification examples. The classification procedure is a supervised learning example. For instance, a training set of correctly identified observations is available [19]. At the same time, clustering is an unsupervised procedure that involves categorizing data using some inherent similarity or distance measure. Regression-based classifiers, generative classifiers, and discriminative classifiers are the three categories of classifiers [19]. KNN is one of the most popular classifiers in the past few years; however, there are various others such as decision trees, SVM, and Naive Bayes.

7. Literature Review

This section will look at a few works on classification systems that use a hybrid swarm algorithm to choose the best features among them (2018–2021):

In 2018, Rajamohana and Umamaheswari [20] presented a study in which a hybrid method of enhanced binary particle swarm optimization and shuffled frog leaping algorithm is suggested for reducing the high dimensionality of the feature set and selecting the best feature set. After that, K Nearest Neighbor (KNN), Support Vector Machine (SVM), and Naïve Bayesian (NB) classifiers have been utilized to do spam review. Ott et al. created the dataset, including 1600 reviews of the 20 most significant Chicago hotels, with 20% utilized for testing and 80% utilized for training. The findings were compared with the binary PSO and SFLA algorithms, and all three classification approaches (NB, KNN, and SVM) were shown to be more accurate.

In 2019, Al-Tashi et al. [21] used the strength of the HGWO hybrid algorithm to present a new binary version of it referred to as the BGWOPSO. The study utilized 18 benchmark datasets from the UC Irvine Machine Learning Repository for classifying the system, using a KNN classifier compared to four state-of-the-art techniques (BPSO, bGWO2, WOASAT-2, and BGA).

In 2019, T. Keerthika and K. Premalatha [22] presented a new hybrid optimization algorithm called HFSBEE, which is the hybrid algorithm of two swarm algorithms (the fish algorithm and the artificial bee colony). The new algorithm addresses the shortcomings of the

fish algorithm, which has a slower convergence speed and takes a long time to specify an optimal solution, as well as the ABC algorithm's disadvantage of increasing computational cost. The classification is achieved using the multi-kernel support vector machine approach, and the proposed classification approach is tested on three datasets from the UCI machine learning repository: Hungarian, Swiss, and Cleveland. The newly suggested approach outperforms the traditional fish optimization and ABC algorithms.

In 2020, Sundaramurthy S. and Jayavel [23] combine PSO with Grey Wolf Optimization to create an efficient Rheumatoid Arthritis (RA) disease prediction system. The method has two main phases. The researchers used the C4.5 classifier to predict the RA in the first stage, which relied on PSO to generate the initial population, and the second step relied on GWO for choosing the optimal subset of features. Also, the dataset utilized in this work was acquired from interested patients of Shakthi Rheumatology Unit's outpatient unit in Coimbatore. One thousand patients were evaluated for model prediction, with 375 male and 625 female patients.

In 2020, Sagban et al. [24] utilized the Binary BAT algorithm to optimize feature selection by benefiting from frequency tuning and automatic zooming of the algorithm. They then attempted to do classification using the Ant-Miner classifier with five folds, each using 20% of the test data. The dataset utilized was a cervical cancer dataset that was downloaded from the UCI repository and involved 858 patients.

In 2021, Thawkar S. [25] suggested a hybrid approach (TLBO-SSA) that is used to analyze mammograms of breast cancer patients in this research. The Salp swarm algorithm improves the efficiency and convergence of the basic learning-based optimization technique. This change merges the SSA's update approach with the TLBO's primary structure. The classifier employed is the adaptive neuro-fuzzy inference system, which has a precision of 98.46%. The Breast Cancer Wisconsin (WBC) Diagnostic Dataset was utilized.

In 2021, Abdel-Basset M et al. [26] presented a hybrid approach (HHOBSA) that combines the Harris Hawks algorithm with a simulated annealing algorithm for classifying various datasets, such as computer, biology, life, financial, statistical, and physical. While simulated annealing (SA) improves the HHOBSA algorithm's performance and helps avoid local optima, two bitwise operations (AND and OR) can randomly transfer the most useful properties from the optimal solution to the other solutions in the population to improve their quality. The KNN classifier is used to rate the solutions' quality. The highest classification obtained was 85%.

In 2021, Fahad et al. [27] suggested an approach (ACO-SU) that combined ACO and symmetric uncertainty. The suggested method assesses the utility of incoming features and discards those that are not required. The algorithm updates the acquired feature set when a new feature is identified. The new technique was tested using 14 medical image datasets from the UCI repository. Three classifiers are employed to evaluate the new approach's quality (JRip, J48, and the decision table). The average accuracy of classification was 72.69%.

In 2021, L. Meenachi and S. Ramakrishnan [28] hybridized the ant colony optimization algorithm with the local search algorithm Tabu search and a fuzzy rough set to predict cancer by selecting the best features from microarray gene expression data. The ant colony algorithm hybridized with the fuzzy rough set to find global optimal features. Following that union of local, Tabu search is hybridized with a fuzzy rough set to find local optimal features. The fuzzy rough nearest neighbor classifier compares the results with the typical ANT algorithm and Tabu search and achieves better results with all types of datasets and classifiers used. The medical

datasets used are small-round blue-cell tumor (SRBCT), diffuse large B-cell lymphoma (DLBCL), breast and leukemia cancer datasets, and non-medical dataset swarm behavior from the UCI ML repository, which was used to demonstrate the generalization capabilities of the proposed algorithms.

In 2021, Dharmalingam and Kumar [29] hybridized multi-objective with the Tabu search algorithm to address the shortcomings of conventional single-objective PSO by generating a set of optimum solutions, which are then utilized to select the best features extracted from lung chest tomography (CT) images. In the classification level, the KNN classification approach with normal distribution and class probability is used. Lung CT images from Stanley Medical Hospital and other scan locations were included in the dataset.

In 2021, Adamu et al. [30] created an alternative wrapper-based feature selection approach by combining CSA and PSO algorithms. The former was altered by using a chaotic map to handle the diversity issue, and then enhanced CSA was combined with PSO to create the Enhanced Chaotic Crow Search Particle Swarm Optimization Algorithm (ECCSPSOA). To solve the local optimization problem, the KNN classifier and the opposition-based learning (OBL) local search method are used. The study made use of 15 well-known benchmark datasets from UCI data.

In 2021, Sathiyabhama et al. [31] hybridized the GWO algorithm with RST for selecting the best feature subsets for the highest classification accuracy using a rough set of positive regions and dependency functions (i.e., fitness functions). The authors used a J48, decision table, Naive Bayes, and IBK classifier, but their best classification performance was with a decision table using five datasets: GLCM 0 with 96.3 % classification accuracy, ISF with 96.4 % classification accuracy, GLCM 90 with 91.9 % classification accuracy, GLCM 45 with 93.3 % classification accuracy, and GLCM 135 with 94.1 % classification accuracy.

In 2021, Al-Wajih et al. [32] suggested a hybrid approach (HBGWOHHO) that combines the Harris Hawks and Binary Grey Wolf algorithms, and 18 standard UCI benchmark datasets have been utilized to validate the suggested approach. The quality of the selected features is assessed using a wrapper-based KNN. The Binary Grey Wolf Optimizer (BGWO), Binary Harris Hawks Optimizer (BHHO), Binary Particle Swarm Optimization (BPSO), Binary Genetic Algorithm (BGA), and Binary Hybrid BWOPSO were used to compare the performance of the suggested hybrid technique. Using the suggested approach, the average accuracy was 92%.

In 2021, for solving the problem of feature selection, Kitonyi and Segera [33] suggested an approach (HGDGWO) that combined the well-known meta-heuristic population-based optimizer, the Grey Wolf algorithm, with an iterative optimization gradient descent algorithm. In the classification stage, 6 medical data sets from the UCI machine learning repository were employed. Also, the accuracy of the suggested technique was tested five times using SVM and KNN classifiers, with certain datasets reaching 100% accuracy. Compared with several existing approaches (BGWOPSO, BGWO), the suggested approach could not attain maximum accuracy across all datasets.

In 2021, Chen et al. [34] suggested a BP-PSO approach that combines an adaptive PSO with a backpropagation NN to perform feature selection. To evaluate the proposed method's quality approach, a set of classifiers and ranking systems was applied (SVM-RFE, DTree, Wilcoxon test, RF, t-test). The datasets utilized for classifying cancer samples were taken from

the TCGA database. The results indicate that BP-PSO has an average accuracy of 8.65% greater compared to the suboptimal NDFs model in various datasets and that its performance is 2.31–18.62% higher than the benchmark approach in all data sets.

Table 2 shows a comparison of using the hybrid method with various swarm algorithms.

Table 2: Comparison between using the hybrid method with different swarm algorithms

Ref	Proposed method	Algorithms or methods used	Dataset used	Classifier used	Used for	Accuracy
[20], 2018	iBPSO_SFLA	improved binary particle swarm optimization and shuffled frog leaping	Ott. et al. dataset	Naive Bayes and KNN and SVM	Detecting spammy reviews	NB classifier with 94.97 % KNN classifier with 92.12% SVM classifier with 91.25 %
[21], 2019	BGWOPSO	Grey Wolf Optimization and Particle Swarm Optimization	UC Irvine Machine Learning Repository	KNN Classifier	Classification of multiple datasets	97 %
[22], 2019	HFSBEE	Fish algorithm and Artificial Bee Colony	Cleveland dataset, Hungarian dataset and Switzerland dataset	MULTI-KERNEL SUPPORT VECTOR MACHINE	Heart Disease Prediction	Cleveland with 97.68% Hungarian with 95.91 % Switzerland with 97.56 %
[23], 2020	HGWO-C4.5	Grey Wolf Optimization and Particle Swarm Optimization	Dataset of Shakthi Rheumatology Unit	C4.5 Classifier	Rheumatoid Arthritis (RA) disease prediction	86.36%
[24], 2020	Hybrid bat-ant colony	Bat algorithm and ant colony algorithm	cervical cancer dataset	Ant-Miner classifier	Diagnosis of cervical cancer	Hinselmann test with 95.93% Schiller's test with 90.91 % Cytology test with 94.88 % Biopsy test with 95.8

Ref	Method	Technique	Dataset	Classifier	Application	Accuracy
[25], 2020	TLBO-SSA	Teaching-learning based optimization and Salp swarm algorithm	Breast Cancer Wisconsin (WBC) Diagnostic dataset	Adaptive neuro-fuzzy inference system	classification in digital mammography	98.46%
[26], 2020	HHOBSA	Harris Hawks optimization algorithm with simulated annealing	diabetic, EEG-eye-state, fri_c0_1000_10, fri_c1_1000_10, kc1, and page blocks and Lichman UCI repository dataset.	KNN Classifier	Classify multiple types of datasets including (biology, computer, financial, life, physical, and statistical)	89%
[27], 2020	ACO-SU	Ant colony optimization and symmetric uncertainty	UCI medical image datasets	J48, JRip, decision table	Medical Image Diagnosis	72.69%.
[28], 2021	ACTFRO	Ant Colony Optimization and Tabu search with Fuzzy Rough set	The medical dataset used are namely, small-round-blue-cell tumor (SRBCT), Dif-fuse Large B-Cell Lymphoma (DLBCL), Breast and Leukemia cancer dataset. The non-medical dataset swarm behavior from UCI machine learning repository.	fuzzy rough nearest neighbor classifier	Classification of multiple datasets	88%
[29], 2021	MOPSO + TS	multi-objective PSO and Tabu search	Stanley Medical hospital and other scan centers lung CT images.	KNN Classifier	classification of lung disorders	90.588
[30], 2021	ECCSPSOA	Enhanced Chaotic Crow Search Particle Swarm Optimization	UCI repository	KNN Classifier	Classification of multiple datasets	89.76 %

[31], 2021	GWORS	Grey wolf optimizer with rough set theory	ISF, GLCM 0, GLCM 45, GLCM 90, and GLCM 135 Dataset	Decision table classifier	mammogram image analysis	96.4 %
[32], 2021	HBGWOHHO	Binary Grey Wolf and Harris Hawks algorithm	UCI benchmark datasets	KNN Classifier	Classification of multiple datasets	92%
[33], 2021	HGDGWO	Gradient Descent algorithm and Grey Wolf Optimizer	UCI medical datasets	KNN and SVM	Classify multiple types of medical datasets	94%
[34], 2021	BP-PSO	Adaptive Particle Swarm Algorithm and Back propagation neural network	TCGA database	DTree,SVM-RFE, RF.t-test, Wilcoxon test	classify cancer samples	86.5%

8. Discussion

The literature review that was introduced in the previous section showed that the hybrid swarm algorithms used for feature selection can enhance classification accuracy [23]. [20–22, 24–25], [29], and [31–32] achieved accuracy between [90%–99%] while the other research achieved less than 90% classification accuracy. Table 3 represents a comparative study of classification accuracy between various feature selection techniques and the hybrid method of different swarm algorithms for feature selection.

Table 3: Comparison of classification accuracy between existing and hybrid methods of different swarm algorithms for feature selection

Ref	Classifier	Technique	Accuracy	Hybrid technique	Accuracy
			Existing Feature selection techniques	**Hybrid swarm algorithms feature selection techniques**	
[20]	NB	BPSO	82.5%	iBPSO_SFLA	94.97%
	KNN		85.05%		
	SVM		88.38%		92.12%
	NB	SFLA	85.5%		
	KNN		87.07%		91.25%
	SVM		89.84%		
[21]	KNN	bGWO2	83 %	BGWOPSO	97%
		BPSO	82 %		
		BGA	83%		
		WOASAT-2	92%		
[22]	SVM	FSO	Cleveland 96.03 % Hungarian 93.19% Switzerland 94%	HFSBEE	Cleveland with 97.68% Hungarian with 95.91 %
		ABC	Cleveland 93.72 % Hungarian 92.17% Switzerland 93%		Switzerland with 97.56 %

Ref	Classifier	Algorithm	Accuracy	Hybrid	Hybrid Accuracy
[23]	C4.5	C4.5	72.29 %	HGWO-C4.5	86.36%
		PSO-C4.5	81.25 %		
		GWO-C4.5	80.42 %		
[25]	ANFIS	TLBO	95.22 %	TLBO-SSA	98.46%
[26]	KNN	HHO	84%	HHOBSA	89%
		HHOB	87%		
[28]	FRNNC	TS	78%	ACTFRO	88%
		DE	79%		
		PSO	80%		
		FS-JMIE	78%		
		DEGR	81%		
		GLO	84%		
[29]	KNN	MOPSO	78.82%	MOPSO + TS	90.58%
		Firefly	61.18%		
		Bee colony	72.94%		
[30]	KNN	PSO	87.75%	ECCSPSOA	89.76 %
		BPSO	89.32%		
		CSA	87.23%		
		CCSA	87.5%		
[31]	SVM	GSO	94%	GWORS	96.4%
		Minimum redundancy maximum relevance (mRMR)	87%		
	Decision Table				
[32]	KNN	BGWO	88%	HBGWOHHO	92%
		BHHO	93%		
		BPSO	82%		
		BGA	83%		

9. Conclusion

This research provided a thorough survey regarding hybrid SI-based feature selection algorithms. The major purpose of this study is to examine hybrid algorithms for feature selection tasks and the classification accuracy of hybrid algorithms in comparison with original techniques and state-of-the-art algorithms. From this survey, a collection of facts was concluded that can help in choosing the best hybrid approaches for the problem, as mentioned below:

1- Since feature selection is a binary problem, most algorithms employ binary versions when creating hybrid techniques.

2- The hybrid approaches were utilized to solve the problems of controlling algorithmic exploitation and exploration.

3- Most hybrid algorithms are employed to improve the fitness value, which improves classification accuracy and feature selection.

4- Due to its strength and great capacity for hybridization, KNN is the most widely utilized approach for classification, and PSO is the most widely utilized algorithm for hybridization.

10. References

[1] P. Agrawal, H. F. Abutarboush, T. Ganesh, and A. W. Mohamed, "Metaheuristic algorithms on feature selection: A survey of one decade of research (2009-2019)," IEEE Access, vol. 9, pp. 26766–26791, 2021, doi: 10.1109/ACCESS.2021.3056407.

[2]   M. Rostami, K. Berahmand, E. Nasiri, and S. Forouzande, "Review of swarm intelligence-based feature selection methods," Engineering Applications of Artificial Intelligence, vol. 100. 2021. doi: 10.1016/j.engappai.2021.104210.

[3]   B. H. Nguyen, B. Xue, and M. Zhang, "A survey on swarm intelligence approaches to feature selection in data mining," Swarm and Evolutionary Computation, vol. 54. 2020. doi: 10.1016/j.swevo.2020.100663.

[4]   M. Mehri, R. Chaieb, K. Kalti, P. Héroux, R. Mullot, and N. E. ben Amara, "A comparative study of two state-of-the-art feature selection algorithms for texture-based pixel-labeling task of ancient documents," Journal of Imaging, vol. 4, no. 8. 2018. doi: 10.3390/jimaging4080097.

[5]   K. Tadist, F. Mrabti, N. S. Nikolov, A. Zahi, and S. Najah, "SDPSO: Spark Distributed PSO-based approach for feature selection and cancer disease prognosis," Journal of Big Data, vol. 8, no. 1. 2021. doi: 10.1186/s40537-021-00409-x.

[6]   S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," Advances in Engineering Software, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[7]   Omid Tarkhaneh, Thanh Thi Nguyen, and Samaneh Mazaheri, "A Novel Wrapper-Based Feature Subset Selection Method using Modified Binary Differential Evolution Algorithm," Information Sciences, Volume 565, pp. 278-305, July 2021. doi: 10.1016/j.ins.2021.02.061.

[8]   W. Mohamed, ALI WAGDY MOHAMED, H. F. ABUTARBOUSH, A. A. HADI, and ALI KHATER MOHAMED, "Gaining-Sharing Knowledge Based Algorithm With Adaptive Parameters for Engineering Optimization," in IEEE Access, vol. 9, pp. 65934-65946, 2021. doi: 10.1109/ACCESS.2021.3076091.

[9]   Kumar and A. Jaiswal, "Swarm intelligence based optimal feature selection for enhanced predictive sentiment accuracy on twitter," Multimedia Tools and Applications, vol. 78, no. 20. pp. 29529–29553, 2019. doi: 10.1007/s11042-019-7278-0.

[10] Sourabh Katoch, Sumit Singh Chauhan, and Vijay Kumar, "A review on genetic algorithm: past, present, and future," Multimedia Tools and Applications, Vol. 80, pp. 8091–8126, 2021. doi: doi.org/10.1007/s11042-020-10139-6.

[11] Arabinda Pradhan, Sukant Kishoro Bisoy, and Amardeep Das, "A survey on PSO based meta-heuristic scheduling mechanism in cloud computing environment," Journal of King Saud University - Computer and Information Sciences, Vol. 34, No. 1, 2021. doi: 10.1016/j.jksuci.2021.01.003.

[12] Peter J. M. Laarhoven, Emile H. L. Aarts , "Simulated annealing: Theory and applications," A chapter in book :Mathematics and Computers in Simulation, vol. 30, no. 1–2. p. 200, 1988. doi: 10.1016/0378-4754(88)90140-1.

[13] D. Manjarres et al., "A survey on applications of the harmony search algorithm," Engineering Applications of Artificial Intelligence, vol. 26, no. 8. pp. 1818–1831, 2013. doi: 10.1016/j.engappai.2013.05.008.

[14] R. v. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching-Learning-Based Optimization: An optimization method for continuous non-linear large scale problems," Information Sciences, vol. 183, no. 1. pp. 1–15, 2012. doi: 10.1016/j.ins.2011.08.006.

[15] S. M. Abdulhamid, M. S. A. Latiff, S. H. H. Madni, and O. Oluwafemi, "A survey of league championship algorithm: Prospects and challenges," Indian Journal of Science and Technology, vol. 8. pp. 101–110, 2015. doi: 10.17485/ijst/2015/v8iS3/60476.

[16] Lucija Brezočnik, J. Iztok Fister, and V. Podgorelec, "Swarm Intelligence Algorithms for Feature Selection : A Review," Applied sciences, vol. 8, no. 9, pp. 1521, Sep. 2018, doi: 10.3390/app8091521

[17] A K Varudharajulu and Y Ma, "Data Mining Algorithms for a Feature-Based Customer Review Process Model with Engineering Informatics Approach." Journal of Physics: Conference Series, 2019. doi: 10.1088/1742-6596/1267/1/012013.

[18] Alawi Alqushaibi, S. J. Abdulkadir, H. Md Rais, and Q. Al-Tashi, "An Optimized Recurrent Neural Network for Metocean Forecasting," 2020 International Conference on Computational Intelligence (ICCI), pp. 190-195, 2020, doi: 10.1109/ICCI51257.2020.9247681

[19] Q. Al-Tashi, H. Rais, and S. J. Abdulkadir, "Hybrid Swarm Intelligence Algorithms with Ensemble Machine Learning for Medical Diagnosis," 2018 4th International Conference on Computer and

Information Sciences: Revolutionizing Digital Landscape for Sustainable Smart Society, ICCOINS 2018 – Proceedings, 2018. doi: 10.1109/ICCOINS.2018.8510615.

[20] S. P. Rajamohana and K. Umamaheswari, "Hybrid approach of improved binary particle swarm optimization and shuffled frog leaping for feature selection," Computers and Electrical Engineering, vol. 67. pp. 497–508, 2018. doi: 10.1016/j.compeleceng.2018.02.015.

[21] Q. Al-Tashi, S. J. Abdul Kadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Binary Optimization Using Hybrid Grey Wolf Optimization for Feature Selection," IEEE Access, vol. 7, no. c, pp. 39496–39508, 2019, doi: 10.1109/ACCESS.2019.2906757.

[22] T. Keerthika and K. Premalatha, "A hybrid fish – Bee optimization algorithm for heart disease prediction using multiple kernel SVM classifier," International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 2. pp. 729–737, 2019. doi: 10.35940/ijitee.I1152.0789S219.

[23] S. Sundaramurthy and P. Jayavel, "A hybrid Grey Wolf Optimization and Particle Swarm Optimization with C4.5 approach for prediction of Rheumatoid Arthritis," Applied Soft Computing Journal, vol. 94. 2020. doi: 10.1016/j.asoc.2020.106500.

[24] R. Sagban, H. A. Marhoon, and R. Alubady, "Hybrid bat-ant colony optimization algorithm for rule-based feature selection in health care," International Journal of Electrical and Computer Engineering, vol. 10, no. 6. pp. 6655–6663, 2020. doi: 10.11591/ijece.v10i6.pp6655-6663.

[25] S. Thawkar, "A hybrid model using teaching–learning-based optimization and Salp swarm algorithm for feature selection and classification in digital mammography," Journal of Ambient Intelligence and Humanized Computing, vol. 12, no. 9. pp. 8793–8808, 2021. doi: 10.1007/s12652-020-02662-z.

[26] M. Abdel-Basset, W. Ding, and D. El-Shahat, "A hybrid Harris Hawks optimization algorithm with simulated annealing for feature selection," Artificial Intelligence Review, vol. 54, no. 1. pp. 593–637, 2021. doi: 10.1007/s10462-020-09860-3.

[27] L. G. Fahad, S. F. Tahir, W. Shahzad, M. Hassan, H. Alquhayz, and R. Hassan, "Ant colony optimization-based streaming feature selection: An application to the medical image diagnosis," Scientific Programming, vol. 2020. 2020. doi: 10.1155/2020/1064934.

[28] L. Meenachi and S. Ramakrishnan, "Metaheuristic Search Based Feature Selection Methods for Classification of Cancer," Pattern Recognition, Vol. 119, 108079, 2021. https://doi.org/10.1016/j.patcog.2021.108079

[29] V. Dharmalingam and D. Kumar, "Hybrid feature selection model for classification of lung disorders," Journal of Ambient Intelligence and Humanized Computing, vol. 13, pp. 5609–5625, 2022. doi: doi.org/10.9790/3021-04140105.

[30] A. Adamu, M. Abdullahi, and S. B. Junaidu, "An hybrid particle swarm optimization with crow search algorithm for feature selection," Machine Learning with Applications, Vol. 6, 100108, 2021, https://doi.org/10.1016/j.mlwa.2021.100108. 2021

[31] B. Sathiyabhama, S. U. Kumar, J. Jayanthi, T. Sathiya, and A. K. Ilavarasi, "A novel feature selection framework based on grey wolf optimizer for mammogram image analysis," Neural Comput & Applic, vol. 33, pp. 14583–14602, 2021. https://doi.org/10.1007/s00521-021-06099-z 2021

[32] R. Al-Wajih, S. J. Abdulkadir, N. Aziz, Q. Al-Tashi, and N. Talpur, "Hybrid binary grey Wolf with Harris hawks optimizer for feature selection," IEEE Access, vol. 9. pp. 31662–31677, 2021. doi: 10.1109/ACCESS.2021.3060096.

[33] P. M. Kitonyi and D. R. Segera, "Hybrid Gradient Descent Grey Wolf Optimizer for Optimal Feature Selection," Biomed Res Int., Aug 2021. doi: doi.org/10.1155/2021/2555622.

[34] B. Chen, H. Chen, and M. Li, "Feature Selection Based on BP Neural Network and Adaptive Particle Swarm Algorithm," Mobile Information Systems, vol. 2021, 2021, doi: 10.1155/2021/6715564.