

An Analysis of Feature Selection Algorithm and their Optimization - A Scrutiny

S. Bharani Nayagi

Research Scholar, Department of Computer Science, SRM University, Kattankulathur, India.

E-mail: itbharansiva@gmail.com

T.S. Shiny Angel

Assistant Professor, Department of Software Engineering, SRM University, Kattankulathur, India.

E-mail: shinyant@srmist.edu.in

Received October 04, 2020; Accepted November 12, 2020

ISSN: 1735-188X

DOI: 10.14704/WEB/V18SI02/WEB18008

Abstract

The eradication of correlated evidence of the enormous volume of the directory is designated as data mining. Extracting discriminate knowledge associate with the approach is performed by a feature of knowledge. Knowledge rejuvenation is carried out as features and the process is delineated as a feature selection mechanism. Feature selection is a subset of features, acquired more information. Before data mining, Feature selection is essential to trim down the elevated dimensional information. Without feature selection pre-processing techniques, classification required interminable calculation duration which might lead to intricacy. The foremost intention of the analysis is to afford a summary of feature selection approaches adopted to evaluate the extreme extensive features.

Keywords

Feature Selection, Filter Method, Wrapper Method, Hybrid Method, Optimization Algorithm, BCO, CS, PSO, ACO and GA.

Introduction

In recent years, scrutiny of Feature Selection Methods has been optimized. It is implemented in an immense assortment of applications. Feature selection eradicates ominous features such as noisy, superfluous, and extraneous. It regulates the input dataset with lacking values into a single or feature subset [1]. Feature selection aspires is to arbitrate subset of a feature as limited as possible. It is the fundamental preprocessing step proceeding to bestow the task of data mining. The function of feature selection is to select the original feature subset, without any loss of appropriate intelligence. Hence, the

devaluation of feature subset demonstrates similar results as with primary elevated magnitude information. Supreme influences of Feature Selection are diminishing cache condition, averting over forcing, promoting knowledge, determination, fastening the execution of mining algorithms [2]. Feature selection is the refinement of evacuating superfluous or extraneous features from the standard source. By removing extraneous features, it decreases the classifier execution time, immense amplitude, and data handling expenditure. With feature selection methodology, it upgrades the accuracy and the intelligible [3].

The selection of a few feature selection algorithms and optimization algorithm approaches are discussed in this paper. The comparisons are done based on their benefits and losses which assist to figure out the current dispute and controversy in the research field. This paper systematized as follows, section 1 defines the essentials about feature selection and optimization techniques, section 2 analyzes the existing feature selection and optimization algorithm and section 3 states the work's conclusion.

Feature Selection Process

The algorithm of feature selection is branched into three divisions; filters, wrappers and hybrid selectors.

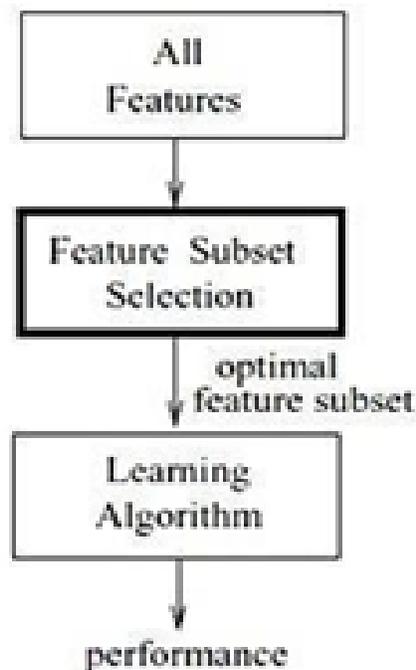


Figure 1.1 Filter Method

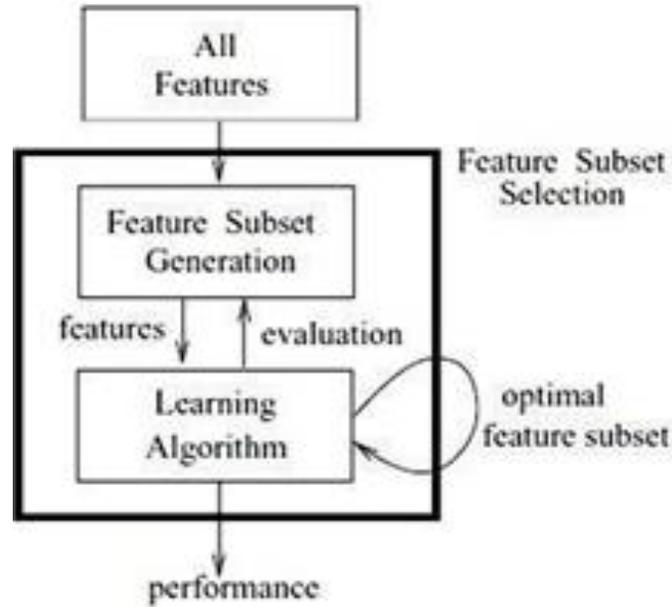


Figure 1.2 Wrapper Method

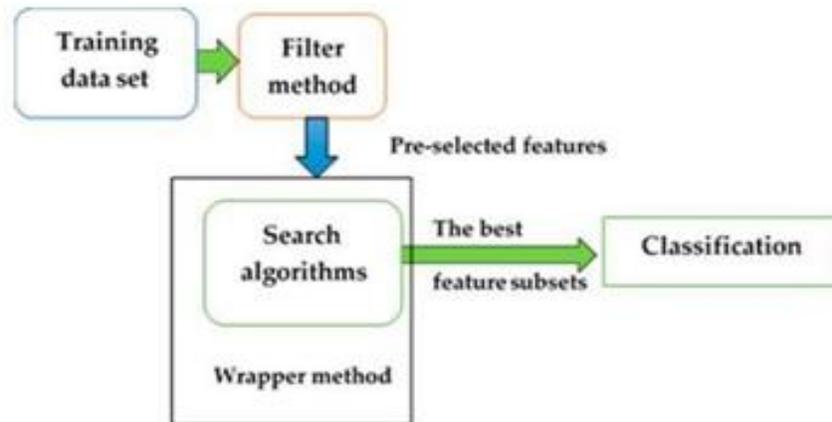


Figure 1.3 Hybrid Method

Filters estimate every feature nonpartisan from the classifier, class the estimated feature subset and select the excellent rank. The filter method produces universality and huge estimation adaptability. The Wrapper method tests the feature subset with the classifier performance. A Higher efficiency classifier subset is selected. Therefore, wrapper methods are determined by the corresponding classifier. It takes reasonable processing time and cache with higher predictable, which influence efficiency due to the classification algorithm. Hybrid Method associates both Filter and Wrapper to accomplish the influence of both independent measure and a mining algorithm. By this process, the filter method reduces the search space and the wrapper method achieves the perfect feature Subset [2].

As the realm dimensionality wider, the count of features subset N is enhanced. The selection of a perfect subset among many feature subsets is performed by the feature selection process as shown in fig 2. It is carried out by the following steps [4],

- Subset formation - Construct adjacent candidate subset through the origination operation.
- Subset assessment - Through analysis of evaluating subset.
- Stopping benchmark - Setup termination criteria.
- Validation - Verify the subset.

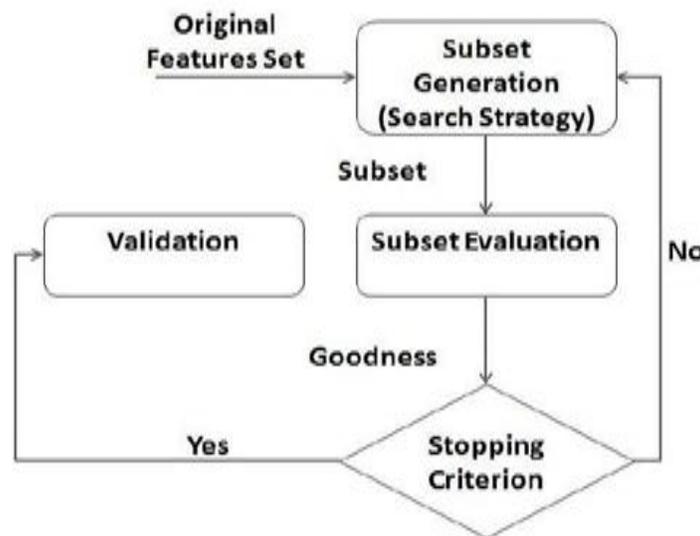


Figure 2 Feature Selection Process

Subset formation generates candidate feature subsets for assessment. It is hinged on a well-defined search strategy. The best candidate subset is selected by ranking feature subsets. The process is continued until the stopping state is satisfied. Feature rank determines the respective feature implication and neglecting their potential communication. Ranking methods depend on statistics, information theory, or on some functions of classifier outputs[4].

Optimization Algorithm

Periodically Optimization experienced plenty of controversy in engineering progress. An optimization algorithm is an action of periodic execution with various solutions to obtain an optimum solution. The optimization algorithm is classified into two methods such as deterministic and stochastic methods respectively. The deterministic algorithms desire enormous computational attempts, which may deteriorate as a complication, escalate. This method persuades metaheuristic stochastic optimization algorithms are most effective. The

characteristics of Metaheuristics algorithms are disputing self-sufficient and perform on the constant enrichment of either a multiple results or a distinct result and operate randomization and local search to provide results for optimization complications. Therefore, nature-inspired metaheuristic optimization algorithms give better performance on problem-solving issues in engineering progresses and influence in the advancement of artificial intelligence [5]. Metaheuristics algorithms implement Evolutionary feature selection algorithms.

The purpose of Evolutionary feature selection algorithms is to search space with the assessment of global and local functions. The optimal solution is obtained by search space. Hence, it arbitrates the probabilistic approximation on evaluation analysis efficiently. From the optimal period, a comparative solution can be attained. An algorithm involves in an evolutionary technique are as follows [6].

- Bee Colony Optimization.
- Cuckoo Search Optimization
- Particle Swarm Optimization.
- Ant Colony Optimization.
- Genetic Algorithm

Related Work

Feature Selection

Filter Method

Filter method implemented for the preprocessing process, which is self-sufficient in any machine learning algorithms. From the enormous analytical analysis, its score represents the feature selection. It selects variants of the unconcerned model. Prediction is done by varying correlation. The minimum appealing variables are restrained and other variations involve in classification or regression to classify and predict. It required a minimum computation period and overfitting robustness. Here some of the filter-based algorithms are reviewed:

- Fast Clustering-Based Feature Selection Algorithm(FAST).
- Fast Correlation Based Filtering(FCBF).
- Relief-F algorithm.
- Sequential Learning Framework for Feature Subset Selection (SLSS).
- Score-based Criteria Fusion(SCF)

The filter methods are independent of the classifier. Qinbao Song et al [8] proposed a Fast Clustering - Based Feature Selection Algorithm (FAST) filtering method for feature selection. It is constructed for an extremely spatial data set. At FAST, Initially, features are clustered by graph-theoretic clustering methods and the highest relevant clustered feature that is highly connected to target is selected. With the comparison of other filter method author found that FAST creates the smallest feature subset and advance classifier performance. He assured that the FAST efficiency by graph-theoretic clustering methods implemented on features using particular Minimum- Spanning Tree (MST) grouping method.

Paris Senlio et al [9] proposed Fast Correlation Based Filtering (FCBF) for feature selection. It prefers a different ordered feature that results in greater precious classifiers. The performance is verified with 8 datasets with 5 different subset sizes. The classification accuracy is tested by the 10-fold cross, which helps to validate 3-NNC classifiers. He accomplishes experiments with the MRMR algorithm and analyzes FCBF's performance.

Yuhang Wang et al [10] implement feature filtering algorithms of the Relief-F algorithm to calculate the efficiency of classification datasets. For classification, he implements the linear SVM and the KNN and evaluates Relief-F performance with other filtering method. Comparison is done with Information Gain, Gain Ratio, and χ^2 -statistic. He concluded Relief-F Algorithm is proportionate.

The number of features and the number of observations are together scalable with a Sequential Learning Framework for Feature Subset Selection (SLSS). Gregory Ditzler et al [11] proposed multi-arm bandit algorithms to find the variable subset and each feature authorized with a level of importance. Here, he proposed SLSS to measure high-dimensional dataset and estimate the same data with a minimum time period. It is an optimization self-reliant and performs the scoring function faster and effectively. He does the experiment with the real- time dataset. It reduces computation complexity. It is essential in a binary learning problem.

Feature selection plays a vital role to attempt a compress and distinguish feature subset. Wenjun Ke et al [11] proposed Score-based Criteria Fusion feature selection method (SCF) method for cancer prediction. SCF is a better and an excellent combination of features along with competing methods. His experiment was done with five datasets of GEM and three low-level datasets. SVM and KNN classifier does the selected dataset characteristics. SCF is efficient to find more number of particular features and evaluate

the importance between classes and features. SCF determines overabundance and eliminates, provide a high correlation.

Wrapper Method

The wrapper method is established by a greedy search algorithm. It is estimated all feasible subset of feature and selects the best subset. It is a computational expense. The Wrapper methods classified into three steps: Step forward feature selection, Step backward feature selection, and Exhaustive feature selection. Some of the wrapper method algorithms are

- SVM with kernel tricks.
- Cosine Similarity Measure Support Vector Machines (CSMSVM).
- BIG-F algorithm.
- Support Vector Machine (SVM).

Sebastian Maldonado et al [13] introduce a wrapper method for feature selection where they implemented Support Vector Machine with kernel tricks. It is established on sequential backward selection to amplitude the number of failures in the validation subset. It helps to accommodate a better data set. In each emphasis, dataset is divided randomly which helps to prevent overfitting. Here stopping criteria are defined explicitly hence elimination of features influences the classifier accomplishment. Therefore, 2-3 run of the algorithm has been permitted and the features are eliminated if they are omitted in more than one run. The Performances are verified and improve the effectiveness in need. It is robustness. It is expensive when a number of dataset increase. In such cases, performance is enhanced by applying the filter method before the wrapper method.

Gang Chen et al [14] implement the Cosine Similarity Measure Support Vector Machines (CSMSVM) wrapper method for feature selection. The CSMSVM implements the margin verses cosine distance ratio, through which features weight has been added to increase the margin verses cosine distance ratio in some cases, Feature selection leads to information loss, which leads to an error in the classifier. Hence, CSMSVM minimizes the classification error rate by attaining an optimal feature subset and higher the maximum a posterior (MAP). It is achieved from the novel methodology of Bayesian interpretation. It provides better computational effectiveness where it works with both feature selection and classification concurrently.

Breast cancer is a major problem in women where feature selection plays a vital role in establishing a cancer classifier. Qiqige Wuniri et al [15] proposed a wrapper method of the

BIG-F algorithm with a kernel-based Bayesian classifier for this purpose. The Dataset includes both discrete and continue where he deals discrete with a naive method and continue with kernel probability density estimation method respectively. This method is greater interpretable and concurrence. However, which he obtained an attainable, precious, and effective result.

A.H. Roslina et al [16] have classified with Support Vector Machines (SVM). She implements wrapper methods to eliminate irrelevant features in a feature dataset. It reduces dimensionality. Through this method, she maximized the accuracy and reduces the computation period and expenses.

Michael E. Farmer et al [17] proposed a wrapper feature selection method to provide better segmentation and classification. He had done segmentation with classification measurement. He considered a wrapper method as an image segmentation framework. This method is essential for real-time operation due to their robustness. Throughout his experiment, he achieved 91% of accuracy.

Hybrid Method

To enhance the feature selection performance, the hybrid method includes both filter and wrapper methods respectively. It accomplished by two steps. In the first step, the filter method orders the feature and in the second step select the best features by wrapper method. This method is implemented in a standard engineering model and a proven hybrid method provides better adaptability and precision to optimize the efficiency of feature selection. [39]. Here we reviewed a few hybrid algorithms:

The filter method prevents overabundance and wrapper method advanced features that are required to perform better classification performance. Yuan Jiang et al [18] implements the hybridization of the Mutual Information Feature Selection (MIFS) filters method and Modified Binary Cuckoo Search (MBCS) wrapper method for feature selection.

He implements a K-NN classifier for the fitness function. Throughout his experiment, he obtained greater classification efficiency and lower computation period as well a lower number of features respectively.

Jihong Liu et al [19] implemented a hybrid method for the feature selection process. As hybrid includes the filter and wrapper method, the selected feature subset is affiliated and applicable. Hence it provides superior classification achievement. It includes three stages, in the first stage, mutual information (MI) grade the features by implementing filtering a

filter model and select superior features. In the second stage, the wrapper methods implement Sheley the model to estimate feature subset classification. From the result, the HFS method provides superior classification achievement. MI eliminates the estimation expenditure adequately. Hybrid Feature Selection Method is extremely significant to classification performance.

B.M. Vidyavathi [20] proposed a hybrid method where the wrapper method is an unattainable and the filter approach does not prove better accuracy she implemented a peculiar method of Artificial Neural Network weight Analysis (ANNWA) for filter method and wrapper method introduced Genetic Algorithm to eliminate inconsequential and superfluous features.

It reduces the rate of dimensional features without affecting classification and prediction. It includes two phases. In the first phase, the filter method omitted inconsequent feature which is essential in bypassing the computational problem in a second phase wrapper method. She certified prospective mechanisms on a real-world data sets. She suggests this method is economical and increase prediction and classification accuracy.

Feature selection reduces the complexity of data mining and pattern recognition. Yuefeng Zheng et al [21] implement the Maximum Pearson Maximum Distance Improved Whale Optimization Algorithm (MPMDIWOA) of the hybrid feature selection algorithm. He proposed Maximum Pearson Maximum Distance (MPMD) filter algorithm to adopt pertinence and repetition weight respectively.

Modified Whale Optimization wrapper algorithm. He combines both as MPMDIWOA. Before combining, he proposed the Maximum Value without Change (MVWC) and the threshold to alter the execution order and frequency. Through which he obtained excellent classification efficiency.

Single Nucleotide Polymorphisms (SNPs) cause many diseases in humans. Raid Alzuri et al [22] implement a hybrid feature selection of the Conditional Mutual Information Maximization (CMIM) filter method and Support Vector Machine Recursive Feature Elimination (SVM-RFE) wrapper method to produce better classification accuracy with high-dimensional space is essential for analyzing the disease and helps to provide better medication. This method provides better time effective.

Comparison of Some Existing Feature Selection Algorithms

S. No	Algorithm	Method	Methodology	Merits	Demerits
1[8]	Fast Clustering-Based Feature Selection Algorithm (FAST).	Filter	Graph-Theoretic Clustering Method.	Minimize amplitude.	Applicable only for Microarray data.
2[9]	Fast Correlation Based Filtering (FCBF)	Filter	Symmetric Uncertainty (SU).	Minimize amplitude.	Lack of feature over abundance.
3[10]	Relief-F algorithm	Filter	Application Estimation	Efficient for high dimensionality.	Omission of overabundance is impossible
4[11]	Sequential Learning Framework for Feature Subset Selection (SLSS)	Filter	Neyman–Pearson test-based feature selection (NPFS). Scalable and Accurate OnLine Approach (SAOLA)	Measure high-dimensional data set with minimum computation period	Extension of SLSS for large-scale distributed data is required.
5[12]	Score-based Criteria Fusion (SCF)	Filter	SU and relief LIBSVM toolbox Matlab	Determines overabundance and eliminates. Provide high correlation	Difficulty in assignment of the weighted parameter
6[13]	SVM with kernel tricks	Wrapper	Classification and prediction in LibSVM.	Influences the classifier performance. Robustness	Improvement in effectiveness needed. Expensive in high dimensionality.
7[14]	Cosine Similarity Measure Support Vector Machines (CSMSVM)	Wrapper	Multi-class Bayes Decision Rule.	Better computational effectiveness	Improve the computational efficiency is required. Should enhance the optimization process. Undesirable effects caused by unbalance data sets.
8[15]	BIG-F algorithm	Wrapper	Kernel-based Bayesian classifier. AUC as the fitness function	Greater interpretable and concurrence. Increase performance.	Need to increase feasibility, the accuracy and the efficiency in the other dataset.
9[16]	Support Vector Machine (SVM)	Wrapper	Wrapper Methods in WEKA system. Classification and prediction in LibSVM.	Reduce dimensionality, computation period and expense. Increase accuracy	Enhancing and producing a technique for choosing the efficient learning algorithm is required to increase performance.

10[17]	EB algorithm && DDMCMC algorithm	Wrapper	NHTSA specification.	Robustness Accuracy. Real world application.	Wrapper-based image segmentation algorithm efficiency should be advanced. Test with Alternative blob combining algorithms. The eradication of various substances from an identical picture should be performed by enhancing the EM algorithm.
11[18]	Mutual Information Feature Selection (MIFS) && Modified Binary Cuckoo Search (MBCS)	Hybrid	K-Nearest classifiers UCI machine learning repository datasets	Greater classification efficiency and lower computation period.	Computation period is more
12[19]	Mutual Information (MI) &&Sheley model	Hybrid	K-Nearest classifiers	Eliminate the expense and significant to classification performance.	Depends upon combination of different feature method
13[20]	Artificial Neural Network weight Analysis (ANNWA) && Genetic Algorithm	Hybrid	UCI Machine Learning Repository datasets. Matlab on Pentium 4 2.66 GHz PC's with 256 MB memory.	It reduces the rate of dimensional features. Applicable to real-world data set. Economical and more accuracy.	To get an optimal reduction is an NP-hard problem. Time consuming
14[21]	Maximum Pearson Maximum Distance Improved Whale Optimization Algorithm (MPMDIWOA)	Hybrid	UCI Machine Learning Repository.	The threshold alters execution order and frequency provides excellent classification efficiency.	An Experiment done with 20 - 100 dataset.
15[22]	Conditional Mutual Information Maximization (CMIM) && Support Vector Machine Recursive Feature Elimination (SVM-RFE)	Hybrid	SVM, Naive Bayes (NB), Linear Discriminant Analysis (LDA) and k-NN	Better classification accuracy with high-dimensional space and time effective. Better prediction method	Study of e genetic basis of diseases should be done.

Optimisation Algorithm

Bee Colony Algorithm

Bee algorithm is similar to a honeybee grab for resources. It deals with dissipate complication optimization. The Bee algorithm is optimal to get better accuracy. It should specify the number of parameters. Bee algorithm rectifies complications without realm knowledge. It is essential to estimate the robustness [40]. It has the capability to eliminate

local optima. Artificial Bee Colony (ABC) optimization deals with real-world complications. The modifications of ABC for pursuing constant, congregational, lean on, multifunction, and extensive optimization complication and numerous engineering requisition of ABC[41].

Bee Colony Algorithm is like food rummage honeybee. It is an excellent bobble dance for keeping the source perfect and the search for a new one. BA is essential to deal with optimization complications. Lake Ozbakir et al [23] gives a brief explanation of BA and steps involve to compute Generalized Assignment Problem (GAP) and his influences bee-inspired algorithm efficiency from BA on NP-hard complication. BA includes Meta-Heuristic algorithms. It is a speedy and valuable method to search pattern concerning for to time and space. It deals with excellent for a limited complication, efficient for enormous and restrained for GAP.

Optimization is done by Swarm intelligence. It is done with different animals like ants, termites, bees, birds and fishes attitude. The environment and resources are utilized by collective intelligence are named as a swarm. Dervis Karaboga [24] explains briefly about the algorithm and do the survey. The high-level response concerning for to low-level interaction is done by Self-organization. It allocates tasks and behaviors concerning for to atmosphere adjustment. Drones, queen, and workers are three categories of Bee Colony algorithms.

Yannis Marlnakls et al [25] proposed Artificial Bee Colony (ABC) and Greedy Randomized Adaptive Search Procedure (GRASP) to perform grouping. ABC does feature selection discrete complication and Clustering complication is performed by GRASP. He compared the performance with another algorithm and concluded DABC and GRASP together provide better consequence and accomplishment.

Cuckoo Search

Cuckoo algorithm is persuaded by the birdlife. Egg laying and breeding stimulate the current advanced optimization algorithm. The Algorithm applies to some standard operation and a real difficulty has justified its ability to handle different optimization complications. It provides greater efficiency and better refinement [42]. It collaborates regulation and volatility. Adjusting parameter by upgrading cuckoo algorithm is to magnify the efficiency and concurrent rate. Enhanced Cuckoo Algorithm by Lévy Flights random walk [43].

Xin-She Yang et al [26] proposed Cuckoo Search (CS) for optimization complications. CS combines offspring parasitic attitude with birds Levy flight attitude. He had done the algorithm estimation with a test function and measure the efficiency by Genetic Algorithm and Particle Swarm Optimization. From the result, he concluded Cuckoo Search is an excellent method for multimodal objective functions. It is a comprehensive and strong method for optimization complications.

Numerical optimization complication is dealing with CS. Concerning Levy flight better solution has been obtained. Jiatang Cheng et al [27] proposed cuckoo search with neighborhood attraction (NACS). It is a ring topology to mentor exclusive evolution. The Degree of exclusive evolution is monitored by the step size. He combines NACS and levy flight to develop potential aspirant results. From the experimental results, he demonstrated NACS aggressively makes perfect and switching parameter and neighborhood size settings are interrogated.

CS is an aggressive optimization algorithm. Revised CS transfer data between the best results. S. Walton et al [28] validate the revised CS through Standard optimization benchmarking functions. He observed that revised CS provides better performance than the other optimized algorithm. It gives a maximum conflux rate even at the larger dataset. Revised CS gives a path where the estimation of the objective function is a costly calculation. It can be done with a local search only. The implementation setup is simple with two parameters. It provides the best solution for engineering optimization complications.

Particle Swarm Optimisation

Chunkai Zhang et al [29] proposed artificial neural networks (ANN) along with particle swarm optimization (PSO) algorithm. ANN deals with architecture and weight. The best ANN is obtained by repeating the steps. From the experiment, he obtained a better efficiency and generalization capability. It efficiently lessens the clamorous fitness estimation complication. With additional node and a limited training, innovation helps to manage the association between parent and child. It also averts annihilation. PSO acquires cache and effective association among individual.

Even though PSO is acknowledged optimization innovation, it endures in complete concurrence and molecules are captured by local meridian. AREzaee Jordehi [30] proposed Enhanced Leader PSO (ELPSO) to overcome these drawbacks. ELPSO From the outcome of the experiment he concluded ELPSO accomplished well at certainty,

extensibility, and concurrence rate. It consists of consecutive five-step metamorphosis approach for each cycle of swarm leader. Through the cyclic process, it selects the best fitness value.

Mauricio Zambrano-Bigiarini et al [31] proposed a Standard Particle Swarm Optimization algorithm to describe, distinguish conference on Real-Parameter individual goal Optimization at CEC-2013. From the experiment, he achieved superior accomplishment by SPSO-2011. It provides vast convergence in an optimal worldwide area with a limited $10E+03$ -performance estimation. It provides guidelines for further PSO development. Best efficiency is achieved by unimodal and separable functions. It provides constant performance up to dimension $D=50$ with a constant population. It is essential for a large dataset where the calculation period might be restricted.

Genetic Algorithm

GA includes reproduction, crossover, and mutation operation. It maneuvers the population's genetic composition. During reproduction operation, extremely ranked features in the present generation are reproducing for the next generation. During the crossover operation, two parent's intelligences are recombined to generate two offspring. Each individual candidate gene values are altered in mutation. GA operation includes two stages. In the first stage, choosing the number of crossing sites with their parents. In the second step, transferring alternate selection pairs generate two new entities [34].

For optimization of a complex problem, class and parameter of optimization are essential. It is decided based on the application. It is adjusting to get better performance. Genetic Algorithm (GA) provides the best result for mathematical optimization complication. John J. Grefenstette [32] tested a Genetic Algorithm on image registration complication. From the experiment, he fined the GA parameterized space to determine the performance of mathematical optimization operation. . The data set also optimizes GA control parameters. It helps to improve efficiency. It is done with a range of the GA control parameter range. The data set proposed an advanced establishment between the control parameters. It helps to advance theoretical awareness, responsibility.

Yiu-Wing Leung et al [33] intensify the Genetic Algorithm, which is prosperous with mathematical logic. To adjust the a traverse operator, Orthogonal is appropriate only to a discrete variable so he proposed quantization technique along with orthogonal to deal with a continuous variable. It assigns the primary population and the crossover function of minute point as possible offspring. Here he proposed GA to deal with a 15-criterion

complication with 30 to 100 measurements and enormous local margin and achieve excellent results.

Chia-Feng Juang [34] proposed an evolutionary learning algorithm of the Hybrid Genetic Algorithm (HGA) and Particle Swarm Optimization (PSO) defined as HGAPSO. A recent creation of an individual is not only designed by crossover and metamorphosis function in GA, likewise by PSO. PSO improvised elites. HGAPSO is applied to the fuzzy neural network. He did the efficiency comparison with GA and PSO. PSO provides the cache and good result where GA provides better communication between particles in the swarm. FCRNN gives the solution in temporal sequence production and TRFN illustrates the dominance of HGAPSO above GA and PSO respectively.

Ant Colony Optimisation Algorithm

Ant Colony Algorithm is a meta-heuristic. It resolves discrete optimization complications and combinatorial optimization problems. ACO behaves exactly blind ant searches for resources. It produces the lowest expense possible results. It has a cache to store the instruction [38]. Hard combinatorial optimization complications are resolved by the ANT algorithm. Hence, it is a novel inspired method. The inspiring is exactly similar to rummage action of ant.

The eradication of classification rule from data is Ant -Miner. Rafael S. Parpinelli et al [35] implements the Ant Colony Optimization (ACO) Algorithm and analyze the efficiency with CN2. In the case of anticipating precisions, Ant-Miner is aggressive with CN2. It is an easy and simple implementation. ACO algorithm is an easy, prosperous operator to determine superior results for enormous search space complications. It reduces the discovery rules which help to reduce the computation complexity.

Yun-Chia Liang et al [36] rectify the Redundancy Allocation Problem (RAP) by applying an ant colony meta- heuristic optimization method. RAP is a NP-hard problem. It is an architecture for combinatorial complications. The efficiency of an ant colony algorithm is a bit unstable in case of a complication situation. ACO gives better performance than GA. ACO provides numerous results for every emphasis. Some of the features of GA such as adaptability, durability, application comfort are also contributing by ACO.

Marco Dorigo et al [37] survey about Ant Colony Optimization. First analyzed about some concurrent solution. Second, he explained about an Ant Colony Optimization with the other optimization algorithm. Third, he concentrates on the performance of the Ant Colony Optimization Algorithm. The interior of model-based search demonstrates the relation

between ACO, hypothetical slope domination and cross-entropy approach. ACO is applicable for optimal discipline and reinforcement study.

Comparison of Some Optimisation Algorithms

S.No	Algorithm	Factors/ Approaches Used	Benefit	Drawback
1[23]	Bee Colony Optimization	Compute Generalized Assignment Problem (GAP)	Speedy, valuable method to search pattern. Limited complication. Efficient.	No parameter optimization is carried out. Improvement in performance with large dataset is required.
2[24]	Bee Colony Optimization	Bee Swarm Intelligence.	High-level response with respect to low-level interaction	Parameters need to be tuned. Random initialization.
3[25]	Bee Colony Optimization	Greedy Randomized Adaptive Search Procedure (GRASP) to perform clustering.	Deals with discrete complication. Better consequence and accomplishment	The Algorithm should be implemented for the both feature selection and clustering algorithm phase.
4[26]	Cuckoo Search Optimization	Lévy Flights random walk. GA and PSO.	Excellent method for multimodal objective functions. Comprehensive.	Focus on the sensitivity and parameter studies is required.
				Increase in convergence rate is required.
5[27]	Cuckoo Search Optimization	Neighborhood attraction (NACS)	Develop potential aspirant result.	NACS done only by ring topology. A flexible alteration is done to enlarge the explore capability.
6[28]	Cuckoo Search Optimization	Standard optimization benchmarking functions	Provides better performance. Gives maximum conflux.	Large-scale global optimization is required.
7[29]	Particle Swarm Optimization	ANN'S evolved by PSONN. UCI machine learning benchmark repository datasets	Better efficiency and generalization capability. Averts annihilation	PSO can be enhanced by using a multiple group of particle
8[30]	Particle Swarm Optimization	Enhanced Leader PSO	Work well at certainty, extensibility, and concurrence rate.	Less conversion efficiency. Higher steady state oscillations.
9[31]	Particle Swarm Optimization	Standard PSO (SPSO-2011)	Best efficiency is achieved by unimodal and separable function	Improvement in PSO for further optimization is required.
10[35]	Ant Colony Optimization	Ant-Miner with CN2	Reduces the discovery rules. Reduce the computation complexity	Ant-Miner to limited with continuous attributes. Low performance in other heuristic function
11[36]	Ant Colony Optimization	ACO-RAP algorithm	Adaptability, durability, application comfort	Designing swarm-based systems is hard
12[37]	Ant Colony Optimization	Ant Colony Optimization	Applicable for optimal discipline and reinforcement study.	Difficult to design initial design parameter. Not applicable for scattering problems.
13[32]	Genetic Algorithm	Genetic Algorithm	Determine the performance of mathematical optimization operation	Neglected multipoint crossover and inversion operator. Unconstrained optimization procedures
14[33]	Genetic Algorithm	Orthogonal genetic algorithm with quantization (OGA/Q)	Deal with complication and provide better results	Limited with continuous variable.
15[34]	Genetic Algorithm	Hybrid Genetic Algorithm Particle Swarm Optimization (HGAPSO)	Better communication between particles in the swarm	Not applicable to all optimization problems.

Conclusion

From the review of existing feature selection algorithms, some algorithms considered relevant feature selection where an increase in dimensionality occurs due to redundancy and few algorithms select the features without the negation of noisy data. This leads to less efficiency and more computation period. The implementation of both filter and wrapper method in future work will be more effective. Few optimization algorithms related to feature selection are also discussed in this paper. It aims to prefer the optimal procedure and approaches functioning with continuous swarm algorithms for discrete complications. From the study, the most of the classical combinatorial complications are dealt with PSO.

References

- Trabelsi, M., Meddouri, N., & Maddouri, M. (2017). A new feature selection method for nominal classifier based on formal concept analysis. *Procedia Computer Science*, 112, 186-194.
- Sutha, K., & Tamilselvi, J.J. (2015). A review of feature selection algorithms for data mining techniques. *International Journal on Computer Science and Engineering*, 7(6), 63.
- Karabulut, E.M., Özel, S.A., & Ibrikci, T. (2012). A comparative study on the effect of feature selection on classification accuracy. *Procedia Technology*, 1, 323-327.
- Kavipriya, P., & Karthikeyan, K. (2017). A Comparative Study of Feature Selection Algorithms in Data Mining. *International Journal of Advanced Research in Computer and Communication Engineering*, 6(11).
- Ravi, R.V., Subramaniam, K., Roshini, T.V., Muthusamy, S.P.B., & Venkatesan, G.P. (2019). Optimization algorithms, an effective tool for the design of digital filters; a review. *Journal of Ambient Intelligence and Humanized Computing*, 1-17. <http://doi.org/10.1007/s12652-019-01431-x>.
- Sheik Abdullah, A., Ramya, C., & Priyadharsini, V. (2017). A Survey on Evolutionary Techniques for Feature Selection. *IEEE Conference on Emerging Devices and Smart Systems (ICEDSS2017)*.
- Vidyavathi, B.M. (2011). A new approach to feature selection for data mining. *International Journal of Computational Intelligence Research*, 7(3), 263-269.
- Qinbao, S., Jingjie, N., & Guangtao, W. (2013). A Fast Clustering-Based Feature Subset Selection Algorithm for High-Dimensional Data. *IEEE Transactions on Knowledge and Data Engineering*, 25(1), 1-14.
- Senliol, B., Gulgezen, G., Yu, L., & Cataltepe, Z. (2008). Fast Correlation Based Filter (FCBF) with a different search strategy. *In IEEE 23rd international symposium on computer and information sciences*, 1-4.
- Wang, Y., & Makedon, F. (2004). Application of Relief-F feature filtering algorithm to selecting informative genes for cancer classification using microarray data. *In IEEE Proceedings. IEEE Computational Systems Bioinformatics Conference, CSB*, 497-498.

- Ditzler, G., Polikar, R., & Rosen G. (2017). A sequential learning approach for scaling up filter-based feature subset selection. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6), 2530-2544.
- Ke, W., Wu, C., Wu, Y., & Xiong, N.N. (2018). A new filter feature selection based on criteria fusion for gene microarray data. *IEEE Access*, 6, 61065-61076.
- Maldonado, S., & Weber, R. (2009). A wrapper method for feature selection using support vector machines. *Information Sciences*, 179(13), 2208-2217.
- Chen, G., & Chen, J. (2015). A novel wrapper method for feature selection and its applications. *Neurocomputing*, 159, 219-226.
- Wuniri, Q., Huangfu, W., Liu, Y., Lin, X., Liu, L., & Yu, Z. (2019). A Generic-Driven Wrapper Embedded With Feature-Type-Aware Hybrid Bayesian Classifier for Breast Cancer Classification. *IEEE Access*, 7, 119931-119942.
<http://doi.org/10.1109/ACCESS.2019.2932505>.
- Roslina, A.H., & Noraziah, A. (2010). Prediction of hepatitis prognosis using support vector machines and wrapper method. In *IEEE Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, 5, 2209-2211.
- Farmer, M.E., & Jain, A.K. (2005). A wrapper-based approach to image segmentation and classification. *IEEE transactions on image processing*, 14(12), 2060-2072.
- Jiang, Y., Liu, X., Yan, G., & Xiao, J. (2017). Modified binary cuckoo search for feature selection: a hybrid filter-wrapper approach. In *IEEE 13th International Conference on Computational Intelligence and Security (CIS)*, 488-491.
<http://doi.org/10.1109/CIS.2017.00113>.
- Liu, J., & Wang, G. (2010). A hybrid feature selection method for data sets of thousands of variables. In *IEEE 2nd International Conference on Advanced Computer Control*, 2, 288-291.
- Vidyavathi, B.M. (2011). A new approach to feature selection for data mining. *International Journal of Computational Intelligence Research*, 7(3), 263-269.
- Zheng, Y., Li, Y., Wang, G., Chen, Y., Xu, Q., Fan, J., & Cui, X. (2018). A novel hybrid algorithm for feature selection based on whale optimization algorithm. *IEEE Access*, 7, 14908-14923. <http://doi.org/10.1109/ACCESS.2018.2879848>.
- Alzubi, R., Ramzan, N., Alzoubi, H., & Amira, A. (2017). A hybrid feature selection method for complex diseases SNPs. *IEEE Access*, 6, 1292-1301.
<http://doi.org/10.1109/ACCESS.2017.2778268>.
- Özbakir, L., Baykasoğlu, A., & Tapkan, P. (2010). Bees algorithm for generalized assignment problem. *Applied Mathematics and Computation*, 215(11), 3782-3795.
- Karaboga, D., & Akay, B. (2009). A survey: algorithms simulating bee swarm intelligence. *Artificial intelligence review*, 31(1-4), 61.
- Marinakis, Y., Marinaki, M., & Matsatsinis, N. (2009). A hybrid discrete artificial bee colony-GRASP algorithm for clustering. In *IEEE International Conference on Computers & Industrial Engineering*, 548-553.
- Yang, X.S., & Deb, S. (2009). Cuckoo search via Lévy flights. In *IEEE World congress on nature & biologically inspired computing (NaBIC)*, 210-214.

- Cheng, J., & Wang, L. (2019). Cuckoo Search Algorithm with Neighborhood Attraction for Numerical Optimization. *IEEE Access*, 7, 122261-122274.
<http://doi.org/10.1109/ACCESS.2019.2937741>.
- Walton, S., Hassan, O., Morgan, K., & Brown, M.R. (2011). Modified cuckoo search: a new gradient free optimisation algorithm. *Chaos, Solitons & Fractals*, 44(9), 710-718.
- Zhang, C., Shao, H., & Li, Y. (2000). Particle swarm optimisation for evolving artificial neural network. In *IEEE SMC 2000 conference proceedings. IEEE international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions'*, 4, 2487-2490.
- Jordehi, A.R. (2015). Enhanced leader PSO (ELPSO): a new PSO variant for solving global optimisation problems. *Applied Soft Computing*, 26, 401-417.
- Zambrano-Bigiarini, M., Clerc, M., & Rojas R. (2013). Standard particle swarm optimisation 2011 at cec-2013: A baseline for future pso improvements. In *IEEE Congress on Evolutionary Computation*, 2337-2344.
- Grefenstette, J.J. (1986). Optimization of control parameters for genetic algorithms. *IEEE Transactions on systems, man, and cybernetics*, 16(1), 122-128.
- Leung, Y.W., & Wang, Y. (2001). An orthogonal genetic algorithm with quantization for global numerical optimization. *IEEE Transactions on Evolutionary computation*, 5(1), 41-53.
- Juang, C.F. (2004). A hybrid of genetic algorithm and particle swarm optimization for recurrent network design. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(2), 997-1006.
- Parpinelli, R.S., Lopes, H.S., & Freitas, A.A. (2002). Data mining with an ant colony optimization algorithm. *IEEE transactions on evolutionary computation*, 6(4), 321-332.
- Liang, Y.C., & Smith, A.E. (2004). An ant colony optimization algorithm for the redundancy allocation problem (RAP). *IEEE Transactions on reliability*, 53(3), 417-423.
- Dorigo, M., & Blum, C. (2005). Ant colony optimization theory: A survey. *Theoretical computer science*, 344(2-3), 243-278.
- Dorigo, M., & Di Caro, G. (1999). Ant colony optimization: a new meta-heuristic. In *Proceedings of the 1999 congress on evolutionary computation-CEC99 (Cat. No. 99TH8406)*, 2, 1470-1477.
- Min, H., & Fangfang, W. (2010). Filter-wrapper hybrid method on feature selection. In *IEEE Second WRI Global Congress on Intelligent Systems*, 3, 98-101.
<http://doi.org/10.1109/GCIS.2010.235>.
- Ghanbarzadeh, A. (2008). The Bees Algorithm. *5th International Conference on Informatics in Control, Automation and Robotics ICINCO2008*.
- Dervis, K., Beyza, G., Celal, O., & Nurhan, K. (2012). A comprehensive survey: artificial bee colony (ABC) algorithm and applications. *Springer Science+Business Media B.V.*
- Rajabioun, R. (2011). Cuckoo optimization algorithm. *Applied soft computing*, 11(8), 5508-5518.
- Valian, E., Mohanna, S., & Tavakoli, S. (2011). Improved cuckoo search algorithm for global optimization. *International Journal of Communications and Information Technology*, 1(1), 31-44.

- Bai, Q. (2010). Analysis of particle swarm optimization algorithm. *Computer and information science*, 3(1), 180-184.
- Lawrence, N.A., & Shiny Angel, T.S. (2020). Survey of Different fusion Techniques and Quality Metric Assessment for Remote Sensing Image Fusion. *International Journal of Advanced Science and Technology*, 29(6s), 1337–1348.