Polytechnic Journal. 2022. 12(2) 203-214
ISSN: 2313-5727
http://journals.epu.edu.iq/index.php/polytechnic

RESEARCH ARTICLE

REVIEW OF FEATURE SELECTION METHODS USING OPTIMIZATION ALGORITHM (Review paper for optimization algorithm)

Zana O. Hamad [1]

[1] Department of Information Technology, Dukan technical institute, Sulaimani Polytechnic University, Sulaimani, Kurdistan Region, Iraq

***Corresponding author:**
Zana O. Hamad,
Department of Information Technology, Dukan technical institute,
Sulaimani Polytechnic University, Sulaimani,
Kurdistan Region, Iraq.

E-mail:
zana.othman@spu.edu.iq

Received: 26 April 2022
Accepted: 24 November 2022
Published: 1 February 2023

DOI
10.25156/ptj.v12n2y2022.pp203-214

ABSTRACT

Many works have been done to reduce complexity in terms of time and memory space. The feature selection process is one of the strategies to reduce system complexity and can be defined as a process of selecting the most important feature among feature space. Therefore, the most useful features will be kept, and the less useful features will be eliminated. In the fault classification and diagnosis field, feature selection plays an important role in reducing dimensionality and sometimes might lead to having a high classification rate. In this paper, a comprehensive review is presented about feature selection processing and how it can be done. The primary goal of this research is to examine all of the strategies that have been used to highlight the (selection) selected process, including filter, wrapper, Meta-heuristic algorithm, and embedded. Review of Nature-inspired algorithms that have been used for features selection is more focused such as particle swarm, Grey Wolf, Bat, Genetic, wale, and ant colony algorithm. The overall results confirmed that the feature selection approach is important in reducing the complexity of any model-based machine learning algorithm and may sometimes result in improved performance of the simulated model.

KEYWORDS: Features selection, Filter process, Wrappers, embedded, Metaheuristic algorithm.

INTRODUCTION

The complexity of any system is a critical issue, and it has been a promising area for researchers. Dimensional reduction is a way to reduce system complexity. One of the basic requirements for any machine learning application is having a huge dataset. Nowadays, the size of the recorded dataset is become bigger and bigger that leads to achieving a system with high complexity. For this purpose, many approaches have been developed to reduce dimensionality (Van Der Maaten et al., 2009). Dimensional reduction helps in data compression because it reduced storage space and computation time. It is also a way to eliminate redundancies from the dataset. There are two main strategies to conduct Dimensional reduction: feature extraction techniques and feature selection techniques (Khalid et al., 2014).

Moreover, the feature extraction process is a process of extracting and tackling hidden information in the raw signal. Any automatic system based on machine learning, requires a huge dataset in order to learn an unknown pattern. Data can be more manageable by applying feature extraction because it removes all fewer effective features from the data without misplacing any significant or applicable data. Moreover, f eature extraction techniques help to develop a system with less mechanism's exertions and upsurge the rapidity of learning and generality phases in the process of machine learning (Cremona et al., 2009; Sharma et al., 2016).

Furthermore, feature selection is a key to enhance the performance of all machine learning-based models in terms of computation time and memory space because feature selection is a technique for selecting the most important feature among feature space. Furthermore, Redundant features are kicked off from obtained features and sometimes lead to improve the accuracy of the classification model. Feature selection has been adopted in various approaches such as filter, wrapper, embedded, nature-inspired algorithm(Jović et al., 2015).

Consequently, the main purpose of this study is to review all kinds of strategies that have been used to select a proper feature which is called as feature selection process. There are three main ways to conduct a feature selection as shown in figure 1, including filter, wrapper, and embedded. Moreover, more attention has focused on wrapper strategies especially those which meta-heuristic algorithms have been involved. Recently, many Meta-heuristic algorithms have been used for feature selection. Some of them are mentioned and cited. The rest of the research is structured as follows; feature selection and types are introduced in section two. Finally, the conclusion of the paper is placed in section three.

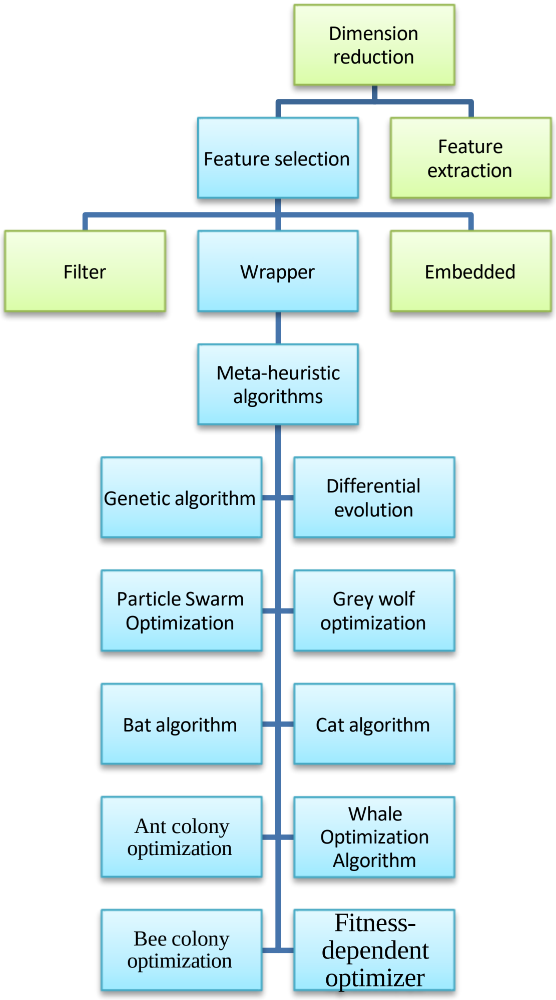

Figure 1:K-Chart for this paper

FEATURE SELECTION

The feature selection process has been widely premeditated and conducted by machine learning applications. Dimension reduction can be adopted into two ways, namely features extraction process and feature selection process. Feature extraction is extracting hidden and effective information from the signal. Subsequently, the feature can be defined as a pattern that is the most representative of the raw signal. However, the feature selection process is choosing the most important and operative characteristics among the obtained features. The feature selection strategy can be divided into four main strategies based on how the most effective features are selected. The strategies are introduced properly, including filter, wrappers, embedded, and Metaheuristic algorithm(Vergara & Estévez, 2014). In this section, all three kinds of feature selection processes are mentioned and some of the mate-heuristic algorithms has been explained.

Filter:

Filter approaches assess the goodness of gene subsections by perceiving individual inherent data features such as statistical procedures. Commonly, a gene or a subsection of genes is appraised regarding the class label. Many filter methods have been developed to feature selection in various applications, including Fast Correlation-Based Filter, Correlation Feature Selection, and (Yu & Liu, 2003) consistency-based filter(Manoranjan Dash & Liu, 2003).

Wrappers:

Wrapper methods are referred to as those algorithms where they measure the effective and efficient feature based on the dedicated classifier performance. This is unlike filter methods because filter methods select the feature based on some statistical value such as information gain, chi-square test, fisher score, and correlation coefficient. Additionally, Wrapper methods feed some features to the classifier and measure the classifier performance, then test other subsets of the features (genes)(Kohavi & John, 1997). wrapper methods have not received less attention from the researcher compared to the filter methods because wrapper methods lead to have high computational costs. What is more, there are many approaches that have been developed for this kind of feature selection, such as sequential feature selection algorithms, recursive feature elimination, and Meta-heuristic algorithm. in this paper, some of the Meta-heuristic algorithms are highlighted that have been used for feature selection.

Metaheuristic algorithm:

Scientists and scholars are increasingly using metaheuristic algorithms to solve and antagonize a widerange of problems. They are unambiguously applied for problems with a large number of constraints or for problems that do not have investigative solutions. Similarly, the presence of the theory of the adequate evolutionary algorithm, the conduct of progressively stimulated algorithms, and the combined of swarm intellect constituent part in nature all contributed to the development of the majority of these algorithms (Bozorg-Haddad, 2018). As there is no such principle as a free lunch rule, it will be problematic to develop an effective universal algorithm for solving almost entirely problems. As a result, researchers all over the world continue to look for new optimization algorithms. Subsequently, according to a recent study conducted by (Diao & Shen, 2015), meta-heuristic algorithms play a vital role in a wide range of optimization problems and applications, including feature extraction. Some meta-heuristic algorithms applied in feature selection have been investigation.

Genetic algorithm:

Genetic algorithm is one of the well-known Meta-heuristic algorithms that was inspired by biological progress. It was developed by John Holland in 1975. the optimization is explored based on four main steps. The First step was selecting parents and passed to the mating pool. The second step is the cross-over process that leads to have a huge change in parents. The third step is the mutation process, which leads to have a small change in the parents(Guo et al., 2005). The last step is the pooling update step (see figure 2). Likewise, the four above processes will be reiterated until the best chromosome will be survived. Genetic algorithm has been participated in many applications and optimization problems such feature generation(Guo et al., 2005) and feature selection(Khalid et al., 2014). Accordingly, the feature subset selection was obtained using a genetic algorithm to progress the performance of the neural system to identify a pattern(J. Yang & Honavar, 1998). In a health monitoring system, a genetic algorithm was used to choose the most vital features from whole gearbox fault features(Hajnayeb et al., 2011).

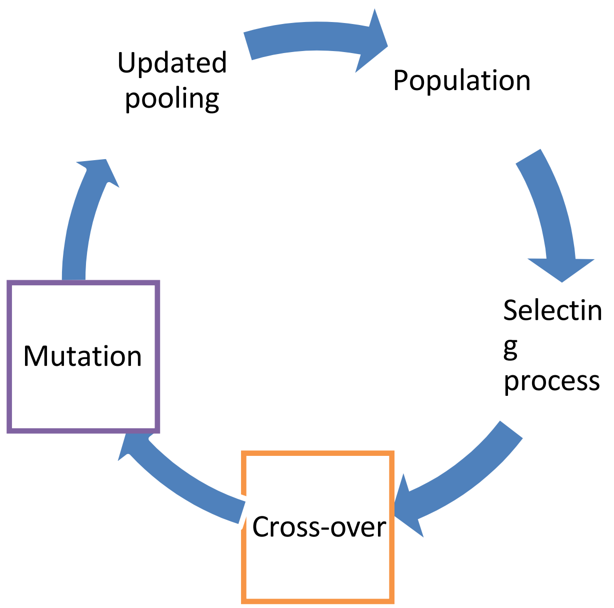

Figure 2. Genetic algorithm

Successively, the feature selection problem is one of the eminent issues that the optimization algorithm plays an important role in solving it. ZorarpacI & Özel, proposed a feature selection model founded on differential evolution and non-natural bee colony. The model was assessed based on various datasets (ZorarpacI & Özel, 2016).

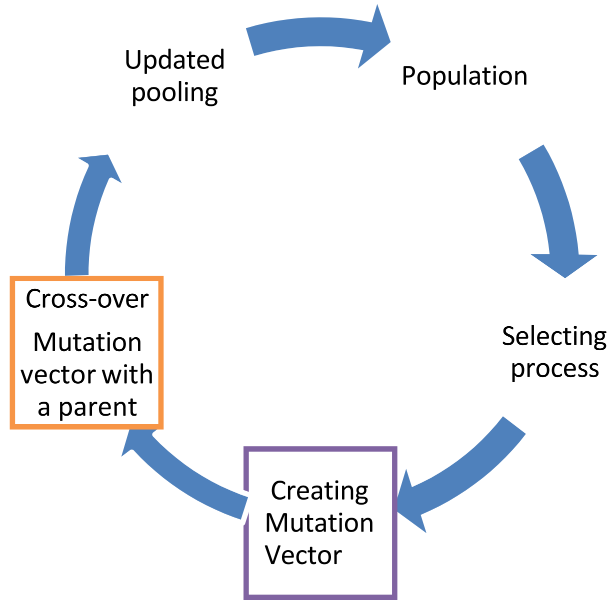

Figure 3: Differential evolution process

Particle swarm optimization:

This kind of optimization is based on population-created optimization method for stochastic exploration in a multidimensional space. It was initially familiarized by Kennedy and Eberhart. Also, Particle Swarm Optimization has so far been conducted successfully for solving a variety of optimization difficulties, including the feature selection process. For that reason, there are two main steps that Particle Swarm Optimization does them during solving a problem. The steps are

Differential evolution algorithm:

The aforementioned algorithm is an established on the population meta-heuristic algorithm that was advanced by Rainer Storn and Kenneth Price. Basically, it is quite inspired by genetic algorithms, and even some researchers considered as an extended genetic algorithm. Besides, Differential evolution algorithm can optimize problem through iteratively improving a candidate solution based on four sequence processes namely, selecting a parent, creating mutation vector based on other parents, cross over between the selected parent and obtained mutation vector, and surviving the best fitness value among new child and selected parent (see figure 3).

finding the new velocity of any point and finding a new position for whole the particles(Sakri et al., 2018). L. F. Chen et al. modified survival rules in Particle Swarm Optimization and then the algorithm used for selecting significant features in various applications (L. F. Chen et al., 2012). A hybrid model based on Correlation-based Feature Selection and improved- Binary Particle Swarm Optimization was invented by (Jain et al., 2018). The model was used for Gene Selection and Cancer Classification.

Grey wolf optimization:

Grey wolf optimization is based on the natural hunting behavior of grey wolves to hunt prey in a supportive manner. Correspondingly, the structure of grey wolf optimization is moderately dissimilar from other meta-heuristic optimization in that three optimal specimens become the base for allencompassing search technique, which are alpha (α) wolf which performances as a pack leader, beta (β) wolf which maintenances the leader, and delta (δ) wolf which monitors the leader and the supportive wolves. The last category of wolf is identified as omega (ω). These wolves differ in terms of accountability, and they have the capacity to be accessible as categorized in such a method that the upper and leading solution is alpha (α), and the second, third, and final solutions are beta (β), delta (δ), and omega (ω) consecutively. Accordingly, omegas are directed through the three preceding categories of wolves (as can be seen in figure 4). Grey wolves optimization has been effectively smeared to several problematic mutual problems such as migrant salesman, feature selection, quadratic assignment(Faris et al., 2018; T. A. Rashid et al., 2018). Emary et al. recommended a binary account of the gray wolf optimization to select the most effective features from Breast cancer images (Emary et al., 2016).

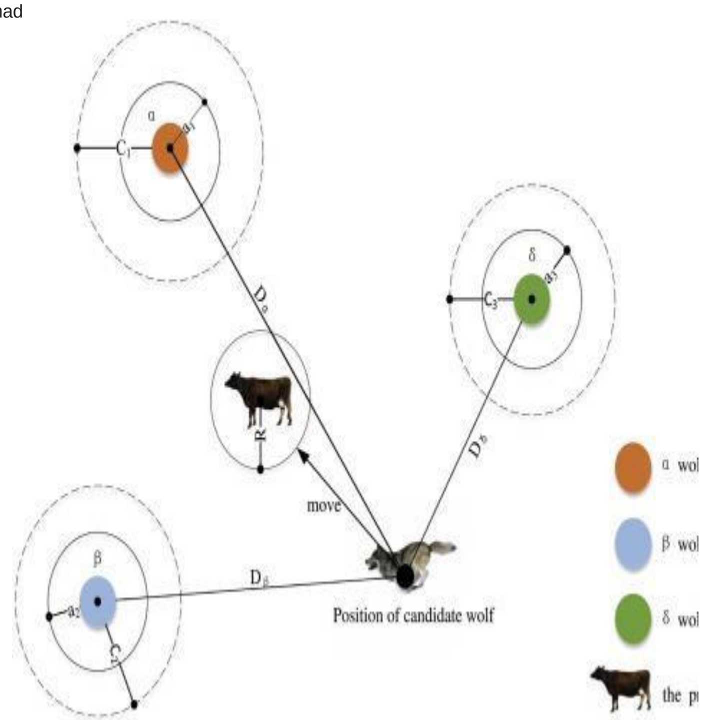

Figure 4: Grey wolf algorithm(Niu et al., 2019)

Ant colony optimization:

Ant colony optimization is a population-based optimization method simulated by ants' behavior of analyzing the unswerving path amid their nest and modern discovered food. As well, in all iterations, the ants spread a certain amount of pheromone values, which causes all the ants who have found a solution in the repetition to be updated. In addition, more pheromone in the path indicates that it is the best and shortest path (Oscar Cordon, Francisco Herrera, 2002). Moreover, many optimization problems, such as feature selection and design mechanical elements, have been optimized and solved by using ant colony optimization. Using ant colony optimization, Aghdam et al. chose a text feature. When compared to the performance genetic algorithm, the selected feature could improve performance (Aghdam et al., 2009). Moreover, according to a study performed by (Ahmed, 2005) investigated a novel technique for selecting feature subset via the Ant Colony Optimization. In addition, a new rough set method to feature selection based on Ant Colony Optimization was proposed in (Y. Chen et al., 2010).

Bat optimization:

Bat algorithm is a population-based algorithm that was developed by Xin-She Yang in 2010, and it is a bio-inspired bat behavior. Based on the literature, the Bat algorithm has been found to be a very efficient algorithm to find out the optimal solution. Respectively, many science and engineering optimizations problem recommend to use bat algorithm to tune or get the optimal solution(X. S. Yang, 2013). And so, numerous researchers have preferred Bat algorithm for adopting feature selection; for instance, Nakamura et al. suggested a binary bat algorithm with a speed of the Optimum-Path Forest classifier for choosing the most effective feature (Nakamura et al., 2012). Taha et al. proposed a hybridization between bat algorithm and a Naive Bayes classifier. Consequently, the proposed model was evaluated in various domains, including time, frequency, and time-frequency (Taha et al., 2013).authors in (Jeyasingh & Veluchamy, 2017) enhanced the bat algorithm and applied it to obtain the most significant features among Breast Cancer' features.

Cat optimization:

According to (Chu et al., 2006) Cat optimization has been considered as an Intelligence optimization algorithm in. furthermore, it is a population-based algorithm stimulated by cat behavior, with a unique procedure for displaying the investigation and manipulation stages. In a similar system, the Cat optimization algorithm has efficaciously unraveled a diversity of optimization problems. Optimizing structures is one of the applications of cat optimizers. For instance, Modified Cat Swarm Optimization was recommended to feature selection and alteration parameters of support vector machine (Bozorg- Haddad, 2018). The Modified Cat Swarm Optimization outcome confirms a substantial enhancement in precision rate (Lin et al., 2015; Ahmedet al., 2020). Additionally, Lin et al. used advanced cat swarm optimization to choose critical features for a text arrangement experiment for giant statistical data (Lin et al., 2016).

Whale optimization:

Whale optimizer is an optimization population-based algorithm that is simulated by means of whale behavior. It was invented by Mirjalili and Lewis in 2016. The process of the whale algorithm is made up of dual chief sections, the initial stage (exploration phase) is surrounding prey and curved updating situation, and the second stage (exploitation phase) is examining for prey (Oscar Cordon, farncisco Herrera, 2002). In view of that, a plenty of applications can be found in the literature that Whale optimizer has been used in it. A whale optimizer was employed to optimize the obtained features from various techniques such as (Sharawi et al., n.d.) and (M. Mafarja & Mirjalili, 2018). A framework constructed on a permutation of the Whale Optimization Algorithm (WOA) And Simulated Annealing algorithm was proposed for optimizing features (M. M. Mafarja & Mirjalili, 2017). Whale's Mechanism for preying is shown in figure 5.

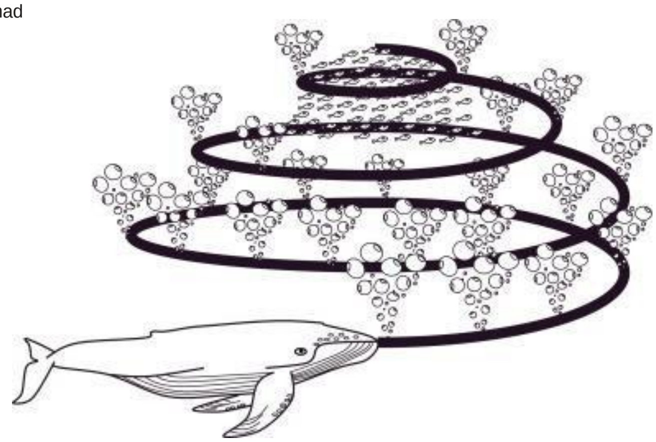

Figure 5: mechanism of whale algorithm(Mirjalili & Lewis, 2016)

Bee colony optimization:

One of the famous optimization algorithms is the bee optimization algorithm, which is inspired by bee swarm behavior and it was introduced by Karaboga in 2005. Three essential components are required to adopt the algorithm, including employed and unemployed bees and food sources. Employed and unemployed bees are in charge of finding the last component, a good food source(KARABOGA, 2005). Oppositely, a comparative study shows that the performance of the bee colony is outperformed by the Genetic algorithm, Particle Swarm, and Differential Evolution algorithm(Karaboga & Akay, 2009). Additionally, there are many NP-hard problems that have been solved using bee colony algorithm like bin packing problem, knapsack problem, and travel sale man. Bee colony has also been used for optimizing features. For instance, Forsati et al. formulated a feature selection by conducting bee colony optimization to improve classification performance. Their approach was evaluated in various datasets like speech signal, breast cancer, and Vehicle (Forsati et al., 2012). Besides, Bee colony optimization was used to solve multi-objective problems; for example, authors in (Hancer et al., 2018) utilized a bee colony algorithm for selecting a feature from Pareto Front.

fitness-dependent optimizer:

The fitness-dependent optimizer is an optimization-based population algorithm and inspired by the bee swarm algorithm. The algorithm was recently developed by Jaza Abdullah in 2019 (Abdullah, 2019). The process of fitness-dependent optimizer can be explored in two steps; namely, scout bee penetrating step and detective bee updated step. In the first step or process, the algorithm tries to find a suitable solution among plenty of solutions. Then, in the scout bee updated step, the scout update position using a random walk and a fitness weight mechanism. Thus, the performance of the technique was appraised based on function19 and 10. The result shows that the fitness-dependent optimizer is outperformed most of the nature-inspired algorithm. Besides, the algorithm is very new; it has been applied for many optimization complications, including feature selection for improving classification performance(Muhammed et al., 2020).

Embedded:

The main advantage of the filter method is lower time consumption compare to the wrapper methods, which have a high computational rate. However, the filter method may lead to obtaining a bad performance as it does not interact with classifiers. To overcome this problem, researchers have investigated an intermediate solution and called embedded methods(M. Dash & Liu, 1997). Therefore, in these methods, all features are examined in a core of the classifier itself to recognized rank features. the feature with high rank will be selected to be the best features among the left features. Vector Machine based on Recursive Feature Elimination is one of the most well-known embedded methods. First Order Inductive Learner and iterative perturbation method are also embedded methods (Shrivastava et al., 2017).

Summary

After collecting the majority of the research publications in which the metaheuristic algorithm is used to select the appropriate features, the observation has been recorded and represented in the table (1) and figure (6). Based on both presentations (table (1) and figure (6)), it is evident that the PSO, DF, and ABC have been employed extensively for feature selection when compared to the other algorithms presented in this paper. The researchers' frequent adoption of these methods could be attributed to their simplicity. The cat algorithm is ranked second, while BAT, WOA, and FDO are ranked last.

Table 1: The most research articles where the metaheuristic algorithm has been used for feature selection process

Algorithms	Ref.
PSO	(Zahran and Kanaan, 2009; Liu *et al.*, 2011; Sahu and Mishra, 2012; Xue, Zhang and Browne, 2012; Aghdam and Heidari, 2015; Ahmad, 2015; Kumar Gupta *et al.*, 2015; Brezočnik, 2017; Abualigah, Khader and Hanandeh, 2018; Qasim and Algamal, 2018)
WOA	(Zheng *et al.*, 2018; Bui *et al.*, 2019; Nematzadeh *et al.*, 2019; Mandal *et al.*, 2021; Too, Mafarja and Mirjalili, 2021)
GA	(Emary *et al.*, 2015; Li *et al.*, 2017; Nirmala Sreedharan *et al.*, 2018; Too *et al.*, 2018; Johari and Gupta, 2021; Kitonyi and Segera, 2021)
DF	(Khushaba, Al-Ani and Al-Jumaily, 2008, 2011; Al-Ani, Alsukker and Khushaba, 2013; Bhattacharyya *et al.*, 2014; Zorarpac\i and Özel, 2016; Vivekanandan and Iyengar, 2017; Dixit, Mani and Bansal, 2020; Hancer, 2020; Zhang *et al.*, 2020; Hancer, Xue and Zhang, 2022)
FDO	(Abdulkhaleq *et al.*, no date; Guha *et al.*, 2020; Chiu *et al.*, 2021; Abbas *et al.*, 2022; Salih, Mohammed and Abdul, 2022)
ABC	(Hancer *et al.*, 2018; Arslan and Ozturk, 2019; Zhang *et al.*, 2019; Wang *et al.*, 2020; Almarzouki, 2022)
Ant colony	(Al-Ani, 2005; Kanan, Faez and Taheri, 2007; Aghdam, Ghasem-Aghaee and Basiri, 2009; Deriche, 2009; Kabir *et al.*, 2009; Moradi and Rostami, 2015; Aghdam, Kabiri and others, 2016; Dadaneh, Markid and Zakerolhossein i, 2016; Fallahzadeh, Dehghani-Bidgoli and Assarian, 2018; Jayaprakash and KeziSelvaVijila, 2019)
CAT	(Ahdesmäki and Strimmer, 2010; Lin, Zhang and Hung, 2014; Lin *et al.*, 2015, 2016; Alarifi *et al.*, 2020; Bansal *et al.*, 2020; Gao *et al.*, 2021; Gomathy, 2021)
BAT	(Nakamura *et al.*, 2012; Yang, 2013; Jeyasingh and Veluchamy, 2017; Tawhid and Dsouza, 2018; Saleem, Zafar and Sabzwari, 2019)
GWO	(Emary *et al.*, 2015; Li *et al.*, 2017; Nirmala Sreedharan *et al.*, 2018; Too *et al.*, 2018; Johari and Gupta, 2021; Kitonyi and Segera, 2021)
DE	(Khushaba, Al-Ani and Al-Jumaily, 2008, 2011; Al-Ani, Alsukker and Khushaba, 2013; Bhattacharyya *et al.*, 2014; Zorarpac\i and Özel, 2016; Vivekanandan and Iyengar, 2017; Dixit, Mani and Bansal, 2020; Hancer, 2020; Zhang *et al.*, 2020; Hancer, Xue and Zhang, 2022)

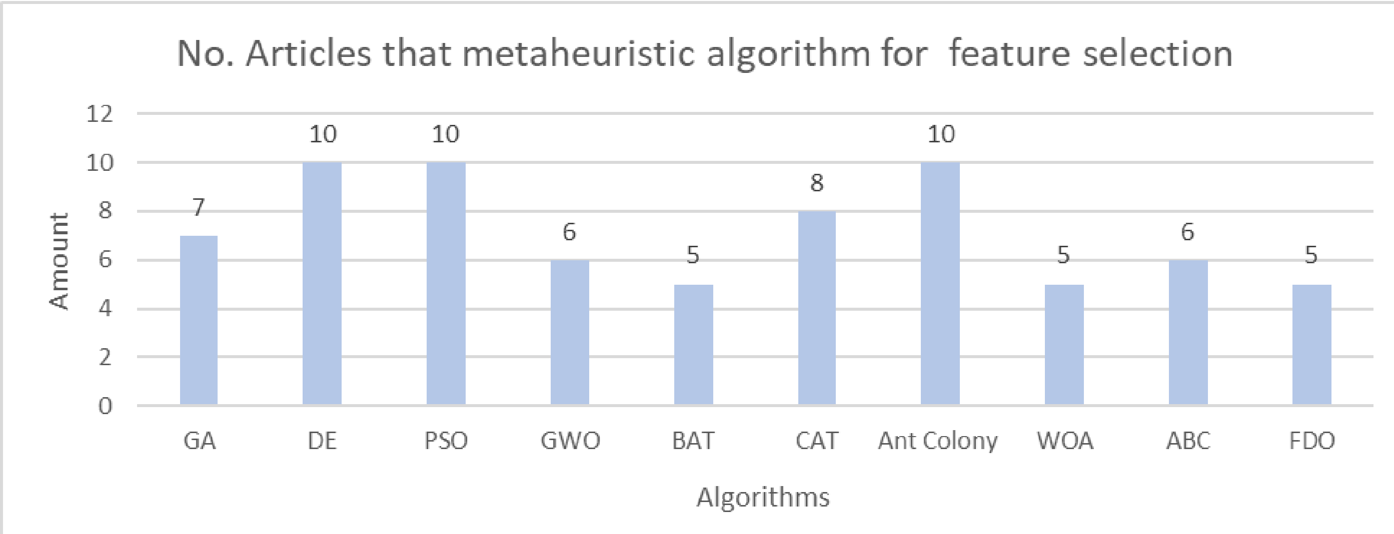

Figure 6: redundancy of metaheuristic algorithms used in feature selection

.

3. CONCLUSION

The feature selection approach plays an important role in reducing the complexity of any model-based machine learning algorithm and sometimes may lead to advance performance of the simulated model. For this purpose, many approaches have been developed and used in many applications of machine learning. In this study, a feature selection approach grounded on the optimization algorithm has been reviewed. Currently, Meta-heuristic algorithms have received more attention for feature selection compared to the other approaches like a filter, wrapper, and embedded, due to Meta-heuristic algorithms interact with the classifiers a lot. However, more computation requires more applying Meta-heuristic algorithms compare to the rest approaches.

REFERENCES

Abdullah, J. M., & Ahmed, T. (2019). Fitness dependent optimizer: inspired by the bee swarming reproductive process. IEEE Access, 7, 43473-43486.

https://doi.org/10.1109/ACCESS.2019.2907012Aghdam, M. H., Ghasem-Aghaee, N., & Basiri, M. E. (2009). Text feature selection using ant colony optimization. Expert Systems with Applications, 36(3 PART2), 6843–6853.

https://doi.org/10.1016/j.eswa.2008.08.022Ahmed, A.-A. (2005). Feature subset selection using ant colony optimization. International Journal of Computational, 2(1), 53–58. http://epress.lib.uts.edu.au/research/handle/10453/6002Ahmed, A. M., Rashid, T. A., & Saeed, S. A. M. (2020). Cat swarm optimization algorithm: a survey and performance evaluation. Computational intelligence and neuroscience, 2020.Bozorg-Haddad, O. (2018). Studies in Computational Intelligence - Advanced Optimization by Nature- Inspired Algorithms. In Studies in Computational Intelligence.

Chen, L. F., Su, C. T., & Chen, K. H. (2012). An improved particle swarm optimization for feature selection. Intelligent Data Analysis, 16(2), 167–182. https://doi.org/10.3233/IDA-2012-0517

Chen, Y., Miao, D., & Wang, R. (2010). A rough set approach to feature selection based on ant colony

optimization. Pattern Recognition Letters, 31(3), 226–

233. https://doi.org/10.1016/j.patrec.2009.10.013

Cremona, C., Yang, S., & Dieleman, L. (2009). Overview of feature extraction techniques. IOMAC 2009 - 3rd International Operational Modal Analysis Conference, September, 381–390.

https://doi.org/10.2139/ssrn.3568122

Dash, M., & Liu, H. (1997). Feature selection for classification. Intelligent Data Analysis, 1(3), 131–156. https://doi.org/10.3233/IDA-1997-1302

Dash, Manoranjan, & Liu, H. (2003). Consistency-based search in feature selection. Artificial Intelligence, 151(1– 2), 155–176. https://doi.org/10.1016/S0004-

3702(03)00079-1

Diao, R., & Shen, Q. (2015). Nature inspired feature selection meta-heuristics. Artificial Intelligence Review, 44(3), 311–340. https://doi.org/10.1007/s10462-015-

9428-8

Emary, E., Zawbaa, H. M., & Hassanien, A. E. (2016). Binary grey wolf optimization approaches for feature selection. Neurocomputing, 172, 371–381. https://doi.org/10.1016/j.neucom.2015.06.083

Faris, H., Aljarah, I., Al-Betar, M. A., & Mirjalili, S. (2018). Grey wolf optimizer: a review of recent variants and applications. Neural Computing and Applications, 30(2), 413–435. https://doi.org/10.1007/s00521-017-3272-5

Forsati, R., Moayedikia, A., Keikha, A., & Shamsfard, M. (2012). A Novel Approach for Feature Selection based on the Bee Colony Optimization. International Journal of Computer Applications, 43(8), 13–16. https://doi.org/10.5120/6122-8329

Guo, H., Jack, L. B., & Nandi, A. K. (2005). Feature generation using genetic programming with application to fault classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 35(1), 89–

99. https://doi.org/10.1109/TSMCB.2004.841426Hajnayeb, A., Ghasemloonia, A., Khadem, S. E., & Moradi,

M. H. (2011). Application and comparison of an ANNbased feature selection method and the genetic algorithm in gearbox fault diagnosis. Expert Systems with Applications, 38(8), 10205–10209. https://doi.org/10.1016/j.eswa.2011.02.065

Hancer, E., Xue, B., Zhang, M., Karaboga, D., & Akay, B. (2018). Pareto front feature selection based on artificial bee colony optimization. Information Sciences, 422, 462–479. https://doi.org/10.1016/j.ins.2017.09.028

Jain, I., Jain, V. K., & Jain, R. (2018). Correlation feature selection based improved-Binary Particle Swarm Optimization for gene selection and cancer classification. Applied Soft Computing, 62, 203–215. https://doi.org/10.1016/j.asoc.2017.09.038

Jeyasingh, S., & Veluchamy, M. (2017). Modified bat algorithm for feature selection with the Wisconsin Diagnosis Breast Cancer (WDBC) dataset. Asian Pacific Journal of Cancer Prevention, 18(5), 1257– 1264. https://doi.org/10.22034/APJCP.2017.18.5.1257

Jović, A., Brkić, K., & Bogunović, N. (2015). A review of feature selection methods with applications. 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings, 1200–

1205. https://doi.org/10.1109/MIPRO.2015.7160458

KARABOGA, D. (2005). AN IDEA BASED ON HONEY- BEE SWARM FOR NUMERICAL OPTIMIZATION.

Globalisasi Dan Kebidayaan Lokal: Suatu Dialektika Menuju Indonesia Baru, 7(2), 147–173. http://dx.doi.org/10.1016/j.

Karaboga, D., & Akay, B. (2009). A comparative study of Artificial Bee Colony algorithm. Applied Mathematics and Computation, 214(1), 108–132. https://doi.org/10.1016/j.amc.2009.03.090

Khalid, S., Khalil, T., & Nasreen, S. (2014). A survey of feature selection and feature extraction techniques in machine learning. Proceedings of 2014 Science and Information Conference, SAI 2014, August 2014, 372–

378. https://doi.org/10.1109/SAI.2014.6918213

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. Artificial Intelligence, 97(1–2), 273–

324. https://doi.org/10.1016/s0004-3702(97)00043-x

Lin, K. C., Huang, Y. H., Hung, J. C., & Lin, Y. T. (2015).

Feature Selection and Parameter Optimization of Support Vector Machines Based on Modified Cat Swarm Optimization. International Journal of Distributed Sensor Networks, 2015.

https://doi.org/10.1155/2015/365869

Lin, K. C., Zhang, K. Y., Huang, Y. H., Hung, J. C., & Yen,

N. (2016). Feature selection based on an improved cat swarm optimization algorithm for big data classification. Journal of Supercomputing, 72(8), 3210–3221. https://doi.org/10.1007/s11227-016-1631-0

Mafarja, M. M., & Mirjalili, S. (2017). Hybrid Whale Optimization Algorithm with simulated annealing for feature selection. Neurocomputing, 260,302–312. https://doi.org/10.1016/j.neucom.2017.04.053

Mafarja, M., & Mirjalili, S. (2018). Whale optimization approaches for wrapper feature selection. Applied Soft Computing, 62, 441–453. https://doi.org/10.1016/j.asoc.2017.11.006

Mirjalili, S., & Lewis, A. (2016). The Whale Optimization Algorithm. Advances in Engineering Software, 95, 51–

67. https://doi.org/10.1016/j.advengsoft.2016.01.008

Muhammed, D. A., Saeed, S. A. M., & Rashid, T. A. (2020). Improved Fitness-Dependent Optimizer Algorithm. IEEE Access, 8, 19074–19088. https://doi.org/10.1109/ACCESS.2020.2968064

Nakamura, R. Y. M., Pereira, L. A. M., Costa, K. A.,

Rodrigues, D., Papa, J. P., & Yang, X. S. (2012). BBA: A binary bat algorithm for feature selection. Brazilian Symposium of Computer Graphic and Image Processing, 291–297. https://doi.org/10.1109/SIBGRAPI.2012.47

Niu, P., Niu, S., liu, N., & Chang, L. (2019). The defect of the Grey Wolf optimization algorithm and its verification method. Knowledge-Based Systems, 171, 37–43. https://doi.org/10.1016/j.knosys.2019.01.018

Oscar Cordon, farncisco Herrera, T. (2002). 1-Cordon- Herrera-Stuetzle.

Rashid, T. A., Fattah, P., & Awla, D. K. (2018). Using accuracy measure for improving the training of LSTM with metaheuristic algorithms. Procedia Computer Science,

Sakri, S. B., Abdul Rashid, N. B., & Muhammad Zain, Z. (2018). Particle Swarm Optimization Feature Selection for Breast Cancer Recurrence Prediction. IEEE Access, 6, 29637–29647.

https://doi.org/10.1109/ACCESS.2018.2843443140, 324-333.

Sharawi, M., Zawbaa, H. M., Emary, E., & Hossamzawbaagmailcom, E. (n.d.). Feature Selection Approach Based on Whale Optimization Algorithm.

Sharma, A., Amarnath, M., & Kankar, P. K. (2016). Feature extraction and fault severity classification in ball bearings. JVC/Journal of Vibration and Control, 22(1), 176–192.

https://doi.org/10.1177/1077546314528021

Shrivastava, P., Shukla, A., Vepakomma, P., Bhansali, N., & Verma, K. (2017). A survey of nature-inspired algorithms for feature selection to identify Parkinson's disease. Computer Methods and Programs in Biomedicine, 139, 171–179.
https://doi.org/10.1016/j.cmpb.2016.07.029

Taha, A. M., Mustapha, A., & Chen, S. Der. (2013). Naive Bayes-guided bat algorithm for feature selection. The Scientific World Journal, 2013.

https://doi.org/10.1155/2013/325973

Van Der Maaten, L. J. P., Postma, E. O., & Van Den Herik,

H. J. (2009). Dimensionality Reduction: A Comparative Review. Journal of Machine Learning Research, 10, 1–

41. https://doi.org/10.1080/13506280444000102Vergara, J. R., & Estévez, P. A. (2014). A review of feature selection methods based on mutual information. Neural Computing and Applications, 24(1), 175–186. https://doi.org/10.1007/s00521-013-1368-0Yang, J., & Honavar, V. (1998). Feature subset selection using genetic algorithm. IEEE Intelligent Systems and Their Applications, 13(2), 44–48.

https://doi.org/10.1109/5254.671091

Yang, X. S. (2013). Bat algorithm: Literature review and applications. International Journal of Bio-Inspired Computation, 5(3), 141–149.
https://doi.org/10.1504/IJBIC.2013.055093

Yu, L., & Liu, H. (2003). Feature Selection for High- Dimensional Data: A Fast Correlation-Based Filter Solution. Proceedings, Twentieth International Conference on Machine Learning, 2, 856–863.

ZorarpacI, E., & Özel, S. A. (2016). A hybrid approach of differential evolution and artificial bee colony for feature selection. Expert Systems with Applications, 62, 91–103. https://doi.org/10.1016/j.eswa.2016.06.004

Abbas, D. K. et al. (2022) 'Using fitness dependent optimizer for training multi-layer perceptron', arXiv preprint arXiv:2201.00563.Abdulkhaleq, M. T. et al. (no date) 'Fitness Dependent Optimizer with Neural Networks for COVID-19 Patients'Abualigah, L. M., Khader, A. T. and Hanandeh, E. S. (2018) 'A new feature selection method to improve the document clustering using particle swarm optimization algorithm', Journal of Computational Science, 25, pp. 456–466.

Aghdam, M. H., Ghasem-Aghaee, N. and Basiri, M. E. (2009) 'Text feature selection using ant colony optimization', Expert systems with applications, 36(3),

pp. 6843–6853.Aghdam, M. H. and Heidari, S. (2015) 'Feature selection using particle swarm optimization in text categorization', Journal of Artificial Intelligence and Soft Computing Research, 5.

Aghdam, M. H., Kabiri, P. and others (2016) 'Feature selection for intrusion detection system using ant colony optimization.', Int. J. Netw. Secur., 18(3), pp. 420–432.Ahdesmäki, M. and Strimmer, K. (2010) 'Feature selection

in omics prediction problems using cat scores and false nondiscovery rate control', The Annals of Applied Statistics, 4(1), pp. 503–519.

Ahmad, I. (2015) 'Feature selection using particle swarm optimization in intrusion detection', International Journal of Distributed Sensor NetworksAl-Ani, A. (2005) 'Feature subset selection using ant colony optimization', International journal of computational intelligence.Al-Ani, A., Alsukker, A. and

Khushaba, R. N. (2013) ‘Feature subset selection using differential evolution and a wheel based search strategy’, Swarm and Evolutionary Computation, 9, pp. 15–26.Alarifi, A. et al. (2020) ‘A big data approach to sentiment analysis using greedy feature selection with cat swarm optimization-based long short-term memory neural networks’, The Journal of Supercomputing, 76(6), pp. 4414–4429.Almarzouki, H. Z. (2022) ‘Deep-Learning-Based Cancer Profiles Classification Using Gene Expression Data Profile’, Journal of Healthcare Engineering, 2022.

Arslan, S. and Ozturk, C. (2019) ‘Multi hive artificial bee colony programming for high dimensional symbolic regression with feature selection’, Applied Soft Computing, 78, pp. 515–527.

Bansal, P. et al. (2020) ‘A hybrid grasshopper and new cat swarm optimization algorithm for feature selection and optimization of multi-layer perceptron’, Soft computing, 24(20), pp. 15463–15489.

Bhattacharyya, S. et al. (2014) ‘Automatic feature selection of motor imagery EEG signals using differential evolution and learning automata’, Medical \& biological engineering \& computing, 52(2), pp. 131– 139.

Brezočnik, L. (2017) ‘Feature selection for classification using particle swarm optimization’, in IEEE EUROCON 2017-17th International Conference on Smart Technologies, pp. 966–971.

Bui, Q.-T. et al. (2019) ‘Whale Optimization Algorithm and Adaptive Neuro-Fuzzy Inference System: a hybrid method for feature selection and land pattern classification’, International journal of remote sensing, 40(13), pp. 5078–5093.

Chiu, P. C. et al. (2021) ‘Hybrid Sine Cosine and Fitness Dependent Optimizer for Global Optimization’, IEEE Access, 9, pp. 128601–128622.

Dadaneh, B. Z., Markid, H. Y. and Zakerolhosseini, A. (2016) ‘Unsupervised probabilistic feature selection using ant colony optimization’, Expert Systems with Applications, 53, pp. 27–42.

Deriche, M. (2009) ‘Feature selection using ant colony optimization’, in 2009 6th International Multi- Conference on Systems, Signals and Devices, pp. 1–4.

Dixit, A., Mani, A. and Bansal, R. (2020) ‘Feature selection for text and image data using differential evolution with SVM and Na{\"\i}ve Bayes classifiers’, Engineering Journal, 24(5), pp. 161–172.

Emary, E. et al. (2015) ‘Multi-objective gray-wolf optimization for attribute reduction’, Procedia Computer Science, 65, pp. 623–632.

Fallahzadeh, O., Dehghani-Bidgoli, Z. and Assarian, M. (2018) ‘Raman spectral feature selection using ant colony optimization for breast cancer diagnosis’, Lasers in medical science, 33(8), pp. 1799–1806.

Gao, X.-Z. et al. (2021) ‘Multi-objective optimization of feature selection using hybrid cat swarm optimization’, Science China Technological Sciences, 64(3), pp. 508– 520.

Gomathy, M. (2021) ‘Optimal feature selection for speech emotion recognition using enhanced cat swarm optimization algorithm’, International Journal of Speech Technology, 24(1), pp. 155–163.

Guha, R. et al. (2020) ‘Embedded chaotic whale survival algorithm for filter--wrapper feature selection’, Soft Computing, 24(17), pp. 12821–12843.

Hancer, E. et al. (2018) ‘Pareto front feature selection based on artificial bee colony optimization’, Information Sciences, 422, pp. 462–479.

Hancer, E. (2020) ‘New filter approaches for feature selection using differential evolution and fuzzy rough set theory’, Neural Computing and Applications, 32(7),

pp. 2929–2944.

Hancer, E., Xue, B. and Zhang, M. (2022) ‘Fuzzy filter cost-sensitive feature selection with differential evolution’, Knowledge-Based Systems, 241, p. 108259.

Jayaprakash, A. and KeziSelvaVijila, C. (2019) ‘Feature selection using ant colony optimization (ACO) and road sign detection and recognition (RSDR) system’, Cognitive Systems

Research, 58, pp. 123–133.

Jeyasingh, S. and Veluchamy, M. (2017) ‘Modified bat algorithm for feature selection with the wisconsin diagnosis breast cancer (WDBC) dataset’, Asian Pacific journal of cancer prevention: APJCP, 18(5), p. 1257.

Johari, P. K. and Gupta, R. K. (2021) ‘Improved Feature Selection Techniques for Image Retrieval based on Metaheuristic Optimization’, International Journal of Computer Science \& Network Security, 21(1), pp. 40– 48.

Kabir, M. et al. (2009) ‘An efficient feature selection using ant colony optimization algorithm’, in International Conference on Neural Information Processing, pp. 242– 252.

Kanan, H. R., Faez, K. and Taheri, S. M. (2007) ‘Feature selection using ant colony optimization (ACO): a new method and comparative study in the application of face recognition system’, in Industrial conference on data mining, pp. 63–76.

Khushaba, R. N., Al-Ani, A. and Al-Jumaily, A. (2008) ‘Feature subset selection using differential evolution’, in International Conference on Neural Information Processing, pp. 103–110

Khushaba, R. N., Al-Ani, A. and Al-Jumaily, A. (2011) ‘Feature subset selection using differential evolution

and a statistical repair mechanism’, Expert Systems with Applications, 38(9), pp. 11515–11526.

Kitonyi, P. M. and Segera, D. R. (2021) ‘Hybrid gradient descent grey wolf optimizer for optimal feature selection’, BioMed Research International, 2021.

Kumar Gupta, D. et al. (2015) ‘Pso-asent: Feature selection using particle swarm optimization for aspect based sentiment analysis’, in International conference on applications of natural language to information systems,

pp. 220–233.

Li, Q. et al. (2017) ‘An enhanced grey wolf optimization based feature selection wrapped kernel extreme learning machine for medical diagnosis’, Computational and mathematical methods in medicine, 2017.

Lin, K.-C. et al. (2015) ‘Feature selection and parameter optimization of support vector machines based on modified cat swarm optimization’, International Journal of Distributed Sensor Networks, 11(7), p. 365869.

Lin, K.-C. et al. (2016) ‘Feature selection based on an improved cat swarm optimization algorithm for big data classification’, The Journal of Supercomputing, 72(8),

pp. 3210–3221.

Lin, K. C., Zhang, K. Y. and Hung, J. C. (2014) ‘Feature selection of support vector machine based on harmonious cat swarm optimization’, in 2014 7th international conference on ubi-media computing and workshops, pp. 205–208.

Liu, Y. et al. (2011) ‘An improved particle swarm optimization for feature selection’, Journal of Bionic Engineering, 8(2), pp. 191–200.

Mandal, M. et al. (2021) ‘A tri-stage wrapper-filter feature selection framework for disease classification’, Sensors, 21(16), p. 5571.

Moradi, P. and Rostami, M. (2015) ‘Integration of graph clustering with ant colony optimization for feature selection’, Knowledge-Based Systems, 84, pp. 144–161.Nakamura, R. Y. M. et al. (2012) ‘BBA: a binary bat

algorithm for feature selection’, in 2012 25th SIBGRAPI conference on graphics, patterns and images, pp. 291–297.

Nematzadeh, H. et al. (2019) ‘Frequency based feature selection method using whale algorithm’, Genomics, 111(6), pp. 1946–1955.

Nirmala Sreedharan, N. P. et al. (2018) ‘Grey wolf optimisation-based feature selection and classification for facial emotion recognition’, IET Biometrics, 7(5),

pp. 490–499.

Qasim, O. S. and Algamal, Z. Y. (2018) ‘Feature selection using particle swarm optimization-based logistic regression model’, Chemometrics and Intelligent Laboratory Systems, 182, pp. 41–46.

Sahu, B. and Mishra, D. (2012) ‘A novel feature selection algorithm using particle swarm optimization for cancer microarray data’, Procedia Engineering, 38, pp. 27–31.

Saleem, N., Zafar, K. and Sabzwari, A. F. (2019) ‘Enhanced feature subset selection using Niche based bat algorithm’, Computation, 7(3), p. 49.

Salih, J. F., Mohammed, H. M. and Abdul, Z. K. (2022) ‘Modified Fitness Dependent Optimizer for Solving Numerical Optimization Functions’, IEEE Access, 10,

pp. 83916–83930. doi: 10.1109/ACCESS.2022.3197290.

Tawhid, M. A. and Dsouza, K. B. (2018) ‘Hybrid binary bat enhanced particle swarm optimization algorithm for solving feature selection problems’, Applied Computing and Informatics.

Too, J. et al. (2018) ‘A new competitive binary grey wolf optimizer to solve the feature selection problem in EMG signals classification’, Computers, 7(4), p. 58.

Too, J., Mafarja, M. and Mirjalili, S. (2021) ‘Spatial bound whale optimization algorithm: an efficient high- dimensional feature selection approach’, Neural Computing and Applications, 33(23), pp. 16229–16250.

Vivekanandan, T. and Iyengar, N. C. S. N. (2017) ‘Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease’, Computers in biology and medicine, 90, pp. 125–136.

Wang, X. et al. (2020) ‘Multi-objective feature selection based on artificial bee colony: An acceleration approach with variable sample size’, Applied Soft Computing, 88,

p. 106041.

Xue, B., Zhang, M. and Browne, W. N. (2012) ‘Particle swarm optimization for feature selection in classification: A multi-objective approach’, IEEE transactions on cybernetics, 43(6), pp. 1656–1671.

Yang, X. S. (2013) ‘Na{\"\i}ve Bayes Guided Bat algorithm for featureselection’,

TheScientificWorldJournalVolume20, 13.

Zahran, B. M. and Kanaan, G. (2009) ‘Text feature selection using particle swarm optimization algorithm 1’.

Zhang, Y. et al. (2019) ‘Cost-sensitive feature selection using two-archive multi-objective artificial bee colony algorithm’, Expert Systems with Applications, 137, pp. 46–58.

Zhang, Y. et al. (2020) ‘Binary differential evolution with self-learning for multi-objective feature selection’, Information Sciences, 507, pp. 67–85.

Zheng, Y. et al. (2018) ‘A novel hybrid algorithm for feature selection based on whale optimization algorithm’, IEEE Access, 7, pp. 14908–14923.

Zorarpac\i, E. and Özel, S. A. (2016) ‘A hybrid approach of differential evolution and artificial bee colony for feature selection’, Expert Systems with Applications, 62, pp. 91–10