Journal of Computer Science
https://dergipark.org.tr/en/pub/bbd
Anatolian Science

ISSN,e-ISSN: 2548-1304
Volume:IDAP-2023, pp:58-66, 2023
https://doi.org/10.53070/bbd.1351629
Research Paper

Evaluation of Performance of Feature Selection of Meta-Heuristic Optimization Methods in Medical Data

Hüseyin GÜNDOĞDU[1], Osman ALTAY[2],

[1] Yazılım Mühendisliği Bölümü, Celal Bayar Üniversitesi, Manisa, Türkiye

[2] Yazılım Mühendisliği Bölümü, Celal Bayar Üniversitesi, Manisa, Türkiye

(221201016@ogr.cbu.edu.tr, osman.altay@cbu.edu.tr)

Received:Aug.29,2023 Accepted:Oct.16,2023 Published:Oct.18,2023

Abstract—In knowledge discovery, the processes of applying data cleaning, data integration, data selection-transformation, and data mining methods and obtaining meaningful information from the obtained patterns are performed, respectively. In recent years, it has become quite common to use metaheuristic optimization methods in the data selection phase. In this study, the nearest neighbor algorithm, support vector machine, and decision tree algorithms from machine learning algorithms were used on health data obtained from the University of California, Irvine. The whale optimization algorithm, salp swarm optimization, slime mould optimization, particle swarm optimization, and Harris Hawks optimization methods were used for feature selection. The obtained results were compared in detail.

Keywords: Feature selection, meta-heuristic optimization algorithms, Whale optimization algorithm, Harris Hawks optimization

1. Introduction

With rapidly developing information processing and data storage technologies in recent years, the amount of data in our world and the rate of increase in data are increasing rapidly. Since advanced computers and electronic devices make the data storage process extremely easy, many data items that human beings did not benefit from keeping in the past are stored today. Because there are electronic devices everywhere, the knowledge gleaned from historical data by data mining techniques influences our future decisions and alters our lives (Suparyanto dan Rosad (2015, 2020).

Machine learning is programming computers to optimize a performance criterion by evaluating data or experience (Castellanos-garzón et al., 2019). Computer operations that optimize the parameters of the model we train with experience and training data using computers are defined as learning. A model can be predictive to make predictions in the future, descriptive to learn from data, or both (Baştanlar & Ozuysal, 2014).

With the development of technology, in many cases in the field of health, it is tried to produce solutions to various health problems by using machine learning algorithms (Castellanos-garzón et al., 2019; Kononenko, 2001; Magoulas & Prentza, 2001; Shailaja & Scholar, 2018). Machine learning has been used in many cases where it is difficult or impossible to use classical methods to produce solutions to any health problem. In the literature; Machine learning algorithms are used in different health applications such as cardiovascular diseases (Iqbal et al., 2022), urinary system infection (Taylor et al., 2018), arrhythmia detection (Zhang et al., 2022), investigation of suicidal behaviors (Nordin et al., 2022), anomaly detection (Yuan et al., 2023), early diagnosis of gestational diabetes (Xiong et al., 2022). Machine learning methods used in the health sector provide decision support to doctors in the diagnosis and treatment of diseases.

Machine learning approaches are examined under two headings: supervised learning, in which the model is trained with data with a certain output, and unsupervised learning, in which we divide the data with unknown labels into groups. While supervised learning is used to train the model with the data of diagnosed patients and make predictions about the diagnosis of new patients, banks use the unsupervised learning approach when they want to divide their customers into certain clusters according to their credit risk status (Shailaja & Scholar, 2018).

Feature selection is used by machine learning algorithms to work more efficiently and improve the results obtained. Feature selection is a preprocessing methodology that aims to identify the most relevant characteristics of a given dataset. It has traditionally been applied to a wide variety of problems, including biological data processing, market analysis, finance, and intrusion detection. Feature selection has also been used successfully in medical applications (Nagarajan et al., 2021; Selvakuberan et al., 2011).

In feature selection, it is mainly tried to determine which variable is effective on the result and to what extent. Feature selection has become a necessity in most problems, as the datasets have become very large recently. Although many feature selection methods have been used since the 1970s, meta-heuristic feature selection algorithms have gained superiority over classical methods for the last 20 years (Dokeroglu et al., 2022).

In this study, after the introduction, the machine learning algorithm is defined in the second section, three of the most used methods in machine learning are explained in detail, third section the meta-heuristic optimization methods to be applied are explained. In the fourth section, the medical datasets used are introduced, and the performances of the meta-heuristic optimization methods used are compared. In the fifth section, the conclusion is presented. This article aims to evaluate the performance of meta-heuristic optimization methods and machine learning methods on medical data.

2 Machine Learning Algorithms

Artificial intelligence, which is a part of computer science, aims to make computers more intelligent. One of the most basic requirements of intelligence is learning. Most researchers today deny the existence of a non-learning intelligence. Therefore, machine learning, one of the main components of artificial intelligence, is developing rapidly. (Kononenko, 2001). In the next part, the K-Nearest neighbor (KNN), decision trees (DT), and support vector machines (SVM) from machine learning algorithms used in this study are to be explained.

2.1. K-Nearest Neighbor Algorithms

The KNN first appeared in the 1950s [26]. The algorithm's labor-intensive nature, particularly when applied to large training sets, hindered its widespread use until the 1960s, when advancements in computer power made it more feasible. It has since been widely used in classification problems [27]. KNN is one of the most basic supervised learning approaches. This algorithm is predicted by looking at the distance to which class the sample tested belongs, using data in the dataset, to which class it belongs.

2.2. Decision Tree

DT starts with a root node and uses the tree structure to analyze data by applying a set of rules to that node. Each rule is associated with a threshold on a feature selection and splits the data in a particular direction. This process creates the next node and branch of the tree. The node that contributes the most to the result is the root node. When the decision tree reaches the leaf node, each leaf contains a class label or predictive value (Mitchell, n.d.).

In the late 1970s and early 1980s, machine learning researcher J. Ross Quinlan developed a decision tree algorithm known as ID3 (Recursive Dichotomiser). This work extends previous work on concept learning systems described by E. B. Hunt, J. Marin, and P. T. Stone.

2.3. Support Vector Machine

The SVM was introduced by Vapnik and has been successfully implemented in classification and regression problems. SVM uses structural risk minimization (SRM), a statistically based approach.

SRM, minimizes the error of an upper bound on the expected risk in the training data (Altay & Tezi, 2020).

Data stacks that cannot be separated linearly in low dimensions with the help of support vector machines are moved to higher dimensions and then separated with the help of a plane. When determining this plane, the place that is farthest from the elements of the two divided parts should be taken. Thus, it is tried to prevent the problem of overfitting. (Cervantes et al., 2020). It can classify datasets that can or cannot be separated linearly using SVM. The kernel function is used for non-linearly separated datasets (Brereton & Lloyd, 2010).

3. Meta-Heuristic Optimization Methods

The applications used today are constantly producing more and more data in terms of both the number of samples and features. The rapid increase in the concept of big data reduces the processing power and speed of machine learning algorithms. Feature selection is one of these data preprocessing processes, in which we remove noisy and unnecessary data. (Dokeroglu et al., 2022). Inspired by nature, meta-heuristic methods have been developed for feature selection, without focusing on a specific local solution and constantly seeking the best solution in a chaotic fashion.

3.1 Particle Swarm Optimization

Particle swarm optimization (PSO) is basically based on two main methodologies. The first is its ties to artificial life in general, and to bird-fish swarm and flock theory in particular. It is also associated with evolutionary computing and has ties to both genetic algorithms and evolutionary programming. PSO algorithm is based on a straightforward principle and may be easily implemented using few lines of computer code. Requires only primitive mathematical operators and is inexpensive in terms of memory requirements, speed, and computation (James Kennedy and Russell E, 2011).

Working principle of PSO; one of the well-known meta-heuristic optimization methods, simulates the behavior of bird flocks. Every particle (every element in the swarm) has two properties expressed by position and velocity vectors. PSO is influenced by the best personal and the best global position when performing the movement of particles. PSO method is described in detail in the corresponding reference(Xue et al., 2013).

3.2 Whale Optimization Algorithm

Whale Optimization (WOA) is a meta-heuristic optimization method inspired by the hunting strategies of humpback whales in nature. This method was proposed by Seyedali Mirjalili in 2016 (Mirjalili & Lewis, 2016). WOA basically consists of three stages:

Encircling prey: Humpback whales can locate prey and circle their prey. Since the location of the optimum design in the search space is not known in advance, the WOA algorithm assumes that the best available candidate solution is the target hunt or is close to the optimum. Once the best solution has been identified, other search agents will try to update their positions towards the best search agent.

Bubble-net attacking Phase: This behavior is modeled in two parts: shrinking the predator chain around the prey and circular motion.

Search for prey (Update) Phase: The new locations of each humpback whale are determined around a randomly selected humpback whale instead of the best known spot for a global resolution (TANYILDIZI & CİGALI, 2017).

WOA method has advantages over other meta-heuristic optimization methods such as effective discovery, fast convergence and suitability for various problems (Gharehchopogh & Gholizadeh, 2019). The relevant citation can be examined for details(Zheng et al., 2019).

3.3. Slime Mould Algorithm

Slime Mould algorithm (SMA) method was first proposed by Mirjalili in 2016 (Mirjalili & Lewis, 2016). This optimization method was inspired by a large single-celled cell with intriguingly intelligent behavior called 'Physarum polycephalum' (Chen et al., 2023). These organisms are known for the web-like structures they form to reach their food source by the shortest route. Slime molds respond to chemical stimuli in their environment using their ability to grow and move. These responses are determined by factors such as distance from their source and density.(ALTAY & VAROL ALTAY, 2022).

There are 4 stages in the slime mold optimization method: Approaching Food, Containing Food, Oscillation and Analysis of Complexity. SMA method is described in detail in the corresponding reference (Li et al., 2020).

3.4. Salp Swarm Optimization

Salp Swarm optimization (SSA) was proposed by Mirjalili in 2017 (Mirjalili et al., 2017). This optimization method was inspired by the feeding and wayfinding behaviors of salp creatures living in the oceans. Salp herds consist of a leader and other salps following him. Salp in the center of the herd leads the herd towards food as the leader. SSA method is described in detail in the corresponding reference (Abualigah et al., 2020; Faris et al., 2018; Hegazy et al., 2020). The position of the leader in salp flock optimization changes according to the food. Other creatures in the herd determine their positions relative to the leader.

3.5. Harris Hawks Optimization

Harris hawks optimization (HHO) was first introduced by Heidari in 2019 (Heidari et al., 2019). This optimization method was developed by being inspired by the hunting methods of Harris hawks. Harris hawks can track and detect their prey with their powerful eyes, but sometimes prey is not easily seen. Therefore, hawks wait, observe, and perhaps watch the desert area a few hours later to spot prey. Two different equations have been developed for two different exploration strategies. The first is the positioning of the hawks relative to the other members of the family, and the other is the positioning of the tall trees in the region. The HHO method is described in detail in the corresponding reference. (Peng et al., 2023) .

4. Results and Discussion

In this study, 3 different machine learning algorithms and 5 different meta-heuristic optimization methods for feature selection were used on 2 different data sets.

4.1 Data sets description

The first of the datasets used was Heart, organized for a study conducted at Pakistan State University in 2015. (Chicco & Jurman, 2020). It includes data on 299 patients, 105 female and 194 male patients. All the patients are over 40 years old and have heart diseases.

The other dataset used in the study was created to evaluate the performance of the Lee Silverman Voice Treatmen (LSVT) computer program developed for the treatment of voice distortions, a symptom of Parkinson's disease (Tsanas et al., 2014). The data of 14 patients (8 men and 6 women) was used in the study. 309 different features of 126 different sounds (for example) were recorded.

4.2 Experimental setup

The two data sets used in the study were primarily subjected to a feature selection process with PSO, SMA, WOA, SSA, and HHO methods. The feature-selected data sets are tested with KNN, DT, and SVM algorithms. 80 percent of the data sets were used for training and 20 percent for testing. The training and test datasets were randomly selected. In order to compare the algorithms under equal conditions, the same training and test data were used for all methods. Since there are random coefficients in meta-heuristic optimization methods, the results obtained by testing each optimization method and machine learning algorithm 10 times have been recorded. All experiments were performed on the Matlab 2021a platform licensed by Manisa Celal Bayar University on a computer with the Windows 11 operating system, 16 GB of RAM, and a CPU of the Intel (R) Core i7-1165G7 (2.8 GHz).

For the KNN algorithm, the k value is taken as 1. For the meta-heuristic optimization methods, the number of iterations was determined as 100 and the population as 50. In PSO, the coefficients$c1$ and $c2$ are used as 2 and weight (w) as 1. In the HHO method, the beta value was determined as 1.5. The z value in SMA is taken as 0.03. $a1 = [2,0]$; $a2 = [-2,-1]$ in WOA and b value is taken as 1.

4.3 Experimental results

The results of the tests performed on the Heart and LSVT datasets are the mean value, standard deviation, maximum, and minimum values of the fitness functions in Table 1. The classification metrics and the average of the selected feature numbers are presented in Table 2.

Table 1. Result of Fitness Function Meta-Heuristic Optimization Methods

Dataset	Machine Learning Algorithm	Meta-Heuristic Optimization	Mean	Standart Deviation	Max.	Min.
Heart	KNN	PSO	0,1356	0,0000	0,1356	0,1356
		SMA	0,1525	0,0240	0,2034	0,1356
		WOA	0,1339	0,0091	0,1525	0,1186
		SSA	0,1373	0,0051	0,1525	0,1356
		HHO	**0,1271**	0,0085	0,1356	0,1186
	DT	**PSO**	**0,1017**	0,0000	0,1017	0,1017
		SMA	0,1407	0,0252	0,1864	0,1017
		WOA	0,1068	0,0078	0,1186	0,1017
		SSA	0,1051	0,0068	0,1186	0,1017
		HHO	**0,1017**	0,0000	0,1017	0,1017
	SVM	PSO	0,1864	0,0000	0,1864	0,1864
		SMA	0,2017	0,0169	0,2373	0,1864
		WOA	0,1864	0,0000	0,1864	0,1864
		SSA	0,1898	0,0071	0,2034	0,1864
		HHO	**0,1864**	0,0000	0,1864	0,1864
LSVT	KNN	PSO	0,2400	0,0000	0,2400	0,2400
		SMA	**0,1560**	0,0580	0,2400	0,0800
		WOA	**0,1560**	0,0440	0,2400	0,1200
		SSA	0,2400	0,0000	0,2400	0,2400
		HHO	0,2400	0,0000	0,2400	0,2400
	DT	PSO	0,3040	0,0909	0,4000	0,2000
		SMA	0,2040	0,0479	0,2800	0,1200
		WOA	**0,1369**	0,0337	0,2000	0,0800
		SSA	0,4640	0,0430	0,5200	0,4000
		HHO	0,3720	0,0501	0,4000	0,2800
	SVM	PSO	0,2240	0,0043	0,3200	0,1600
		SMA	0,1600	0,0377	0,2000	0,0800
		WOA	**0,1360**	0,0470	0,2000	0,0400
		SSA	0,2880	0,0454	0,3200	0,2000
		HHO	0,2320	0,0169	0,2400	0,2000

As a result of the tests performed on the Heart data set, the average fitness function value was determined to be 0.1271 in the KNN for the HHO method, 0.1017 in the DT, and 0.1864 in the SVM. When Table 1 is examined, it is seen that HHO has a better fitness function average than other meta-heuristic optimization methods in all 3 machine learning algorithms.

As a result of the tests performed on the LSVT dataset, the average of the fitness function values was determined as 0.1560 in the KNN for the slime mold optimization method and the whale optimization method, 0.1369 in the WOA method DT, and 0.1360 in the SVM. It is seen that the WOA has a better average than other meta-heuristic optimization methods in all 3 machine learning algorithms in terms of fitness function values.

Table 2. Evaluation of Classification Metrics

Dataset	Machine Learning Algorithm	Meta-Heuristic Optimization	Accuracy	Sensitivity	Specificity	Precision	F-Measure	G-Mean	Number Of Feature
Heart	KNN	**PSO**	**0,8814**	0,9500	0,7368	0,8837	0,9157	0,8366	5,4
		SMA	0,8644	0,9500	0,7368	0,8809	0,9047	0,8256	5,2
		WOA	0,8813	0,9750	0,7368	0,8837	0,9156	0,8366	4,9
		SSA	0,8647	0,9750	0.7368	0,8809	0,9024	0,8255	5,6
		HHO	0,8813	0,9500	0.7368	0,8837	0,9157	0,8366	4,6
	DT	PSO	0,7796	0,8250	0.8305	0,8648	0,8354	0.7677	4,9
		SMA	**0.7945**	0.8350	0.7125	0.8559	0.8463	0.7691	5,2
		WOA	0.7830	0,8075	0,7315	0,8641	0,8345	0,7680	5,6
		SSA	0,7796	0,8100	0,7157	0,8578	0,8328	0,7607	5,1
		HHO	0,7796	0,8150	0,7052	0,8536	0,8337	0,7578	4,9
	SVM	PSO	0,7796	0,9375	0,4473	0,7817	0,8523	0,6458	4,3
		SMA	0,7677	0,9475	0,3894	0,7683	0,8474	0,5712	5,4
		WOA	0,7779	0,9300	0,4578	0,7834	0,8503	0,6514	3,9
		SSA	**0,7915**	0,9350	0,4894	0,7942	0,8587	0,6760	5,1
		HHO	0,7796	0,9350	0,4526	0,7829	0,8520	0,6488	5,2
LSVT	KNN	PSO	0,7600	0,6250	0,8235	0,6250	0,6250	0,7174	151,8
		SMA	**0,8440**	0,7250	0,9000	0,7933	0,7502	0,8049	1,5
		WOA	**0,8440**	0,7000	0,9117	0,8079	0,7361	0,7925	5,4
		SSA	0,7600	0,6250	0,8235	0,6250	0,6250	0,7174	154,8
		HHO	0,7600	0,6250	0,8235	0,6250	0,6250	0,7134	143,3
	DT	PSO	0,7360	0,6000	0,8125	0,6378	0,6169	0,6953	147,3
		SMA	0,7600	0,6333	0,8312	0,7306	0,6614	0,7167	2,3
		WOA	0,7360	0,6666	0,7750	0,6255	0,6426	0,7145	9,0
		SSA	**0,7680**	0,6555	0,8312	0,7019	0,6729	0,7366	152,4
		HHO	0,7480	0,6555	0,8000	0,6545	0,6518	0,7217	119,5
	SVM	PSO	0,6800	0,0000	1.0	NAN	NAN	0,0000	147,9
		SMA	**0,6920**	0,2750	0,8883	NAN	NAN	0,2634	2,0
		WOA	0,5840	0,3875	0,6764	NAN	NAN	0,3389	11,5
		SSA	0,6800	0,0000	1,0000	NAN	NAN	0,0000	154,8
		HHO	0,6800	0,0000	1,0000	NAN	NAN	0,0000	139,7

Selected data sets obtained as a result of feature selection using meta-heuristic optimization methods on the Heart dataset were tested with various machine learning algorithms. According to the test results, the highest accuracy value in KNN was obtained from PSO, and it was determined to be 0.8814. The highest accuracy value in the DT was obtained from the SMA and was determined to be 0.7945. In the SVM, the highest accuracy value was found to be 0.7915 in the SSA.

In the LSVT data set, according to the test results, the highest accuracy value in KNN was obtained from SMA and was determined to be 0.84. The highest accuracy value in the DT was obtained from the SSA and was determined to be 0.7670. In the SVM, the highest accuracy value was determined as 0.6800 in the SMA.

5.Conclusion

Today, the size and growth rate of data sets are constantly increasing, which increases the processing power required by machine learning algorithms. Feature selection is used to solve problems between machine learning and large data sets. In this study, the performance of meta-heuristic feature selection methods inspired by nature was evaluated on medical data. The KNN algorithm, DT, and SVM from machine learning methods are explained in detail and used in the study. In this context, in this study conducted for feature selection of medical data, the HHO method gave better results in datasets with a relatively low number of features, such as Heart, and the WOA in datasets with a large number of features, such as LSVT. When evaluated in terms of classification metrics in our study, it is seen that SMA and SSA methods perform better in medical data. However, it should be noted that a different method may yield better results depending on the data set and the nature of the problem.

6. References

Abualigah, L., Shehab, M., Alshinwan, M., & Alabool, H. (2020). Salp swarm algorithm: a comprehensive survey. *Neural Computing and Applications*, *32*(15), 11195–11215.

Altay, O., & Tezi, D. (2020). *KENDİLİĞİNDEN YerleşenÇeliLifli Beton PerformansiniTahmin EdeceSisteminModellenmesi*.

Altay, O., & Varol Altay, E. (2022). Investigation of Slime Mould Algorithm and Hybrid Slime Mould Algorithms' Performance in Global Optimization Problems. *DÜMF Mühendislik Dergisi*, *4*, 661–671.

Baştanlar, Y., & Ozuysal, M. (2014). Introduction to Machine Learning Second Edition. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 1107).

Brereton, R. G., & Lloyd, G. R. (2010). Support Vector Machines for classification and regression. *Analyst*, *135*(2), 230–267.

Castellanos-garzón, J. A., Costa, E., Luis, J., Jaimes, S., & Corchado, J. M. (2019). Knowledge-Based Systems An evolutionary framework for machine learning applied to medical. *Knowledge-Based Systems*, *185*, 104982.

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, *408*, 189–215.

Chen, H., Li, C., Mafarja, M., Heidari, A. A., Chen, Y., & Cai, Z. (2023). Slime mould algorithm: a comprehensive review of recent variants and applications. *International Journal of Systems Science*, *54*(1), 204–235.

Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Medical Informatics and Decision Making*, *20*(1), 1–16.

Dokeroglu, T., Deniz, A., & Kiziloz, H. E. (2022). A comprehensive survey on recent metaheuristics for feature selection. *Neurocomputing*, *494*, 269–296.

Faris, H., Mafarja, M. M., Heidari, A. A., Aljarah, I., Al-Zoubi, A. M., Mirjalili, S., & Fujita, H. (2018). An efficient binary Salp Swarm Algorithm with crossover scheme for feature selection problems. *Knowledge-Based Systems*, *154*(March), 43–67.

Gama, J. (2004). Functional trees. *Machine Learning*, *55*(3), 219–250.

Gharehchopogh, F. S., & Gholizadeh, H. (2019). A comprehensive survey: Whale Optimization Algorithm and its applications. *Swarm and Evolutionary Computation*, *48*(March), 1–24.

Hegazy, A. E., Makhlouf, M. A., & El-Tawel, G. S. (2020). Improved salp swarm algorithm for feature selection. *Journal of King Saud University - Computer and Information Sciences*, *32*(3), 335–344.

Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawks optimization: Algorithm and applications. *Future Generation Computer Systems*, *97*, 849–872.

Iqbal, M., Setiawan, M. N., Irawan, M. I., Khalif, K. M. N. K., Muhammad, N., & Aziz, M. K. B. M. (2022). Cardiovascular disease detection from high utility rare rule mining. *Artificial Intelligence* *in Medicine*, *131*.

J R Quinlan. (1986). Induction of Decision Trees. *Machine Learning*, *1*(1), 81–106.

James Kennedy and Russell E. (2011). Particle Swarm Optimization. *The Industrial Electronics Handbook - Five Volume Set*, 1942–1948.

Kononenko, I. (2001). Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine*, *23*(1), 89–109.

Li, S., Chen, H., Wang, M., Heidari, A. A., & Mirjalili, S. (2020). Slime mould algorithm: A new method for stochastic optimization. *Future Generation Computer Systems*, *111*, 300–323.

Magoulas, G. D., & Prentza, A. (2001). *Machine Learning in Medical Applications*. 300–307.

Mirjalili, S., Gandomi, A. H., Zahra, S., & Saremi, S. (2017). *Advances in Engineering Software Salp Swarm Algorithm : A bio-inspired optimizer for engineering design problems*. *114*, 163–191.

Mirjalili, S., & Lewis, A. (2016). *Advances in Engineering Software The Whale Optimization Algorithm*. *95*, 51–67.

Mitchell, T. M. (n.d.). *Machine Learning*.

Nagarajan, S. M., Muthukumaran, V., Murugesan, R., Joseph, R. B., & Munirathanam, M. (2021). Feature selection model for healthcare analysis and classification using classifier ensemble technique. *International Journal of Systems Assurance Engineering and Management*.

Nordin, N., Zainol, Z., Mohd Noor, M. H., & Chan, L. F. (2022). Suicidal behaviour prediction models using machine learning techniques: A systematic review. *Artificial Intelligence in Medicine*, *132*(August), 102395.

Peng, L., Cai, Z., Heidari, A. A., Zhang, L., & Chen, H. (2023). Hierarchical Harris hawks optimizer for feature selection. *Journal of Advanced Research*, *xxxx*.

Selvakuberan, K., Kayathiri, D., Harini, B., & Devi, M. I. (2011). An efficient feature selection method for classification in health care systems using machine learning techniques. *ICECT 2011 - 2011 3rd International Conference on Electronics Computer Technology*, *4*, 223–226.

Shailaja, K., & Scholar, M. T. (2018). *Machine Learning in Healthcare : A Review*. *Iceca*, 9–13.

Suparyanto dan Rosad (2015. (2020). Data Mining: Practical Machine Learning Tools and Tecniques. In *Suparyanto dan Rosad (2015* (Vol. 5, Issue 3).

Tanyıldızı, E., & Cigali, T. (2017). Kaotik Haritalı Balina Optimizasyon Algoritmaları. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, *29*(1), 307–317.

Taylor, R. A., Moore, C. L., Cheung, K., & Brandt, C. (2018). Predicting urinary tract infections in the emergency department with machine learning. *Plus One*, *3*, 1–15.

Tsanas, A., Little, M. A., Fox, C., & Ramig, L. O. (2014). Objective automatic assessment of rehabilitative speech treatment in Parkinson's disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(1), 181–190.

Xiong, Y., Lin, L., Chen, Y., Salerno, S., Li, Y., Zeng, X., & Li, H. (2022). Prediction of gestational diabetes mellitus in the first 19 weeks of pregnancy using machine learning techniques. *Artificial Intelligence in Medicine*.

Xue, B., Zhang, M., & Browne, W. N. (2013). Particle swarm optimization for feature selection in classification: A multi-objective approach. *IEEE Transactions on Cybernetics*, *43*(6), 1656–1671.

Yuan, Z., Chen, B., Liu, J., Chen, H., Peng, D., & Li, P. (2023). Anomaly detection based on weighted fuzzy-rough density. *Applied Soft Computing*, *134*, 109995.

Zhang, X., Wu, H., Chen, T., & Wang, G. (2022). Automatic diagnosis of arrhythmia with electrocardiogram using multiple instance learning: From rhythm annotation to heartbeat prediction. *Artificial Intelligence in Medicine*, *132*(August), 102379.

Zheng, Y., Li, Y., Wang, G., Chen, Y., Xu, Q., Fan, J., & Cui, X. (2019). A Novel Hybrid Algorithm for Feature Selection Based on Whale Optimization Algorithm. *IEEE Access*, *7*, 14908–14923.