Optimal Feature based Selection of IOMT data based on Feature whale optimization Algorithm

Sayyed Johar[1] and Dr. Manjula G R[2]

' Department of Computer Science and Engineering, JNN College of Engineering, Shivamogga, Karnataka.

sayyedjohar@jnnce.ac.in

[2] Department of Computer Science and Engineering, JNN College of Engineering, Shivamogga, Karnataka.

grmanjula@jnnce.ac.in

Abstract

Patient records in medical datasets frequently contain huge pointless and repetitive attributes, many hundreds of which are not really necessary for medical decision-making. The effectiveness of classifiers is hampered by the amount of data, which also increases dimensionality. This matter has been resolved recently, demonstrating worth of feature selection in solving it. By deleting unnecessary and redundant characteristics, feature based selection approaches primarily aim to minimize the volume of data and boost learning algorithms' efficiency. In this study, the FBWOA novel-based meta-heuristic algorithm for feature-based selection is introduced. The algorithm, which was impacted by the methods used to hunt humpback whales, has following stages: surrounding of prey, launching s bubble-net sprial attacks, and locating. Four common medical datasets are employed for the assessment of the algorithm's performance: "Diabetes, hepatitis, original Wisconsin breast cancer, and the Pima Indians". The results highlight the algorithm's capacity to substantially reduce dataset dimensionality while retaining acceptable disease diagnostic.

Keywords:*Internet of Medical Things (IoMT), Feature based selection, "Whale optimization algorithm (WOA)", Dimensionality reduction, Diseases diagnosis; K-Nearest Neighbour .*

1.Introduction

Medical data mining, which uses analysis, pattern recognition, machine learning, and statistical analysis to find relationships and patterns hidden in patient information, has emerged as one of the most important challenges in recent years. All medical procedures can generally be broken down into following categories: screening, diagnosis, therapy, prognosis, supervision, and administration [1]. When diagnosing and predicting diseases, accuracy and sensitivity are particularly crucial. Positive comments can encourage the clinician to use analysis to improve the diagnostic and prognosis process. As a result, it is possible to lower treatment costs while raising the standard of health in society. Real-world medical databases are frequently packed with pointless and redundant characteristics that raise their dimensions or make them susceptible to the curse of dimensionality. After that, the learning process' accuracy, cost calculations, and speed are all impacted. To address this problem, dimensionality reduction techniques have been developed. Feature-based selection is one of the widely used dimensionality reduction approaches. Choosing a useful subset of features from a larger collection of characteristics that may include overlapping and potentially inconsequential traits with varying degrees of significance is the goal of the feature-based subset selection issue[2, 3].

According to how heavily they rely on classification algorithms, features-based selection strategies can be split into two broad groups: model-free and model-based [4]. The feature extraction process used in model-free approaches is independent of any particular data model and is based on statistical functions. The Maximum relevance with minimal redundancy, information gain, correlation function, and F-score criterion strategy are a few popular model-free methods. The selection strategy and feature extraction in the model-based approach depend on the predictor's output. In terms of speed, scalability, and computational cost, model-free feature extraction is superior to model-based feature extraction since it is independent of a particular data model. But in contrast to model-based approaches, this independence reduces performance. To extract the most useful based subset of features, the issue space must often be thoroughly searched. However, due to high dimensionality, particularly in NP-hard issues, it is unfeasible to employ all look for most of the real-world circumstances. Obviously, it is quite expensive in terms of computational complexity and reaction time to explore the entire issue space and evaluate every state. In order to find the best answers, numerous meta-heuristic methods have been suggested. Description of the foraging habits of the wildlife in the natural world. They primarily center their consideration of the a compromise between computing Time and complexity of swarm intelligence, which is distributed via agent rivalry and cooperation. As a result, some effective meta-heuristic software for feature selection have been devised. Examples include Artificial Bee Colony (ABC) [7] and "Particle Swarm Optimization" Ant Colony Optimization Algorithm (ACO) [5] are two examples. Recently, they submitted proposals to a variety of applications in the medical sciences [8, 9].

In this study, we propose a meta-heuristic algorithm called FBWOA, which stands for "Feature-based selection on dataset using "Whale Optimization Algorithm"". The primary goal of FBWOA is to reduce the number of dimensions in medical data. This FBWOA actually follows the "hunting Humpback techniques on whales," which include three Three main strategies are used: " spiral bubble-net attack, enveloping prey, and looking for prey.."

2. Related Work and Literature Review

In classification problems, feature based selection techniques are commonly employed to condense the feature base set and improve the effectiveness and precision of the classifier [10]. Exhaustive evaluation is not workable when taking care of many features, which calls for use with meta-heuristic search approaches [11]. These approaches use agent-based tactics that involve competition and cooperation to handle real-world problems, and various studies have focused on feature selection within the context of swarm intelligence. Advanced Binary ACO (ABACO), which aims to simultaneously select features and reduce dimensions, is a noteworthy method [11].When using the Attribute based AWAIS, or Artificial Immune Weighted System to identify heart and diabetes disorders in 2005, ahan et al. demonstrated the negative effects of irrelevant features on disease diagnosis [12].

In a different work, Huange developed a hybrid strategy that combined the support vector machine and the ant colony optimization algorithm" for efficient feature based selection [13]. To find relevant features, Inbarani et al. created a system that incorporated rough set techniques and PSO [14]. Heart disease diagnosis model was carried out by Nahar et al. using computational intelligence [15].Another swarm-inspired metaheuristic optimization Technique that uses of humpback whale hunting behavior is the whale optimization algorithm (WOA), which was developed by Mirjalili and Lewis [23].Based on a ten-year history, Lee et al. [22] developed a stroke prediction algorithm in 2018. Their main goal was to use information from the Korean National Health Examination to estimate the risk of stroke occurrence. They also established a customized alarm system that provides information on stroke risk factors and is related to the likelihood of having a stroke. This study aims to inspire medical consumers to improve their health management techniques and subsequently alter their health behaviors for the better.A personalized decision-making technique that can address the issues of missing data in healthcare was introduced in 2019 by Azimi et al. [15]. Their algorithm successfully imputed missing values using IoT-based data resources, producing acceptable results. A genuine trial involving 20 pregnant women who were tracked for seven months to validate the method used heart rate information to gauge maternal health. Comparative analysis showed the model to be more precise, especially when using with large window sizes. Using the UCI Repository's dataset, Kumar et al. [8] presented a novel systematic approach for the diagnosis of diabetes in 2018. To identify and evaluate the severity of the sickness, they developed a brand-new classification model called Rule-based Neural Classifier. Real hospital data also benchmark data coming from the Repository of UCI were used in the investigation. Results convincingly demonstrated the suggested model's superiority over traditional alternatives.

A data protection strategy based on an ENN classifier and incorporating cryptography and authentication was developed in 2019 by Krishnan et al. [18]. This novel method included two separate procedures, one on the client side and the other on the cloud side. Patients' EEG signals were processed on the client-side using a combination of ANC (Adaptive Noise Cancelling) and HWT (Haar Wavelet Transform) models. The ECC (Elliptic Curve Cryptography) protocol was then used to protect the signals from counterfeiting. On the cloud side, attributes from authorized data were retrieved and classified as abnormal or normal. This method was used in an ECG-based health screening scenario and its efficacy was proved by comparison to the OCSVM (One-Class Support Vector Machine) model with IoT for testing and simulation in the cloud. Refer to Table 1 for a succinct assessment of the characteristics and difficulties of current IoT-oriented healthcare data modeling.

In their study published in 2020 [2], Akkaş et al. examine patient healthcare through the utilization of wearable Internet of Things (IoT) devices. The study concentrates on monitoring health parameters such as plethysmograph readings, pulse rate, and relative oxygen levels. Collected vital signal are stored within a centralized dataset employing IoT technology. The acquired data is then effectively utilized to develop a comprehensive health monitoring system

In a study by H. T. Yew et al. published in 2020 [4], real-time patient monitoring systems are developed using Internet of Things (IoT) principles. The primary objective of these systems is to alleviate challenges related to accessing healthcare providers and to reduce associated costs. The patient's cardiac data is acquired through ECG-related sensors. This data is then transmitted using the message queuing telemetry transport protocol. Subsequently, doctors can access the collected information through LAN or WAN connections, enabling them to devise tailored treatments based on the patient's health condition.

In 2020, Motwani et al. [8] employed deep learning in conjunction with innovative cost optimization methods to start a smart healthcare framework and recommendation system. This system utilizes IoT sensors to capture the patient's movements comprehensively. The information obtained is then used to predict blood pressure disorders through a deep learning model. The framework records the patient's activities at 15-minute intervals, aiding in the mitigation of risk factors. Based on the analyzed data, patients are promptly alerted to emergencies.

In 2020, Huang et al. [6] investigated chronic kidney disease in patients through remote patient monitoring (RPM) systems. The study focuses on exploring the correlation between patients' emotional sharing and technical issues during dialysis within the RPM context. The researchers design software integrated with RPM to capture patient sentiments and emoticon-sharing activities. The gathered data is then subjected to logistic regression analysis to establish connections between emotions and sentiments. These identified relationships serve as valuable indicators for the effective identification of patients with chronic kidney disease..

An Internet of Medical Things (IoMT) framework for the management of senior patients' health was developed in 2020 by T. Zhang et al. [6]. This framework focuses on overseeing the cardiac data of older adults by utilizing deep learning (DL) techniques. DL extracts precise cardiac image information from potentially noisy devices, thereby enhancing the accuracy of predicting the elderly individuals' health status. Notably, a self-adaptive power control strategy, bolstered by an efficient aware technique, is integrated into the system. This combination effectively ensures device reliability, prolongs battery lifespan, and minimizes the energy utilization of wearable IoMT devices

3. Proposed Algorithm

The Feature based Selection on whale Optimization algorithm (FBWOA) is a suggested meta-heuristic method that is introduced in this section. The three main components of FBWOA's hunting strategy encircling prey, spiral bubble-net assaulting, and looking for pre are motivated by the actions of humpback whales. Figure 1 illustrates the FBWOA flowchart and key phases.

.

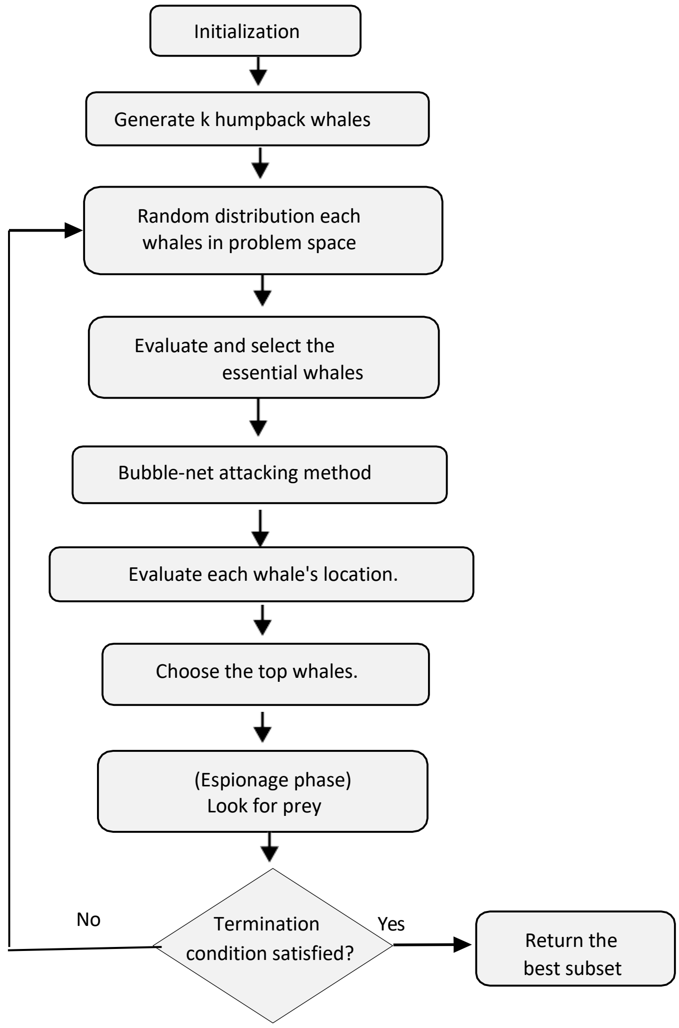

Fig 1: The flow of proposed FBWOA

The first stage involves creating k humpback whales and dispersing them at random around the search space. The best whales are chosen after an evaluation of their postures. The remaining whales alter their positions to face the top whale.Moving on the second stage, humpback whales assault bubble nets by employing two tactics: spiral position update and diminishing encircling. Each whale proposes a feature subset during this step, which is comparable to an exploitation stage. According to the testing set's classifier accuracy, these subgroups are assessed.Humpback whales randomly look for prey according to positions in relation to one another during the third and final stage, called as the exploration phase.The first stage involves creating k humpback whales and dispersing them at random around the search space. Their standings are

The best whales are chosen after being evaluated. The positions of the remaining whales are modified to face the top whale. Moving to the next stage, humpback whales assault bubble nets by employing two tactics: spiral position update and diminishing encircling. Each whale proposes a feature subset during this step, which is comparable to an exploitation stage. Due to testing set's classifier accuracy, these subgroups are assessed.Humpback whales randomly look for prey based on a citation for the relation to one another during the third and final stage, named as the exploration phase.

The “Whale Optimization Algorithm (WOA)” draws an idea from "Humpback whales" and their hunting techniques [23]. Blunt-nosed whales, belonging to the baleen whale family, exhibit a distinctive hunting strategy. Notably, they possess the ability to detect prey and employ an encircling maneuver. This encircling action is represented by Equation (1) and Equation (2), as shown below:

Z = |P C(it) minus C(it)| (1)

C(it + 1) = C(it) minus Q.Z (2)

The location of a vector are P and Q, C is the location of a vector, the form || represents the absolute value, and '.' denotes element multiplication. It is evident that iteration is occurring from Equations (1) and (2). The location of a vector of the best among the results is provided by C. The vectors P and Q are evaluated using equations (4) and (3), where b is carefully lowered from 2 to 0 over all rounds where the random vector rv is between [0, 1].

Q = 2b.rv − b (3)

P = 2.rv (4)

There are two ways to depict humpback whales. whale numerically by bubble-net technique; Lowering the value of b in the circumferential mechanism Equation (3). The difference between the prey's position at the start of the spiral updating position strategy is (X, Y) and the whale's position at (X, Y). Then, to mimic the helix-shaped group of humpback whales, the equation of spiral is established among the whale's position and its prey, according to Equation (20), where Z = |P*C(it)- C(it)| denotes the length among the whale and its prey, an is taken into consideration a random number is indicated by the constant word l between [1, 1], and "." denotes an element analysis.

$$c(it + 1) = Z' \cdot e^{al} \cdot \cos*(2\pi l) + C(it) \qquad (5)$$

Equation (6) illustrates how to update the response using statistics according to the surrounding, receding area method or the spiral approach, q represents random number between [0,1].

$$c(it + 1) = \begin{cases} C(t) - Q*Z & \text{if } q \leq 0.5 \ Z \\ Z' \cdot eal \cdot \cos(2\pi l) + C.(it) & \text{if } q > 0.5 \end{cases} \qquad (6)$$

C has the potential to be utilized for look for the victim vector. Searchers are discouraged from the reference whale by Vector C's random values in the limt [1,1]. Equations (7) and (8), where location vector position is Crand chosen from the current solutions, display the numerical equation.

Z = |P*C rand − C| (7)

c(it + 1) = Crand - Q*Z (8)

The representation of the algorithmic Feature Based WOA shown in Algorithm 1.

Do the population initialization as C_i, where $i = 1, 2, \cdots, ne$
Evaluate the fitness value of every search agent
C^* is the best search agent
it_{max} indicates the maximum number of iteration
while$(it < it_{max})$
  for each search agent
  Update b, Q, P, l, and q
    if $fit(it)$ is better than $fit(it-1)$
      Compute q using Equation (23)
    else
      Compute q using Equation (20)
    End if
    if1$(q < 0.5)$
      if2$(|Q| < 1)$
        Update the solution by Equation (17)
      else if2$(|Q| \geq 1)$
        Select a random search agent (C_{rand})
        Update the solution using Equation (23)
      end if2
    else if1$(q \geq 0.5)$
        Update the solution using Equation (20)
    end if1
  end for
  Make sure if any search agent is going afar from the search space and rectify it
  Evaluate the fitness value of each search agent
  Update C^* if a better occurs
  $it = it + 1$
  end while
return C^*

Algorithm 1: Feature Based WOA shown in Algorithm

4. Experimental Evaluation

In this evaluaion, four medical benchmark datasets that were retrieved from the "machine learning repository" at UCI. [17] are used to assess the outcome of the suggested strategy. These common datasets for feature-based selection issues include “Hepatitis, diabetes, original Wisconsin breast cancer, Statlog, and Pima Indian diseases”. The datasets' statistical data are displayed in Table -1.

Table 1: Statistical Information of Datasets

Dataset	Features	Sample	Classes	Missing data
“Pima Indians Diabetes”	8	769	2	Yes
“Original Wisconsin Breast Cancer”	10	698	2	yes
Statlog	13	272	2	no
Hepatitis	19	156	2	yes

The “Pima Indians Diabetes dataset” (PID) uses clinical and laboratory data to determine if a person has diabetes or not. Breast cancer diagnosis is the objective of the Original Wisconsin Breast Cancer dataset. The Statlog dataset predicts cardiac disease. Predicting whether a person with hepatitis will live or die is the objective of the Hepatitis dataset. Since medical datasets are noisy and incomplete in the actual world, each dataset is first normalized before the propose method is tested.

4.1 Evaluation Functions

The “sensitivity, specificity, precision, accuracy, negative predictability (NPV), and space behind the curve (AUC) are few of the known-well evaluation functions used to assess the subsets of characteristics chosen by the FBWOA algorithm. These are the class samples of both positive and negative that are accurately identified are indicated by the sensitivity and specificity stated in Eqs. (9) and (10) accordingly. Equations (11) and (12) is used to find precision, also known as PPV and NPV, or positive and negative predictive values. AUC measures the ratio of true positives to false positives and ranges in value from 0.0 to 1.0. The suggested algorithm's cost function is determined by the classifier's accuracy, as indicated by Eq. (13), where PT and PF combined denotes the amount of subjects with positive test results and adding NF and NT denotes the amount percent participants with negative test results .

$$\text{True positive rate, or sensitivity} = PT/PT+EN \quad (\%) \qquad (9)$$

$$\text{(True negative rate) Specificity} = NT/NT+PF \quad (\%) \qquad (10)$$

$$\text{Precision (higher likelihood of success)} = PT / PT+PF \quad (\%) \qquad (11)$$

$$\text{negative predictability} = NT / NT+NF \quad (\%) \qquad (12)$$

$$\text{Precision (ACC)} = (PT+NT) / (PT+NT+PF+NF) \quad (\%) \qquad (13)$$

4.2 Experimental Setup

Python is used to implement the suggested technique on an Intel Core i7 computer with 8GB of RAM. Our approach is put to the test 15 times using the evaluation functions mentioned in Section 4.1 to select the ideal subset of attributes. Each time, the the datasets were divided into a training set and a test set., each comprising 70% and 30% of the taken data. After that, the proposed technique evaluates the subset of chosen characteristics using the K=3 K-Nearest Neighbors method with K=3. Also, the starting a population of 30 is assumed., the maximum the set iteration to 60, and the lower and high bounds are set, respectively, to 0 and 1.

4.2 Experimental Results

Tables -1 and 2 in this section contain the experimental findings of the suggested algorithm. The terms mean and maximum refer to the highest and average numbers, respectively. The suggested algorithm's accuracy on several medical datasets is as follows: For Hepatitis, Breast Cancer, Diabetes, and Pima Indians, the rates are 87.10%, 97.86%, 78.57%, and 77.05%, respectively. among particular, the algorithm chooses 6 criteria for diagnosing heart illness, 7 features for diabetes among Pima Indians, 8 features due to hepatitis, and four features due the breast cancer.

The quality based reduction rates are calculated for each dataset. The reduction rates for hepatitis, diabetes among Pima Indians, heart disease, and breast cancer are 57.89%, 60%, 12.5%, and 53.85%, respectively, as shown in Table 3. The dimensionality reduction rate and classifier accuracy for diabetes and heart disease are shown in Figures 2 and 3.

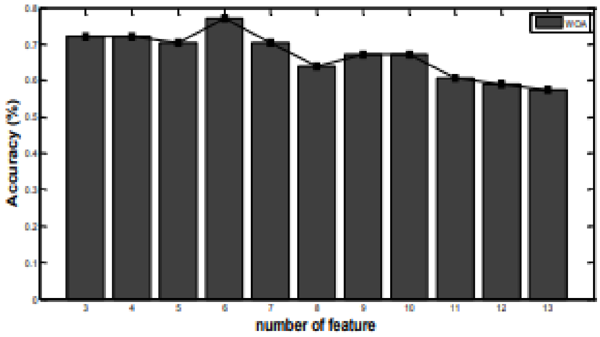

Fig. 2. The accuracy of classifier for different number of features in Heart disease dataset

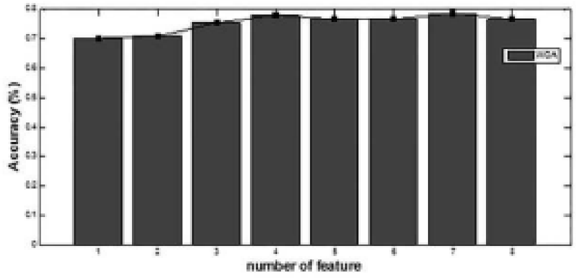

Fig. 3. The accuracy of classifier for different number of features in Pima Indian Diabetes dataset.

TABLE -2
Functions of evaluation (in %) of FBWOA algorithm

Datasets	# NF	Accuracy	sensitivity	specificity
Hepatitis	8			
Max		87.10	100.00	94.12
Mean		73.33	70.51	79.97
Breast Cancer	4			
Max		97.86	98.90	100.00
Mean		96.57	96.60	96.47
Pima Indians Diabetes	7			
Max		78.57	90.65	63.46
Mean		70.87	82.72	48.28
Statlog	6			
Max		77.05	100.00	82.76
Mean		62.84	94.71	64.36

TABLE - 3
Functions of evaluation (in %) of FBWOA algorithm

Datasets	# NF	AUC	PPV	NPV
Hepatitis	8			
Max		0.971	89.89	99.45
Mean		0.752	76.99	76.56
Breast Cancer	4			
Max		0.994	100.00	98.92
Mean		0.965	98.17	93.64
Pima Indians Diabetes	7			
Max		0.771	81.83	71.59
Mean		0.655	76.37	58.67
Statlog	6			
Max		0.913	91.24	99.56
Mean		0.795	77.74	91.53

TABLE -4
Features based subset and rate of reduction of effective characteristics

Dataset	Reduction rate (%)	Effective characteristics
Hepatitis	58.89	[27,5,28,11,15]
Breast Cancer	62.00	[7,2,3,10]
Statlog	54.85	[2,3,6,9,10,11]
PID	13.5%	[5,6,1,3,8,4,7]

5. Conclusion

This word presents FBWOA, a powerful feature selection method that makes use of the whale optimization process. FBWOA seeks to reduce problem space dimensionality while improving learning algorithm performance. The algorithm demonstrated its capacity to discover the best feature subsets and was created expressly for use in medical decision-making. FBWOA demonstrated promising results in lowering dataset dimensions for disease detection while retaining appropriate accuracy levels through experimentation on diverse medical datasets.

Acknowledgment

This project is funded by Karnataka Science and Technology Promotion Society (KSTePS), Vision Group on Science and Technology (VGST), and the Karnataka government and The authors would like to express gratitude to staff and management of JNN College of Engineering for their support.

Acknowledgment

This project is funded by Government of Karnataka, Vision Group on Science and Technology (VGST), and Karnataka Science and Technology Promotion Society (KSTePS) and the authors wish to express their intention to express gratitude to staff and management of JNN College of Engineering for their support.

REFERENCES

1. M. Hassanalieragh et al., "Health Monitoring and Management Using Internet-of-Things (IoT) Sensing with Cloud-Based Processing: Opportunities and Challenges," 2015 IEEE International Conference on Services Computing, New York, NY, USA, 2015, pp. 285-292, doi: 10.1109/SCC.2015.47.
2. Akkaş, M. Alper, Radosveta Sokullu, and H. Ertürk Çetin. "Healthcare and patient monitoring using IoT." Internet of Things 11 (2020): 100173.
3. Archip, N. Botezatu, E. Şerban, P. Herghelegiu and A. Zală, "An IoT based system for remote patient monitoring," 2016 17th International Carpathian Control Conference (ICCC), High Tatras, Slovakia, 2016, pp. 1-6, doi: 10.1109/CarpathianCC.2016.7501056.
4. H. T. Yew, M. F. Ng, S. Z. Ping, S. K. Chung, A. Chekima and J. A. Dargham, "IoT Based Real-Time Remote Patient Monitoring System," 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), Langkawi, Malaysia, 2020, pp. 176-179, doi: 10.1109/CSPA48992.2020.9068699.
5. Rajan Jeyaraj, Pandia, and Edward Rajan Samuel Nadar. "Smart-monitor: patient monitoring system for IoT-based healthcare system using deep learning." IETE Journal of Research (2019): 1-8.
6. Huang, Robin, Na Liu, Mary Ann Nicdao, Mary Mikaheal, Tanya Baldacchino, Annabelle Albeos, Kathy Petoumenos, Kamal Sud, and Jinman Kim. "Emotion sharing in remote patient monitoring of patients with chronic kidney disease." Journal of the American Medical Informatics Association 27, no. 2 (2020): 185-193.
7. Ali, Farman, Shaker El-Sappagh, SM Riazul Islam, Daehan Kwak, Amjad Ali, Muhammad Imran, and Kyung-Sup Kwak. "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion." Information Fusion 63 (2020): 208-222.
8. Motwani, Anand, Piyush Kumar Shukla, and Mahesh Pawar. "Smart Predictive Healthcare Framework for Remote Patient Monitoring and Recommendation Using Deep Learning with Novel Cost Optimization." In International Conference on Information and Communication Technology for Intelligent Systems, pp. 671-682. Springer, Singapore, 2020.

9. Pathinarupothi RK, Durga P, Rangan ES. IoT-Based smart Edge for global health: remote monitoring with severity detection and alerts transmission. IEEE Internet Things J. April 2019;6(2):2449–2462.

10. Kalyani G, Chaudhari S. An efficient approach for enhancing security in internet of things using the optimum authentication key. Int J Comput Appl. 2019;42(3):306–314

11. Limaye A, Adegbija T. HERMIT: A benchmark suite for the internet of medical Things. IEEE Internet Things J. Oct. 2018;5(5):4212–4222

12. Yang S, et al. Iot Structured long-term wearable social sensing for mental wellbeing. IEEE Internet Things J. April 2019;6(2):3652–3662.

13. Tseng Y-J, Shih Y-L. Developing epidemic forecasting models to assist disease surveillance for influenza with electronic health records. Int J Comput Appl. Aug. 2019.

14. Rani S, Ahmed SH, Shah SC. Smart health: a novel paradigm to control the Chickungunya Virus. IEEE Internet Things J. April 2019;6(2):1306– 1311.

15. ImanAzimi TP, Rahmani AM, Niela-Vilén H, et al. Missing data resilient decision-making for healthcare IoT through personalization: A case study on maternal health. Future Gener Comput Syst. July 2019;96: 297–308.

16. SeyedFarhadAghili HM, Shojafar M, Peris-Lopez P. LACO: lightweight three-factor authentication, access control, and ownership transfer scheme for E-health systems in IoT. Future Gener Comput Syst. July 2019;96: 410–424

17. ilela PH, Rodrigues JJPC, Solic P, et al. Performance evaluation of a fogassisted IoT solution for e-health applications. Future Gener Comput Syst. August 2019;97:379–386.

18. krishnan S, Lokesh S, Ramya Devi M. An efficient Elman neural network classifier with cloud supported internet of things structure for health monitoring system. Comput Netw. March 2019;151:201–210.

19. Abbaszadeh B, Haddad T, Yagoub MCE. Probabilistic prediction of Epileptic Seizures using SVM. In Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany,2019.

20. Gaurav D, Tiwari SM, Goyal A, et al. Machine intelligence-based algorithms for spam filtering on document labeling. Soft comput. 2019: 1–14.

21. Mishra S, Sagban R, Yakoob A, et al. Swarm intelligence in anomaly detection systems: an overview. Int J Comput Appl. 2018: 1–10.

22. K. Bache, M. Lichman (2016) UCI machine learning repository. School of Information and Computer Science, University of California, Irvine. Available: http://archive.ics.uci.edu/ml.

23. S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," Advances in Engineering Software, vol. 95, pp. 51-67, 5// 2016.