An IoT Enabled Heart Disease Monitoring System Using Grey Wolf Optimization and Deep Belief Network

Sandhiya S (✉ sksandhiya89@gmail.com)  
University College of Engineering Villupuram  https://orcid.org/0000-0002-8458-4490

Palani U  
IFET College of Engineering

Research Article

Keywords: Heart disease, deep learning, feature selection, Internet of Things, health monitoring system and IoT devices.

Posted Date: January 31st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1058279/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
An IoT Enabled Heart Disease Monitoring System Using Grey Wolf Optimization and Deep Belief Network

*S.Sandhiya*
Teaching Fellow
Dept. of Information Technology
University College of Engineering Villupuram
Tamilnadu, India
sksandhiya89@gmail.com

Dr U.Palani
Professor
Dept. of Electronics and communication Engineering
IFET College of Engineering
Tamilnadu, India
palani@ifet.ac.in

Abstract – Heart disease is a deadly disease today irrespective of age group among people because they are not aware about their heart disease level, and the kind of heart diseases. It is necessary to aware about their type of heart diseases and the regular disease monitoring process in this quick world. This paper proposes a new heart disease monitoring system with the incorporation of Internet of Things and Deep Learning technique for safeguarding the patients. Here, a new feature selection algorithm which is incorporated for performing better classification through deep learning algorithm. In this proposed heart disease monitoring system, we monitor the disease level according to the inputs that are given to the IoT devices. Moreover, it classifies the patient details according to the heart disease types and the severity of the disease. Finally, it gives alarm/message to the patients according to the type of heart disease with the available inputs. The experiments have been demonstrated and proved as the proposed system is better in terms of prediction accuracy.

Keywords – Heart disease, deep learning, feature selection, Internet of Things, health monitoring system and IoT devices.

1. INTRODUCTION

The World Health Organization identified the heart disease is very dangerous and huge volume of deaths around 12 million happened in the world. Moreover, this number is high in all the countries especially in developing countries. Generally, the “heart” disease is also named “cardiovascular disease” and comprehends the assorted diseases which damage the heart. This heart disease is also threatening the adults now due to the increment of sudden deaths in all over the world. One of the statistics declares that every 40 seconds one person meet the death due to the heart disease in United States [22]. Diagnosing the heart disease is
necessary today and it is also a challenging task for the health care department of each nation. This is because patients not aware about the heart disease levels and fails to monitor the disease levels and health conditions.

Heart disease is categorized by different causes such as Coronary artery and vascular disease, Heart rhythm disorders (arrhythmias), Structural heart disease and heart failure. Among them, the coronary artery and vascular disease that happens when the arteries of a heart is blocked. This is a common heart disease and causes like chest pain (angina). In addition, the vascular disease is issues of blood vessels that affect the blood circulation and it affects the heart functionality. Heart rhythm disorders (arrhythmias) that causes because of less heart beat that affect the blood circulation. Structural heart disease refers to abnormalities of the heart’s structure including its valves, walls, muscles or blood vessels near the heart. Heart failure is a very serious condition which develops after the heart becomes weakened. Heart failure will happen due to the high blood pressure (BP). This kind of heart failure can be prevented early [28].

The major symptoms for the heart diseases are including high BP, high cholesterol, diabetes, Sleep apnea, unhealthy diet, no physical activity, abnormal weight, smoking, too much alcohol, stress, age, family medical history, heritage and personal circumstances. Here, the majority people affected with coronary heart disease. For this purpose, this work proposes a new health monitoring system to identify the type of heart disease according to the given inputs and available information and also taken care more if the person affected with coronary type of heart disease.

The deep learning techniques are playing major role in the process of predicting the disease level in the disease monitoring system. It is capable of handling large volume of data faster than the machine learning algorithm. Usually, the machine learning algorithms work well than the deep learning algorithm when used small size of dataset. At the same time, the deep learning algorithms work well than machine learning algorithms when handle very large volume of data with consuming more time. Deep learning algorithms learn the data set repeatedly instead of once in training process. The features of the dataset can be analysed in depth due to the incorporation of repeated learning. For this purpose, it is taking more time than the machine learning algorithm. Three major types of deep learning algorithms are available such as Convolutional Neural Networks (CNN), Recurrent Neural Network (RNN)
and Least Short Term Memory (LSTM). All these algorithms are having many layers in between the input and output layer instead of hidden or processing layer such as convolutional layer, pooling layer, fully connected layer and softmax layer. In general, the softmax layer contains the ReLu Activation function for making final decision over the deep learning algorithm. The Internet of Things (IoT) techniques and devices were used for collecting the large volume of streaming data. The streaming data have been collected from various people in various places around the world. The necessary available devices and sensors have been incorporated with the IoT technology.

Feature selection is playing major role to predict the disease level in the respective type of heart disease. This work proposed a new wolf optimization algorithm-based feature selection method for selecting the most contributed features that are useful for predicting the different kinds of heart diseases. Moreover, an existing deep learning algorithm called Deep Belief Network (DBN) is used for classifying the datasets by conducting the effective training and testing processes over the benchmark and real time datasets. This disease monitoring system identifies the type of heart disease and it monitors the disease level. The rest of this paper is organized as below: Section 2 describes in detail about the relevant works in the direction of feature fusion process, deep learning and IOT. Section 3 explains the overall system architecture that demonstrates that the working process of the proposed model. Section 4 describes the proposed heart disease monitoring system with the necessary subsections that are explained the feature fusion and classification process. Section 5 demonstrates the experimental results and discussion. Finally, this work concludes with necessary highlights and future works.

2. LITERATURE REVIEW

Many heart disease prediction systems have been developed by various researchers in the past. Among them, Hui et al [1] developed for tackling the issue of risk factor identification and prediction for designing the monitored risk factors by applying DBN that learnt by contrastive divergence method. They have identified that the way for designing the uniqueness and the discrimination from various groups of patient’s data. They have achieved 95% prediction accuracy. Sandhiya et al. [2] proposed an disease prediction system which consists of feature selection method that works as incremental in nature named as
Incremental Feature Selection Algorithm (IFSA) which combines the concepts of Intelligent Conditional Random Field (CRF) on feature selection process and the Linear Correlation Coefficient based Feature Selection (ICRF-LCFS) method algorithm and an existing Convolutional Neural Network (CNN) with temporal features (T-CNN). Sethukkarasi et al [3] proposed a new fuzzy cognitive map based intelligent disease prediction system with the consideration of temporal constraints for predicting the heart disease, breast cancer disease and diabetic disease. They have achieved better performance than others. Ganapathy et al [4] introduced a new fuzzy temporal incorporated fuzzy min-max neural network for predicting the heart, cancer and diabetic diseases. They have achieved better result than other works in terms of prediction accuracy. Ahmed and Ayman [5] developed a diagnosis method which works automatically to detect the breast cancer according to the deep belief network that incorporated the pre-training and followed by the incorporation of supervised back propagation neural network. The overall prediction accuracy is achieving 99.68% accuracy with 100% sensitivity and 99.47% specificity over the breast cancer dataset.

Sharawi et al [6] recommended a new kind of Whale Optimization Algorithm which is useful wrappers method for selecting the contributed features. On the other hand, the WOA is also introduced as hybrid method for enhancing the performance in all iterations by Mafarja and Mirjalili [7].They also introduced two more variants such as roulette wheel and tournament selection is also applied instead of the random operator on WOA in 2018. In addition, they have used crossover and mutation operators for achieving better selection process which is capable of producing better classification accuracy. In another work Sayed et al [8], they have applied a chaotic search method with WOA for combating the local optimization and the slow convergence speed issues which considers the feature selection issues too. Zheng et al [9] introduced a new filter approach-based feature selection algorithm that adopts the Pearson correlation coefficient and correlation distance.

Chao et al [10] developed a new pervasive monitoring system which is able to send the signs for the real time application. Their system consists of two parts such as data acquisition and data transmission. Based on the feature analysis, they have analysed the heart disease level according to the inputs in the form of sign signals. Jun Qi et al [11] identified few new researches works and drawbacks in the direction of IoT. Finally, they have considered some successful studies in the mentioned area and also express the possible applications for future. Yixue et al [12] developed a new Recurrent CNN with the adaptation of multi-modal to predict the risk level. They have extracted the fine-tuned features that are structured and non-
structured along with the patient’s dataset and the medical report in the form of text. Moreover, they have applied DBN for obtaining the variance from the structured data to unstructured. In addition, they have collected patient’s data as medical big data from second grade level hospitals in China between the years 2013 and 2015. Finally, their model achieved 96% prediction accuracy which is higher than the available models.

Sandhiya S and Palani U [13] proposed a hybrid feature subset selection algorithm called Genetic Binary Cuckoo Optimization Algorithm (GBCOA) which comprises of two steps, in first step subset generation is performed by using Binary Cuckoo search (BCS) optimization algorithm it will generate n number of subsets from the original large dataset. In second step subset selection is performed by using Genetic Algorithm it will select the best subset from the generated n number of subsets. In this paper the feature selection process is improved by combining the two different algorithms. Vivekanandan et al.[14] have selected most contributed attributes from the available features that are helpful to perform effective heart disease diagnosis. They have applied a modified differential evolution for performing feature selection process over the cardiovascular disease and also optimized the selected features. Moreover, their algorithm has been tested along with the classifiers such as feed forward neural network and fuzzy AHP. Finally, they have achieved 83% prediction accuracy which is better than the available hybrid models in this direction. Jun Huang et al [15] developed a new approach that performs feature selection and classification processes through multi-label learning process. They have considered the pair wise label correlation processes and also built the multi-label classification over the learned low-dimensional data representations. Mahdieh et al [16] proposed a new filter approach named Multivariate Relative Discrimination Criterion(MRDC) to select most valuable features for performing classification. Their approach considers the feature reduction, feature redundancy by applying the redundancy and relevancy. Moreover, their approach considers the document frequencies for each term when estimates the usefulness. They have used three kinds of real-time datasets for assessing the efficiency of their approach and achieved better result than existing filter approaches.

Fajr and Mamoon [17] developed an efficient method for categorizing the ECG signals with high classification accuracy. Generally, the heartbeats are classified into different waveforms that are varying person to person enormously. They have applied a dataset for evaluating the
performance of their approach. Moreover, they have tested with various classifiers such as Random Forests, Decision Tree and Gradient Boosted Trees. Meng Lu et al [18] developed a new feature selection algorithm which is embedded from a sparse learning aspect that is capable of adjusting the unknown heterogeneity. The performance of the selected features over the multi-class classification processes. They have used synthetic and standard datasets for knowing the selection strategy and unknown heterogeneity. Finally, the proposed method achieved higher superiority on various embedding classifiers and statistical models. Saurabh et al [19] conducted an extensive analysis over the machine learning approach along with the datasets such as D1 and D2. Among them, D1 considers conjunctivitis, diarrhoea, stomach pain, cough and nausea related data and D2 contains the standard dataset called WebKB4. The machine learning algorithms namely Radial function incorporated SVM, Multi-layer perceptron (MLP) and the Random Forest (RF). Finally, they have achieved 97% as a highest accuracy in their experimental results. Moreover, they also shown the memorization process and tested in the direction of maintaining stability and the reliability.

Yu Xue et al [20] developed a new feature selection algorithm called self-adaptive particle swarm optimization (SaPSO) to select the most contributed features on very large size of datasets. They have employed encoding method to perform effective feature selection in their SaPSO and also investigated three self-adaptive methods related issues. Their model outperforms than the existing algorithms specifically large-scale feature selection issues. Mohamed et al [21] proposed a grey wolf algorithm-based feature selection method to resolve the feature selection issues. Their method is adopted a two-phase mutation operator and improved the selection process. They have taken steps to enhance the classification accuracy and to identify new feature as additional feature which is capable of achieving highest classification accuracy. Moreover, they have applied a new fitness function to achieve highest classification accuracy. Finally, they have proved that their method is better than the existing optimization algorithms such as PSO, MVO, FA and WA.

Shreshth et al [22] developed a new framework named Health Fog to integrate the ensemble based deep learning technique for deploying the real time application of automatic Heart Disease analysis. All the existing works have been achieved reasonable percentage of accuracy in classification and prediction. Even though, no prediction system has been directed to the patients for knowing their heart disease level and not instructed them to go for
the next level treatment. For this purpose, this paper introduces a new model for knowing their type of heart disease and monitoring the disease severity level.

3. SYSTEM ARCHITECTURE

The proposed heart disease monitoring system architecture is shown in figure 1. which consists of six components such as UCI Repository Dataset, IoT device, Sliding Window, Decision Manager, Feature Selection Module, Classification Module and Rule Base. IoT device collects the data from the patients. Sliding window process the streaming data which are collected through IoT device also it uses feature selection module for selecting the most contributed features in the streaming data and sent it to the decision manager. The decision manager analyses the necessary streaming data from sliding window using deep belief network and making decision over the patient data whether it is affected by the heart disease or not. Moreover, the decision manager will divide the data for finding the type of category from the standard dataset by the help of rule base.

4. PROPOSED WORK

This section explains the proposed feature selection algorithm and classification processes of the proposed heart disease prediction and monitoring model. First, the feature selection
process is explained in detail with the brief working process of the proposed feature selection method.

4.1 Feature Selection

This subsection is explained in detail about the feature selection process which is playing major role to categorize the dataset. This subsection is discussed about the background detail of the existing Grey Wolf Optimization and the proposed Grey Wolf Optimization based Feature Selection Algorithm (GWO-FSA).

4.1.1 Grey Wolf Optimization Algorithm

The grey wolves are identified as predators which is fantastic with high level skills in the process of catching their prey. This Grey Wolves Optimization Algorithm (Mirjalili et al. 2014) is a meta-heuristic search which is used for simulating the characteristics of the grey wolves in the process of searching, hunting and encircling their prey. Moreover, it has four kinds of wolves named as Alpha (α), Beta (β), Delta (δ), and Omega (ω). These wolves encircled their prey while processing the hunting which is modelled mathematically below:

\[
DS = |CD \cdot XY_p - XY(temp)| \tag{1}
\]

\[
XY(temp + 1) = |XY_p(temp) - AB \cdot DS| \tag{2}
\]

Where \(XY_p\) indicate the prey position and \(XY\) indicate the positions of the grey wolf at various iterations and are coefficient vectors are to be formed as below:

\[
AB = |2ab \cdot RND\bar{1} - ab| \tag{3}
\]

\[
CD = 2 \cdot RND\bar{2} \tag{4}
\]

Where \(RND\bar{1}\) and \(RND\bar{2}\) are random vectors in [0 1]. \(\bar{ab}\) indicates flow, which is decreasing linearly from the value two to 0 on different iterations:

\[
aa = 2 - temp * \frac{2}{IT} \tag{5}
\]

Where, IT indicates the maximum number of iterations. The position of the grey wolf (XX, YY) is modified based on the prey (XX*, YY*) position. The best grey wolf position is also modified by changing the vector values of \(AB\) and \(CD\).
For simulating the behaviour of hunting process, the symbols, $\beta$, and $\delta$ are beware of the potential prey locations. These symbols determine the solutions, the remaining wolves are changed their positions based on the three solutions ($\overline{XX}_1, \overline{XX}_2, \text{and} \overline{XX}_3$). Moreover, these are shown below in the form of equations.

$$\overline{XX}(\text{temp} + 1) = \frac{(\overline{XX}_1 + \overline{XX}_2 + \overline{XX}_3)}{3}$$ (6)

$$\overline{XX}_1 = \overline{XX}_\alpha - \overline{A}B_1, (\overline{DS}_\alpha), \overline{DS}_\alpha = |\overline{CD}.\overline{XX}_\alpha - \overline{XX}|$$ (7)

$$\overline{XX}_2 = \overline{XX}_\beta - \overline{A}B_2, (\overline{DS}_\beta), \overline{DS}_\beta = |\overline{CD}_2.\overline{XX}_\beta - \overline{XX}|$$ (8)

$$\overline{XX}_3 = \overline{XX}_\delta - \overline{A}B_3, (\overline{DS}_\delta), \overline{DS}_\delta = |\overline{CD}_3.\overline{XX}_\delta - \overline{XX}|$$ (9)

Here, every wolf is changed its position between its current and prey positions while attacking the prey. Now, $|A|$ is less than the value 1. The $\alpha, \beta$, and $\delta$ wolves are left from every wolves for searching the prey and also covered when happened the prey attack. $\overline{A}B_1$ is able to consider any random values that are less than -1 or greater than 1.

4.1.2 GWO based Feature Selection Algorithm

This subsection is explained that the proposed GWO based Feature Selection Algorithm which is used for selecting the most contributed features. The most contributed features are used to enhance the classification process. The pseudo-code of the proposed feature selection algorithm is given below:

**GWO based Feature Selection Algorithm**

**Input:** Benchmark Dataset or Patient Data

**Output:** Dataset with selected features

Procedure GWO-FSA (features, n, wt)

{ POP_Grey[n] = $X_i$, i=1, 2, ..., n
  A=0, a=0, C=0
  Call Fitness(POP_Grey, wt).
  Assign_Best_Grey($X_\alpha, X_\beta, X_\delta$)
  temp = 0
While (temp<IT) 
{
  For each grey wolf 
  { 
    Position of Grey Wolf using \( \vec{X}(temp + 1) = \frac{(\vec{X}_1 + \vec{X}_2 + \vec{X}_3)}{3} \) 
  }
  Update a, A, and C 
For j = 1 to m do 
  { 
    Calculate_Fitness(GW(i), wt) 
  }
Update the first three grey wolves \( XX_\alpha, XX_\beta, \text{and} \ XX_\delta \). 
} 
Return Find_Best_GW(\( XX_\alpha \)) 

The proposed feature selection algorithm selects the most contributed features that are useful for making effective decision over the dataset. First, find the fitness values for all the population of grey. Then, assigns the best grey values for the three kinds of features in the form of alpha, beta and delta. Find the position of the grey wolf by applying the formula which is given in equation (6). Calculate the fitness values for the values of grey wolves. Finally, identify the first three grey wolves and returns the best grey wolves from the first three grey wolves. According to the best grey wolves and the respective feature is to be identified as most contributed feature. Likewise, the process is repeated until get all the most contributed features from the dataset.

4.2 Classification using DBN

This subsection explains in detail about the classification process of the proposed disease monitoring system. The classification process of the proposed system is executed by applying the existing DBN which is most useful for predicting the disease types and levels. First, the background information of DBN is explained.
4.2.1 Deep Belief Network

The learning process is enhanced in each decade from the introduction of learning process through neural networks. Many versions of neural networks have been introduced and improved the learning process. Even though, the learning algorithms are struggling with weight assignment processes. Here, a general approach is applied to measure the difference between the Expected Output ‘EO’ and the Actual Output ‘AO’ is applying the squared error measure (SEM):

\[ \text{SEM} = (EO - AO)^2 \] (10)

Update the weight that is added into the old weight which is equal to the product of the learning rate and the gradient of the error function, multiplied by−1:

\[ WT_{ij} = -1 (ERR - WT_{ij}) \] (11)

Where, all the considered data is unlabelled. Even though, the Back Propagation Neural network (BPN) needs labelled training data. Therefore, the major drawback of BPN is taken long time for performing training process and also required more number of hidden layers. Vapnik et al (1936) introduced a new classifier called Support Vector Machine (SVM) for categorizing the given input dataset. Later on, [23] recommended that a way for creating nonlinear classification algorithms by using the kernel trick to maximum margin hyperplanes.

The linear classification process of SVM is to calculate the weights for each feature by applying an optimization method. On the other hand, the non-linear classification of SVMs are performed on the kernel trick by applying the mapping process of the given inputs. Generally, the non-linear classification is to be converted into linear classification while considering the high dimensional features.

Deep learning is a group of methods that are available in machine learning algorithms which tries to design a new model with high level abstractions in input data by applying model architectures calm of non-linear transformations [24]. The DBN is a graphical model which is generative and composed of multiple hidden layers along with many connections from one layer to another layer of latent variables but not between the layers [25]. As per the learning process in computation, the DBN works based on greedy manner [26]. For enhancing the performance of DBN, the weight update between visible ‘VS’ and processing layer units ‘PLU’ that unites and simply present as per the equation (12).

\[ \Delta WT_{ij} = \varepsilon (\langle V_i, H_j \rangle^0 - \langle V_i, H_j \rangle^1) \] (12)
Where, 0 indicates the network data and 1 represent the reconstruction state. In general, the DBN is facing different kinds of challenges such as i) DBN is fine-tuned as NNs, ii) DBN has more number of non-linear processing layers, iii) DBN is pre-trained and react as linear dimensionality reduction over the input data, and iv) the DBN trainer is also consider as another one input feature.

The DBN is used to train the input dataset with selected features. Even though, the DBN taken consider the dataset without labels only. The input dataset is an un-labelled dataset which is used for performing the feature selection and classification processes. In this work, the DBN accepts the two kinds of datasets such as benchmark dataset and real-time dataset. In this work, both types of datasets with contributed features are used without labelling for data pre-processing and classification processes. The weights have been assigned for each input features with the help of decision manager that refers the generated rules from the facts available in the rule base.

5. RESULTS AND DISCUSSION

The proposed disease monitoring system has been implemented by using Python programming language. This section explained in detail about the input datasets such as Heart Disease dataset which is taken from UCI Repository and the real time dataset which has been collected from various patients those are remotely located in the entire world through IoT devices. Moreover, the evaluation process can be done by using the standard performance metrics also explained in this section. Finally, the experimental results shown in this section for demonstrating the efficiency of the proposed disease monitoring system which is used to know the types of heart disease by applying rules over the classification process and mention the severity of the particular disease in that level.

5.1 Experimental Setup

The experiments have been conducted by using a newly developed software product and the IoT devices along with Raspberry PI software. The necessary sensors used for measuring the Glucose level, Heart Beat rate and Blood Pressure. These readings are to be considered as input data to the proposed system. Moreover, the heart dataset is available in the standard UCI repository dataset is also considered for performing testing process.
5.2 Data Sets

The real-time dataset is constructed with the features such as age, gender, glucose level, blood pressure and heart beat rate per minute. In this real time dataset is coming to the proposed system continuously as streaming data. The streaming data can be trained and tested by the proposed disease prediction system. On the other hand, the UCI Machine Learning Repository dataset is also used for validating the prediction result of the proposed disease monitoring system. Most of the researchers carried out their experiments by using a subset of 14 features of the 76 features that are available in the processed Cleveland heart disease database. Especially, the machine learning researchers apply only the Cleveland database for their research. The presence of heart disease in the patient is indicated in the “goal” field by means of an integer that can take any value from 0 (no presence) to 4. Distinguishing the disease existence (values from 1 to 4) from non-existence (value 0) has been the focus of the experiments conducted in the Cleveland database [Blake and Mertz 2004]. This dataset contains the various attributes such as Id, age, gender, CP, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal and num.

5.3 Performance Evaluation Metrics

The standard performance metrics used in this work for evaluating the proposed disease prediction and monitoring system. The True Positive (TP) denoted that the patient affected with disease and identified correctly, the False Positive (FP) indicated that the patient not affected with disease but predicted wrongly, the True Negative (TN) represented that the patient is not affected by the disease but test result is wrongly predicted and the False Negative (FN) mentioned that the patient is not affected with disease and the result is also wrongly predicted. These are all used to calculate the specificity, sensitivity and prediction accuracy. The Sensitivity is nothing but the capability to identify the patients with disease correctly. The Specificity is the capability to identify the patients those are not affected by the disease. The classification accuracy is calculated by using the sensitivity and specificity in this work.

\[
Sensitivity = \frac{TP}{TP+FN} \times 100
\]

\[
Specificity = \frac{TN}{TN+FP} \times 100
\]

\[
Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100
\]
These are all the evaluation parameters used for testing the performance of the proposed disease monitoring system.

5.4 Experimental Results

The classification accuracy of the newly proposed heart disease monitoring system is shown in figure 2. Here, the various experiments have been conducted with two different heart datasets such as data streaming and standard heart disease datasets. Moreover, the heart disease monitoring system uses full dataset and the reduced feature dataset. Figure 2 shows the prediction accuracy over the two different heart datasets with the use of full dataset.

![Figure 2. Prediction Analysis with All Features](image)

From figure 2, the efficiency of the proposed heart disease monitoring system which is the combination of newly proposed feature selection algorithm and the standard DBN. Here, the full features of the two datasets were used in this work for carried out the experiments. The reason for the performance enhancement is the use of DBN.

Figure 3 shows the heart disease prediction accuracy on the two different heart disease datasets of reduced feature sets for proposed model by conducting five different experiments with the consideration of accuracy, sensitivity and specificity.
From figure 3, it shows the efficiency of the proposed model which is applied in the proposed disease prediction system in terms of prediction accuracy on reduced datasets. The prediction accuracy of the proposed model is better than other classifiers over the reduced datasets.

Table 1 shows that the comparative analysis based on the performance of proposed model over the all the features and the reduced features in terms of prediction accuracy. Here, the two different kinds of heart disease datasets were used in this work for comparative analysis by considering the prediction accuracy of full dataset and reduced dataset.

**Table 1. Comparative Analysis**

| Datasets               | Prediction Accuracy (%) |          |          |
|------------------------|-------------------------|----------|----------|
|                        |                         | Full Dataset | Reduced Dataset |
| UCI Heart Disease Dataset | 85.4                   | 86.9     |          |
| Streaming Dataset      | 93.23                   | 95.12    |          |

From Table 1, the prediction accuracy of proposed model over the reduced featured two Heart disease datasets are better when compared with the other benchmark full featured data sets including streaming and standard datasets.

The computational time analysis of the proposed disease prediction system is shown in Table 2. It shows the time taken for training and testing for the various medical datasets like two different heart disease datasets. Here, the equal numbers of records have been considered for time analysis.
Table 2. Computational Time Analysis

| Dataset           | Time Taken (Sec) for Proposed Model | Time Taken (Sec) for FTCM |
|-------------------|-------------------------------------|----------------------------|
|                   | Training   | Testing  | Training | Testing |
| UCI Heart Dataset | 0.44       | 0.25     | 0.41     | 0.21    |
| Streaming Dataset | 0.41       | 0.21     | 0.44     | 0.25    |

From table 2, it can be observed that the time taken for performing training and testing processes over the two different heart disease datasets such as streaming and standard heart disease datasets in better and optimal than other classifier called FTCM. The proposed model takes less time only for training and testing processes than FTCM classifier on all the datasets.

Figure 4 shows the performance comparative analysis over the heart disease dataset based on the accuracy of proposed disease prediction system and the domain expert opinions. Here, the various numbers of patient records have been considered for performing the comparative analysis over the heart disease dataset. Moreover, the patient records are considered only that are extracted from the streaming data and UCI machine learning repository. Moreover, different types of heart diseases such as Coronary artery and vascular disease, Heart rhythm disorders (arrhythmias), Structural heart disease and heart failure that are affected patients records have been considered in this work.

![Figure 4. Accuracy analysis between the Proposed Model and the domain expert](image-url)
From figure 4, it can be seen that the proposed model performs well in terms of prediction accuracy than the domain expert. The prediction accuracy over the patient records goes down and stabilized at same percentage of accuracy while increasing the number of patient records from 1000 records. At the same time, the proposed model provides better result than domain expert even the number of records cross the 1000 records. This is because the efficiency of the proposed model has provided this result.

Figure 5 shows the performance based comparative analysis of the proposed IoT enabled health monitoring system with the consideration of different heart disease affected datasets that are collected through IoT devices and also compared with the existing prediction systems including Temporal convolutional neural network [2], new multimodal data-based recurrent convolutional neural network (MD-RCNN) [12], Fuzzy Temporal Cognitive Map (FTCM) [3] and Fuzzy Rule based Classifier [27].

Figure 5 shows the performance based comparative analysis of the proposed IoT enabled health monitoring system with the consideration of different heart disease affected datasets that are collected through IoT devices and also compared with the existing prediction systems including Convolutional neural networks (Sandhiya & Palani 2020), Fuzzy Rule based Classifier (Kanimozhli et al 2019), Vivekanandan and Sriram (2017) and Temporal Fuzzy Min-Max-Particle Swarm Optimization (TFMM-PSO) (Ganapathy et al 2014).

![Figure 5. Performance Analysis](image.png)
From figure 5, it is seen that the performance of the proposed health monitoring system with prediction accuracy is high and it is almost equal to the sensitivity and the specificity. All the experiments are also show better performance in terms of prediction accuracy with the help of sensitivity and specificity. This is due to the fact that the better performance of the proposed model is high due to the use of streaming data and enabled the IoT, feature selection and DBN.

6. CONCLUSION AND FUTURE WORK

A new heart disease monitoring system has been proposed and implemented in this work with the incorporation of Internet of Things and Deep Learning technique for safeguarding the patient detail as records. Here, a new feature selection algorithm is incorporated for performing better classification through deep learning algorithm. In this proposed heart disease monitoring system monitors the disease level according to the inputs that are collected through IoT devices. Moreover, it classifies the patient details according to the heart disease types and the severity in this work. Finally, it gives alarm/message to the patients according to the type of heart disease.

Compliance with Ethical Standards

Funding: No funding for this research work

Conflict of Interest: There is no conflict of interest

Ethical Approval: This article does not contain any studies with animals performed by any of the authors.

REFERENCES

1. Hui Li, Xiaoyi Li, Murali Ramanathan, Aidong Zhang, “Identifying informative risk factors and predicting bone disease progression via deep belief networks”, Methods, Vol. 69, pp. 257–265, 2014.
2. Sandhiya S, Palani U, “An effective disease prediction system using incremental feature selection and temporal convolutional neural network”, Journal of Ambient Intelligence and Humanized Computing, 2020. https://doi.org/10.1007/s12652-020-01910-6.
3. R Sethukkarasi, S Ganapathy, P Yogesh, A Kannan, "An Intelligent Neuro Fuzzy Temporal Knowledge Representation Model for Mining Temporal Patterns", Journal of Intelligent & Fuzzy Systems, Vol. 26, No. 3, pp. 1167-1178, 2014.

4. S Ganapathy, R Sethukkarasi, P Yogesh, P Vijayakumar, A Kannan, "An Intelligent Temporal Pattern Classification System using Fuzzy Temporal Rules and Particle Swarm Optimization", Sadhana, Vol. 39, No. 2, pp. 283-302, 2014.

5. Ahmed M. Abdel-Zaher, Ayman M. Eldeib, “Breast cancer classification using deep belief networks”, Expert Systems With Applications, Vol. 46, pp. 139–144, 2016.

6. M. Sharawi, H. M. Zawbaa, E. Emary, H. M. Zawbaa and E. Emary, "Feature selection approach based on whale optimization algorithm," 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI), Doha, 2017, pp. 163-168, doi: 10.1109/ICACI.2017.7974502.

7. Mafarja, Majdi & Mirjalili, Seyedali. (2017). Whale Optimization Approaches for Wrapper Feature Selection. Applied Soft Computing. 62. 10.1016/j.asoc.2017.11.006.

8. Sayed, G.I., Darwish, A. & Hassanien, A.E. A New Chaotic Whale Optimization Algorithm for Features Selection. J Classif 35, 300–344 (2018). https://doi.org/10.1007/s00357-018-9261-2

9. Zheng, Yue-Feng & Li, Ying & Wang, Gang & Chen, Yu-Peng & Xu, Qian & Fan, Jia-Hao & Cui, Xue-Ting. (2018). A Novel Hybrid Algorithm for Feature Selection Based on Whale Optimization Algorithm. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2879848.

10. Chao Li, Xiangpei Hu, Lili Zhang, "The IoT-based Disease Monitoring System for Pervasive Healthcare Service", Procedia Computer Science, Vol. 112, pp. 2328-2334, 2017.

11. Jun Qi, Po Yang, Atif Waraich, ZhikunDeng, Youbing Zhao,Yun Yang, "Examining sensor-based physical activity recognition and monitoring for healthcare using Internet of Things: A systematic review", Journal of Biomedical Informatics, Vol. 87, pp. 138–153, 2018.

12. Yixue Hao, Mohd Usama, Jun Yang, M. Shamim Hossain, Ahmed Ghoneimb, "Recurrent convolutional neural network based multimodal disease risk prediction", Future Generation Computer Systems, Vol. 92, pp. 76-83, March 2019.

13. S.Sandhiya and U.Palani, “A Novel hybrid genetic binary cuckoo optimization algorithm for feature selection”, International Conference on Recent trends in science, engineering and management (ICRTTSEM), 2019.
14. T. Vivekanandan, N. Ch Sriman Narayana Iyengar, "Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease", Computers in Biology and Medicine, Vol. 90, pp. 125–136, 2017.
15. Jun Huang, Guorong Li, Qingming Huang, Xindong Wu, "Joint Feature Selection and Classification for Multilabel Learning", IEEE Transactions on Cybernetics, Vol. 48, No. 3, pp. 876-889, 2018.
16. Mahdieh Labani, Parham Moradi, Fardin Ahmadizar, Mahdi Jalili, "A novel multivariate filter method for feature selection in text classification problems", Engineering Applications of Artificial Intelligence, Vol.70, pp. 25–37, 2018.
17. Fajr Ibrahem Alarsan and Mamoon Younes, "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms", Journal of Big Data, Vol.6, No.81, pp. 1-15, 2019.
18. Meng Lu, "Embedded feature selection accounting for unknown data heterogeneity", Expert Systems with Applications, Vol. 119, pp.350–361, 2019.
19. Saurabh Kumar Srivastava, Sandeep Kumar Singh, Jasjit S. Suri, "Effect of incremental feature enrichment on healthcare text classification system: A machine learning paradigm", Computer Methods and Programs in Biomedicine, Vol. 172, pp. 35-51, 2019.
20. Yu Xue, Bing Xue and Mengjie Zhang, "Self-Adaptive Particle Swarm Optimization for Large-Scale 1 Feature Selection in Classification", ACM Transactions on Knowledge Discovery from Data, Vol. 13, No. 5, Article 50, pp. 1-31, 2019.
21. Mohamed Abdel-Basset, Doaa El-Shahat, Ibrahim El-henawy, Victor Hugo C. de Albuquerque, Seyedali Mirjalili, "A new fusion of grey wolf optimizer algorithm with a two-phase mutation for feature selection", Expert Systems With Applications, Vol. 139, pp. 1-14, 2020.
22. Shreshth Tulia, Nipam Basumatarya, Sukhpal Singh Gill, Mohsen Kahania, Rajesh Chand Arya, Gurpreet Singh Wander, Rajkumar Buyya, "HealthFog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments", Future Generation Computer Systems, Vol. 104, pp. 187-200, March 2020.
23. Boser, B. E., Guyon, I. M., & Vapnik, V. N., “A training algorithm for optimal margin classifiers”, In Proceedings of the 5th annual ACM workshop on computational learning theory, pp. 144–152, 1992.
24. Bengio, Y., Courville, A., & Vincent, P., “Representation learning: A review and new perspectives”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, pp. 1798–1828, 2013.
25. Hinton, G.E.(2009b). Deep belief networks. <http://www.scholarpedia.org/article/Deep_belief_networks> Accessed 03.09. 14.
26. Hinton, G. E., Osindero, S., & Teh, Y.-W., “A fast learning algorithm for deep belief nets”, Neural Computation, Vol. 18, pp. 1527–1554, 2006.
27. Kanimozhi U, Ganapathy S, Manjula D, Kannan A, "An Intelligent Risk Prediction System for Breast Cancer Using Fuzzy Temporal Rules", National Academy Science Letters, Vol. 42, No. 3, pp. 227-232, 2019.
28. “Types of Heart Disease”, Heart and Stroke foundation. file:///C:/Users/USER/Desktop/Sandhiya_SCOPUS/R_Types%20of%20heart%20disease%20%E2%80%93%20Heart%20and%20Stroke%20Foundation.pdf
29. Maryam I. Al-Janabi, Mahmoud H. Qutqut, Mohammad Hijjawi, "Machine learning classification techniques for heart disease prediction: a review", International Journal of Engineering & Technology, Vol.7, No.4, pp. 5373-5379, 2018.
30. Simon Fong, Raymond Wong, and Athanasios V. Vasilakos, "Accelerated PSO Swarm Search Feature Selection for Data Stream Mining Big Data".