Print ISSN 2710-0952 Electronic ISSN2790-1254

تحسين أداء أنظمة الرعاية الصحية باستخدام إنترنت الأشياء والذكاء الاصطناعي.

الاسم: مينا عادل الحصونة
العمل: جامعة الامام جعفر الصادق /فرع ذي قار
الايميل : mena_adel@ijsu.edu.iq
الاسم : حسين منخي جبر
العمل: مديرية تربية ذي قار، وزارة التربية، العراق
الايميل : cbook586@gmail.com
الاسم : خلف زياد حسين
العمل: جامعة الامام جعفر الصادق /فرع ذي قار
الايميل : Khalaf.zeyad@ijsu.edu.iq

الملخص : يُشكل الطلب على رعاية طبية دقيقة وسريعة، بالإضافة إلى محدودية الموارد وارتفاع الأسعار، تهديدًا متزايدًا لأنظمة الرعاية الصحية. يوفر الذكاء الاصطناعي، جنبًا إلى جنب مع إنترنت الأشياء، طريقة ثورية لتحسين أداء أنظمة الرعاية الصحية. تُقدم مناهج الذكاء الاصطناعي أدوات فعّالة لتقييم البيانات واسعة النطاق، والتنبؤ بالمشكلات الصحية، وتُسهّل الأجهزة المتصلة بإنترنت الأشياء اتخاذ القرارات المتعلقة بعلاج المرضى في الحياة الواقعية، من خلال الربط بين المعدات الطبية والمراقبة المستمرة للمريض. ولزيادة كفاءة الرعاية الصحية، وخفض تكاليف التشغيل، وتحسين جودة رعاية المرضى، تقترح هذه الدراسة إطارًا شاملًا يمزج بين التحليلات القائمة على الذكاء الاصطناعي وجمع البيانات المدعوم بإنترنت الأشياء. تستخدم هذه الورقة البحثية تقنيات معالجة البيانات في الوقت الفعلي، والتحسين، والنمذجة التنبؤية لمعالجة مشكلات مهمة مثل الكشف المبكر عن الأمراض، ومراقبة المرضى، وتخصيص موارد المستشفيات. تُظهر الدراسة المقارنة والنتائج التجريبية أن الاستراتيجية المقترحة تُحسّن أداء نظام الرعاية الصحية بشكل كبير مقارنةً بالتقنيات التقليدية.

كلمات مفتاحية : أنظمة الرعاية الصحية؛ إنترنت الأشياء؛ تقنيات التعلم العميق؛ الذكاء الاصطناعي

Mina adel al-hassona
imam Ja'afar Al-sadiq University
mena_adel@ijsu.edu.iq
Hussein M.Jebur
Education Directorate of Thi-Qar, Ministry of Education, Iraq
cbook586@gmail.com
Khalaf Zeyad Hussein
imam Ja'afar Al-sadiq University
Khalaf.zeyad@ijsu.edu.iq

Optimizing the performance of healthcare systems using IoT and Artificial Intelligence.

Abstract

((The requirement for precise and prompt medicalcare, resource limitations, and growing prices are posing a growing threat to healthcare systems.Artificial IntelligenceCombined with Internet Of Things provides a revolutionary way to maximize healthcare systems' performance. AI approaches offer strong tools for evaluating large-scale data, forecasting health concerns, device connected by the internet of things make decision about patient treatment easier in real life connection between medical equipment, and ongoing patient monitoring. In order to increase healthcare efficiency, lower operating costs, and improve patient care quality, this study suggests a holistic framework that blends AI-driven analytics with IoT-enabled data collecting)). This paper uses realtime, data processing, optimization, techniques, and predictive modeling to tackle important problems like early disease identification, patient monitoring, and hospital resource allocation. Comparative study and experimental results show that the suggested strategy considerably improves healthcare system performance when compared to conventional techniques)).

Keywords: Healthcare Systems; IoT; Predictive Modeling; Deep Learning Techniques; Artificial Intelligence(AI)

1. Introduction:-

IoT is a conception that reflects a collection of connected device (Including humans, Objects, at any time, any Location, Service, and Network). IoT is a new technology able to influence the entire group of businesses and, using the current infrastructure The internet, can be considered as a unique link of smart objects and services that are identifiable . The important of IoT devices typically involve advanced connectivity with services in a machine-to-machine (M¹2M) type of communication.

Therefore, the use of automation in any field is imaginable. Essentially, IoT is a promising technology for accessing health services and can improve medical services. [1] Recently, efforts have been made to design intelligent electronic medical systems for IoT-based applications. IoT, which is mainly used to connect objects worldwide through multiple devices, can be used in healthcare monitoring systems primarily for quick and large-scale access to information. IoT is expected to provide various healthcare services, each of which offers a set of healthcare solutions. IoT provides suitable solutions to many of applications, such as Traffic Congestion, Smart Cities, Organizational Health, Security, Retail, Waste Management, Industrial Control, Healthcare monitoring, Emergency Services, supplies, and Fitness Programs. In more sophisticated cases, IoT devices Monitor and, drugs, and supervise the treatment process, their amounts [2]. Moreover, IoT programs have been designed to allow Physicians to Monitor Patients after they are leaved from the hospital. One of the most important uses of IoT in the healthcare section is to Monitor Vital Signs and specific parameters of patients with chronic diseases such as Parkinson's disease, respiratory diseases, diabetes and heart disease. [3] While having cameras in the home can increase privacy concerns, we believe that this security can be threatened as any resident can turn them off at any time. It should be noted that the use of cameras and image processing can be less intrusive as there is no need for special requests for help from individuals regarding any specific case.To reduce costs and increase patient care, the use of smart devices is increasing day by day. The human tendency to spend daily life without worry and assurance of surveillance, as well as the reluctance to spend long days in hospitals and even some patients to live in small towns or villages without any advanced medical facilities, are factors that use health smart tools. Different medical sensors, devices, diagnostic devices and imaging can be considered as things that make up a significant portion of the Internet of Things or Smart devices. IoT-based healthcare systems are expected to reduce user costs, improve quality of life, and enrich user experience. From the perspective of healthcare providers, IoT has the potential to reduce equipment downtime by providing remote readiness. Additionally, IoT can identify optimal times for replacing various equipment devices for proper and continuous operation. Furthermore, IoT provides a tool for effectively scheduling limited resources and ensuring better service and care for patients [3].

Machine learning are extensively used in This field to develop analytical Models that are integrated, to various healthcare, clinical decision and services support systems, .These models primarily analyze collected data from sensor devices, and other sources, to identify behavioral patterns and clinical conditions of patients. For instance, these models analyze data to identify

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

habits,and abnormalities and improvements, in, the patient's daily routine works are change in eating and sleep ,movement drinking habits eating and digestive patterns. bulit on these patterns, healthcare applications and clinical decision support systems recommend lifestyle changes, treatment plans, and special care for the patient. Physicians and caregivers can also participate in the care planning process to confirm lifestyle recommendations. However, when using machine learning in this field, there is uncertainty and a gray area. lifestyle data, behavioral, and Clinical are very sensitive in nature, and there may be various types of bias in the data collection and interpretation process. The training data model can have an outdated version of the dataset. All of these can lead to the system making incorrect decisions without the user's knowledge. The Internet of Things (IoT) has become an interesting topic in the field of technology research. Essentially, it involves connecting devices to each other through the internet. Alongside its general applications in self-driving cars and smart homes, some of the most notable uses of IoT technology are in the field of healthcare monitoring. Advanced artificial intelligence techniques, besides Machine learning techniques, orgnize style related to health cases and diseases. Currently, Scientific community is focused on enhancing Internet Of Things based applications by integrating the Blockchain technology with machine learning techniques models to advantage from drug tracking, infectious disease tracking and medical report management, etc.

To date, modern techniques such as various machine learning-based attempts in IoT applications have been presented [4]. This paper proposes an artificial intelligence-based approach in healthcare systems to diagnose heart diseases. The proposed method uses data collected from embedded sensors in the hospital environment to create medical diagnosis patterns for heart disease. The proposed method employs several supervised methods, including neural networks , k-nearest neighbors(knn), support vector machines, decision trees, and random forests to classify heart patients and extract patterns for predicting patients at risk of heart disease. Finally, a comparison between supervised methods will be made to find the most effective method for diagnosing heart disease. The contributions of this work can be summarized as follows:

1-We have developed an artificial intelligence database based on data collected by the Internet of Things to improve the performance of healthcare devices.

2-The proposed system can continuously monitor patients through IoT sensors and uses artificial intelligence algorithms to detect abnormalities, predict potential health risks, and provide appropriate solutions to the medical condition in a timely manner.

3-This study presents intelligent optimization models for allocating hospital resources, including patient appointment scheduling, bed management, staff scheduling, and equipment utilization, thereby reducing operational costs and improving service delivery at lower prices and in faster times.

This Paper is explains the following:

Section2 reviews the related, work at Artificial Intelligence and Internet of Thing in healthcare optimization. Section3 display the proposed framework including the system architecture IoT-based data acquisition, and AI-driven analytics. Section4 describes the experimental setup, and evaluation metrics. Also this section discusses the results and provides a comparative analysis with existing, approaches. Section5 the final section show the conclution of thestudy and outlines potential directions for future work

2. Related work

((New prospects for enhancing healthcare systems have been made possible by the quick development of digital technologies, especially through the combination of artificial intelligence and the Internet of Things. While AI offers clever tools for evaluating big and complicated information to support diagnosis, prediction, and decision-making, IoT allows for seamless data collection from patients and medical devices as well as continuous monitoring. IoT and AI together provide a more potent framework for improving patient outcomes, cutting operational costs, and optimizing healthcare performance, despite the fact that numerous studies have examined these technologies separately)). In the sections that follow, we examine previous studies that emphasize the function of AI and IoT in healthcare systems.

In [5], A method called exploratory and duplex deep neural network (EH-DDNN) has been proposed for early prediction of diabetes. First, the EH-PCA algorithm is presented for feature selection using large datasets as input. EHP divides the input data matrix into rows and columns separately. Using the concept of conditional heteroscedasticity evaluation and exploratory techniques, EHP utilizes neighborhood evaluations in subregions while reducing communication time and overhead, thus strongly linking calculations. This in turn removes irrelevant and redundant features, making it easier to predict initial features. Then, a deep duplex neural network (DDNN) is designed for initial prediction with inherent features that analyzes a combination of nonlinear processing and linear response to store large amounts of data. Experiments and validation are performed on benchmark datasets, diabetes datasets from the UCI repository, and Pima Indians diabetes datasets. Comparative analysis of diabetes prediction time, diabetes prediction overhead, and ROC curve analysis are performed.

In [6], A proposed method combines two deep learning architectures, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). CNN is used for detecting lesions in retinal fundus images, and LSTM is used for generating descriptive sentences based on those lesions. In the training and testing process, the output of CNN is used as input to LSTM. The goal of the process is to train a model that can describe retinal fundus images in one sentence. The results of this study using the MESSIDOR dataset show an accuracy of about 90%.

((In [7] In this paper, we propose a deep neural network (DNN) combined with redundant feature elimination (RFE) to provide initial prediction of diabetic retinopathy (DR) based on several individual risk factors obtained from the Internet of Things. The proposed model utilizes irrelevant feature elimination and redundant feature elimination (RFE) to more clearly classify and identify the disease. Data from a cohort of several patients were used to predict DR in the initial stages of the proposed model and several existing models. The proposed model achieved an accuracy of 82.033%, which outperformed the current models significantly)). Therefore, important risk factors for retinopathy can be successfully extracted using RFE. Additionally, to evaluate the robustness and generalization of the proposed prediction model, it was compared with other models and machine learning datasets (neuropathy and hypertension-diabetes). The proposed prediction model helps improve the diagnosis of retinopathy in the early stages based on individual risk factors.

In [8], a new DNN model called DTP is proposed for predicting the type of diabetes. The deep neural network is trained on two datasets, each containing over a thousand records. The number of epochs in the training phase is kept low to ensure the method can work quickly, even on mobile platforms. The experimental results show the effectiveness and adequacy of the proposed DTP model. The best result achieved for the diabetes dataset was 94.02174% accuracy, and for the Pima Indians dataset was 99.4112% accuracy.

In [9], the goal in is to assist physicians in predicting diabetes built on current Health indicators such as age ,blood plasma, pregnancy, insulin level,skin thickness, body mass index, blood pressure , and pediatric conditions . Machine learning tool is having a significant impact on the field of medicine. Predicting future diseases accurately can be challenging, this project uses an efficient CNN system for the classification. This classification is performed as non-diabetic or diabetic based on health indicators. The results have an accuracy of 84%, and the accuracy level can be further improved by using deeper neural networks, which is a step towards further improvement for this project. The multilayer perceptron neural network is an algorithm used for binary diabetes classification. It involves analyzing all eight parameters and determining whether the patient is diabetic or not. This computer system does not require repeated blood tests for patient prediction, saving time, cost, and increasing hospital system efficiency. A graphical user interface has been developed to extract data for analysis. The dataset used here is the Pima Indian Diabetes Dataset, which is a collection of 768 healthy patient records.

In [10], A hybrid model combining an artificial neural network (ANN) and a genetic algorithm (GA) was developed for classifying diabetic patients in [14]. The optimal number of neurons and hidden layers for designing the ANN model architecture were determined. To reduce the mean squared error (MSE) of the network and optimize the accuracy of the diagnostic system, a GA algorithm was combined with the proposed ANN model. For testing purposes, the model was applied to a dataset consisting of 768 samples for the diagnosis of type II diabetes patients among other cases. The results showed an accuracy of 85% for diagnosing type 2 diabetic patients. The proposed structure based on the lower MSE achieved the best performance of the artificial neural network with an MSE rate of 0.155.

In [11], a neural network model was developed and trained to detect gestational diabetes in pregnant women. The model is a four-layer feedforward network that was trained using backpropagation and Bayesian regularization algorithms. The input layer has 8 neurons, two hidden layers each with 10 neurons, and the output layer has one neuron which gives the diagnosis result. The developed model was also included in a web-based application to facilitate its use. Validation with regression shows that the trained network has an accuracy of over 92%.

In [12], a model is presented where smartwatch sensors and ecological sensors, optitrack cameras, are used to, image, and audio signals and collect video with special wearable devices to aggregat physiological variables.

[13],anew smart healthcare model device is proposed, which includes a pathology detection technique using deep learning technique . Pathogens can be identified from the electroencephalogram, signals of the patient. In this model, a smart EEG headset captures EEG signals and sends them to a mobile edge computing server. The server preprocesses the signals and sends them to a cloud server.

In [14], the solution is proposed for predict diffrent human activities using a Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) and wearable sensors implemented on a local fog server and accelerated ((GPU)).

[15], additional sensors are used for motion tracking and SVM and RF classification methods are utilized for predicting motion. Some recent models simulate ML approaches for analyzing physiological data in portable sensors. However, there are issues in predicting anomalies in physiological variables in terms of computational flow structure. In this study, Hierarchical Temporal Memory (HTM) is implemented in a distributed manner. The model is implemented on edge nodes and used for inference.

In [16], a solution for fall prediction based on the LSTM RNN method is presented, which is executed at the edge level. The computation performance of the edge access method is defined along with a case study on electroencephalogram (EEG) data.

Even while machine learning and deep learning have made great strides in the field of disease prediction, the majority of current research is dataset-specific, unscalable, and ignores system-level issues like resource optimization, data security, and real-time monitoring. Despite its potential, IoT-enabled healthcare systems frequently struggle with computing efficiency, privacy protection, and predictive intelligence integration. This paper fills these gaps by presenting an integrated IoT-AI framework that enhances clinical decision-making and the performance of the healthcare system as a whole by fusing safe storage, intelligent optimization, and real-time data collection.

3. Proposed Method

As mentioned, in this study, an artificial intelligence-based approach has been proposed for cardiac disease diagnosis within healthcare systems. The proposed method utilizes data collected from embedded sensors in hospital environments to create diagnostic patterns for cardiac diseases. Several supervised methods, including neural networks, support vector machines, k-nearest neighbors, random forests, and decision trees are employed in the proposed method for classifying cardiac patients and extracting patterns to predict individuals at risk of cardiac disease. Finally, a comparison between the supervised methods will be presented to identify the most effective approach for cardiac disease diagnosis.

((One proposed method for diagnosing cardiac disease in new patients involves using a standard dataset extracted from processing electronic signals of patients in healthcare systems. This dataset contains information about the disease progression and treatment for each patient. In this approach, using classification methods based on the data related to patients in the standard dataset, the diagnosis of new patients is determined)). ((After diagnosis, similar patients are identified for each new sample.))

((For predicting cardiac disease in new patients using the available data in the standard dataset, machine learning methods such as artificial neural networks or decision trees can be employed)).

((In the neural network method, the standard dataset needs to be divided into two parts: training and testing. Then, using the training data, a neural network with an appropriate structure is created and trained to predict the cardiac disease label in new data. Afterward, the network's performance is evaluated using the testing data, and if it performs well, the network is used for predicting the cardiac disease label in new data)).

((Similarly, in the decision tree method, the standard dataset is divided into training and testing sets. Using the training data, a decision tree with an appropriate structure is created and trained to predict the cardiac disease label in new data. Then, the performance of the decision tree is

Print ISSN 2710-0952 Electronic ISSN2790-1254

evaluated using the testing data, and if it performs well, the decision tree is used for predicting the cardiac disease label in new data)).

In both methods, after diagnosing cardiac disease in new data, you can search for similar patients in the training dataset and, consequently, predict the cardiac disease label for the new sample. The overall proposed method is illustrated in Figure 1.

Medical data collection module

Disease diagnosis module

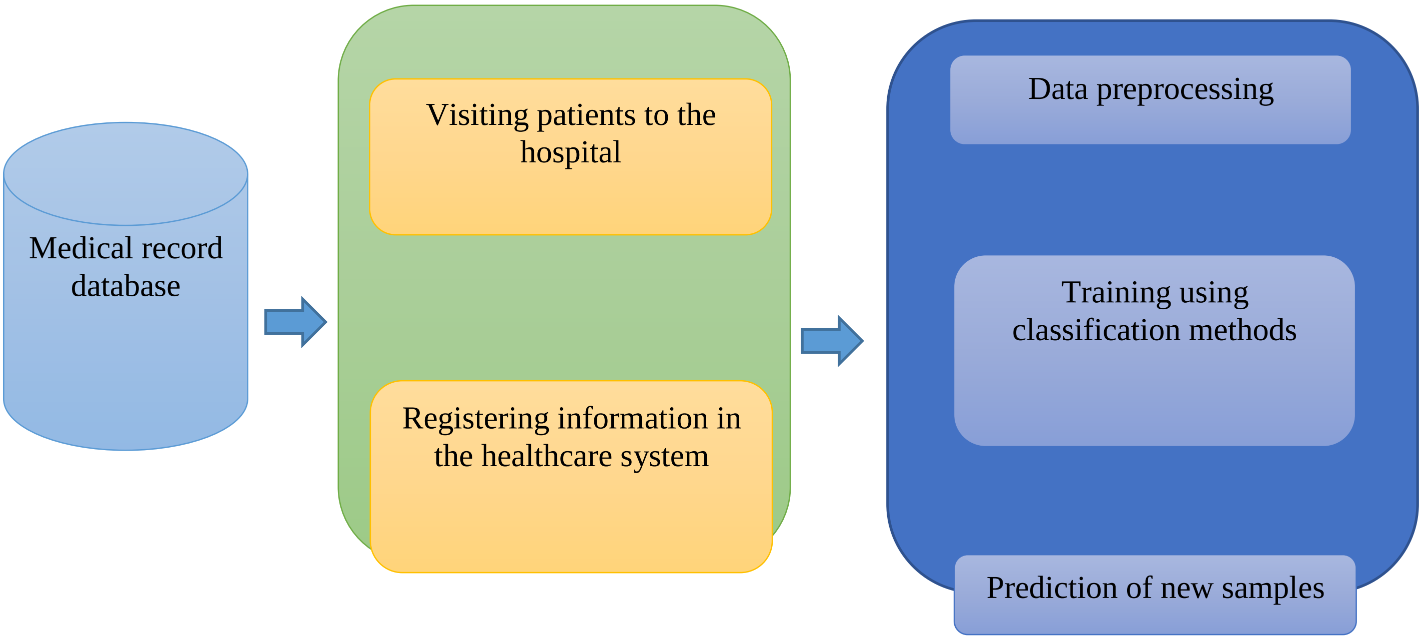

Figure 1 general process of the proposed method

As shown in Figure 1, the proposed method consists of two modules: the data collection module and the disease diagnosis module. We will now elaborate on the methods used in each of these modules.

3.1. Data Collection Module

As mentioned, this paper presents a healthcare framework within the Internet of Things (IoT) based healthcare systems. In the proposed method, smart devices are connected to healthcare systems in the IoT network. Equipped with sensors, these smart devices collect patient-related data and send it to the smart gateway. The smart gateway performs data Collection, inconsistency checks, and data updates before sending it towards the storage unit. The storage unit securely stores the data in the medical record database of the healthcare system. Ultimately, end-users can retrieve and access the stored data by sending their queries to the storage unit.

The proposed method leverages IoT technology to enhance the efficiency of healthcare systems and reduce diagnostic errors based on the data collected from the healthcare system.

The proposed healthcare framework consists of a web-based application with two interfaces: the front-end interface connecting to patients and the back-end interface providing internal communication through the database. A specific request acts as a link between these two interfaces. The proposed healthcare framework offers a user-friendly web-based communication between patients and providers.

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

In the process of communicating with the database, nodes connect to each other as a system. These nodes can include servers or identity verification nodes, while others are executing nodes. In this process, servers' responsibility is to validate or reject transactions to ensure database status is committed or rolled back. In other words, servers are responsible for validating transactions in the network. However, the executing nodes' task is to verify transactions accumulated in the servers. If the transaction is valid, the executing node increases the server's value and sends the transaction to the database. As a result, after validation in the servers and executing nodes, transactions are stored in the database and become part of the history of all transactions in the database. Figure 2 illustrates the architecture of the proposed method.

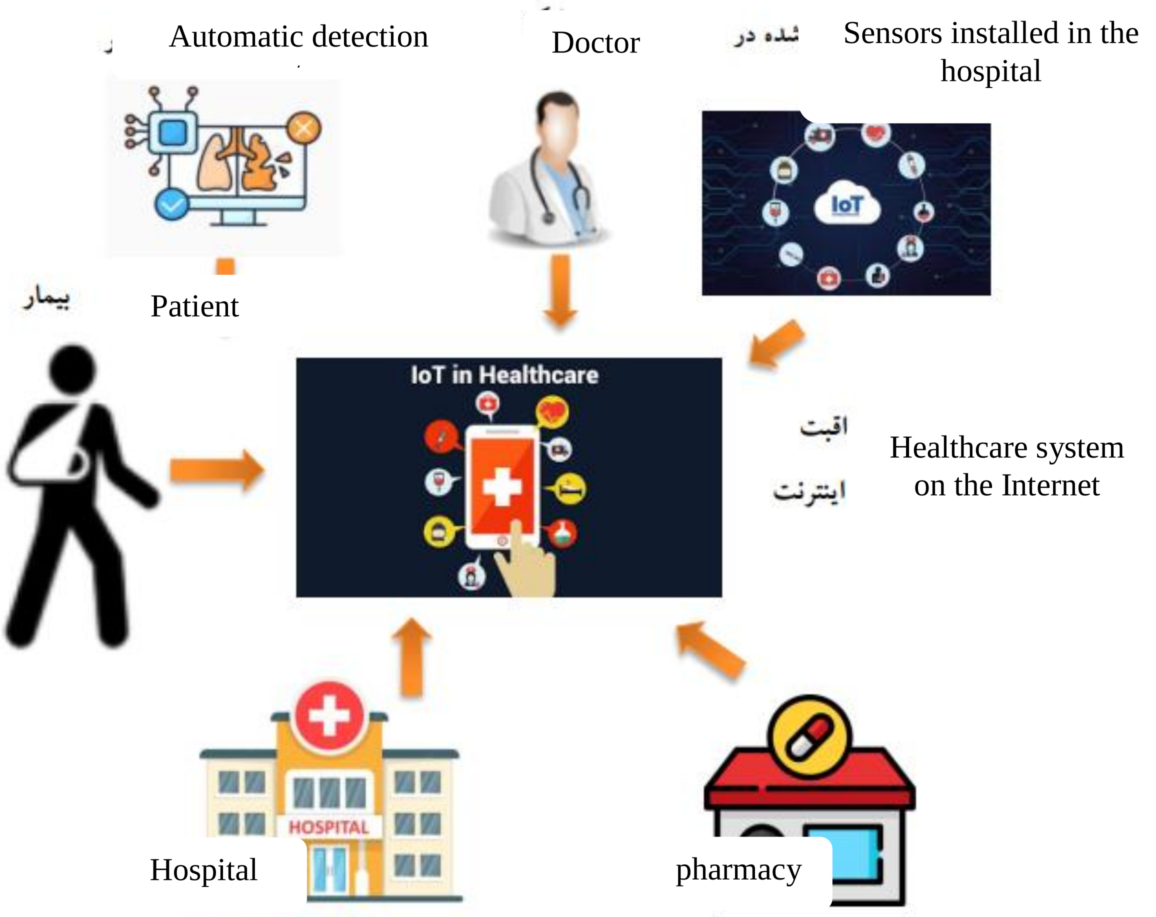

Figure 2 The proposed method Architecture

Use of the IoT in Healthcare system can assist in recording patients' health-related data and monitoring their status. For instance, IoT can be utilized for the real-time monitoring of patients' physical condition, such As blood pressure, body temperature,heart rate which can be recorded in the database.

Moreover, IoT can be effective in disease detection and prevention. By using smart sensors, environmental conditions can be monitored, and potential health hazards for patients can be identified. For example, if air pollutants increase, the IoT system can automatically raise an alert and implement preventive measures if necessary.

Using IoT, it is also possible to automatically record information like the patient's medical history, medication usage, etc., in the database and make this information accessible to physicians and other medical staff. This can significantly enhance the effectiveness of disease diagnosis, treatment, and follow-up.

3.1.1. Registration of Patient Information

In the field of healthcare, whenever a patient visits a doctor, it is possible that after diagnosis, the physician may recommend specific tests or prescribe specific medications to be purchased from specific medical stores. However, the patient may choose to change their doctor and start a different treatment.((In such cases, it is necessary for all reports, laboratory test results, and prescribed medications to remain securely recorded in the blockchain. This ensures that both the patient and the doctor have accurate information about the treatment costs and the expenses incurred for a specific treatment. In other words, the use of blockchain in healthcare provides patients with greater transparency regarding their treatment costs and also enhances the security of their medical data)). ((Indeed, utilizing blockchain in the healthcare and medical sector can contribute to improving transparency and credibility of treatments and their associated costs)). ((By maintaining all records and relevant information related to a patient's treatment in the blockchain, the possibility of fraud and data alteration is reduced, and the database remains continuously updated.))

((In the proposed method, due to the substantial volume of patient-related data, two scenarios are used to store this information in the database. In the first scenario, patient-related information, including data, physician's diagnosis, and results from the automated diagnostic system, are stored in the database)). In this scenario, patient-related information is updated in real-time in the blockchain and can quickly be accessed by physicians and other medical staff.

((In the second scenario, information and status related to medical centers and pharmacies are stored in the database. This information can include drug lists, prices, order history, and pharmaceutical inventory. By maintaining this information in the database, continuous awareness of drug inventory, market prices, and medical orders can be achieved, improving the distribution process of drugs and treatment)).

((Overall, the use of blockchain in the healthcare and medical sector can enhance transparency, treatment quality, and cost reduction. This approach can aid physicians and other medical staff in clinical and managerial decision-making, ultimately leading to an improvement in the quality of patient care)).

3.1.2. The registration of information by medical centers and pharmacies

regarding the transfer of medications to specific medical stores may encounter irregularities due to various intermediaries involved in the process. However, this issue can be resolved using blockchain technology. In this method, when the pharmacy "A" sends medications to the seller company "B," each transaction and its details are recorded in a block. These blocks are directly connected to a network of interconnected computers and operate as a distributed system.

In this system, each block has a unique hash containing information about the transactions within that block. This hash is linked to the next block and continues in the same manner for all subsequent blocks. As a result, each block is sequentially linked to its previous and next blocks, and any change in a block affects all subsequent blocks, meaning that any attempt to alter past transactions would entirely disrupt the system.

Print ISSN 2710-0952 Electronic ISSN2790-1254

Therefore, blockchain technology makes all data related to medication transfers transparent and analyzable. However, for the successful implementation of this technology, all companies involved in the process must work cohesively and accurately record all information in the blockchain.

1. In the proposed process, when the company or sender "A" initiates the product shipment process, they add a new transaction "A → B" to the database and request distribution to available agents.
2. In the next stage, the seller or "B" accepts the request from "A" and adds another query with the new and previous requests to the existing database.
3. Additionally, the delivering entity or "B" transfers its product to the recipient "C," meaning that the final delivery is recorded as a new query, and the function of the previous queries comes to an end.

Indeed, in a blockchain network, each node (computer) holds a copy of the entire blockchain. This implies that every node has complete access to all blocks, including the registered transactions. These copies are accessible throughout the network without any modification or tampering, and when a new node joins the network, it receives a full version of the blockchain and synchronizes its chain with others following the network protocol.

New node joins the network, it receives a full version of the blockchain and synchronizes its chain with others following the network protocol.

In this network, all nodes synchronize with each other to ensure that everyone has a complete and identical version of the blockchain. To do this, each node continuously and independently checks for new blocks that are added to the network.

3.2. Disease Detection Module

The proposed method consists of a detection module in which the following steps are performed based on the standard process of training data mining algorithms.

3.2.1. Data Preprocessing

We are talking about data mining classification models and the data preprocessing process in them. Each data mining classification model deals with a specific type of data and to benefit from its results, the data must be used in the model. Therefore, the data preprocessing process is different for each model.

Three examples of data preprocessing steps are:

1. Feature extraction: In this step, important features of the data are extracted. For example, in image models, features such as shape, color, and dimensions of the image are extracted

2. Data cleansing: In this stage, data that is inappropriate, unclear, or incomplete is removed. Also, correction and completion of incomplete data may be required.

3. Data Transformation: In this stage, data is converted into a format that the data mining classification model can comprehend. For example, in text-based models, textual data is transformed into numerical vectors.

These data preprocessing stages are performed to enhance the accuracy and efficiency of data mining classification models. Generally, the better and cleaner the data, the higher the accuracy and efficiency of the data mining classification models will be.

- **Missing Values**

((The removal of missing values is a crucial step in the data preprocessing process. Missing values may occur due to errors in data collection, data storage issues, or any other reasons that lead to data incompleteness. These missing values could be inherently insignificant for certain records or impossible to gather due to the absence of similar data. Eliminating missing values can be done in two ways. In the first method, rows containing missing values are removed from the data. In the second method, missing values are replaced with values such as the mean, mode, or other measures with the highest frequency.))

((Removing missing values enhances data quality and improves the performance of data mining classification models. Moreover, several solutions to address the issue of missing values include:

- Deleting records with missing values.
- Removing features with missing values from all analyses.
- Estimating the value of the record for the feature with missing values.
- Analyzing all possible values and selecting the one that yields the optimal outcome as the record value for the feature with missing values.

Removing records or features with missing values might lead to reduced accuracy and efficiency of data mining classification models. Consequently, alternative methods exist to handle missing values, including approaches where missing values are replaced with predicted values. Your proposed method of using the "?" symbol as a reasonable substitute for missing values is a commonly used and straightforward approach in data mining classification models. However, this method may result in lower accuracy, particularly if a significant percentage of data contains missing values. In such cases, more advanced methods like imputation and prediction techniques can lead to higher accuracy)).

- **Selection of Feature Subsets**

Selecting a subset of features that are relevant to the class label is a crucial step in the data preprocessing for classification tasks and is recognized as one of the main feature selection methods in data mining.

((In the case of a dataset of heart patients, choosing a subset of features related to the patients' body biochemistry can lead to the best performance in classification models. However, due to the presence of numerous biochemistry features, selecting features with non-discriminatory distributions and minimal impact on determining the class label may result in reduced accuracy and model performance. Therefore, in the data preprocessing stage, removing features with non-discriminatory distributions and less impact on the class label can significantly improve the accuracy and performance of data mining classification models)).

((The primary goal of the data preprocessing stage, including selecting feature subsets and removing irrelevant features and attributes, is to enhance the accuracy and speed of data classification. By reducing the dimensions of the data, the operational and computational complexity of the system is reduced as well. Moreover, by eliminating features and attributes with weak correlations to the class label, the classification accuracy is increased.

Furthermore, selecting a subset of data can identify implicit dependencies between the data and the class label, resulting in significant improvements in the accuracy and speed of classifying new data (test samples). This allows for the execution of classification models on new data with high accuracy and less computational time)).

((In this study, the standard deviation method is used for selecting important features. This method involves calculating the mean and standard deviation for each feature. Then, features with a standard deviation below a certain threshold are removed. Such features, with low

Print ISSN 2710-0952 Electronic ISSN2790-1254

standard deviation, do not demonstrate sufficient differentiation between normal and diseased samples and can be eliminated during data examination and preprocessing.
Using the standard deviation method in data preprocessing can help remove irrelevant features, retaining only those features that play an important role in determining the class label. Moreover, reducing the number of features enhances the speed of data classification. It is essential to consider the range of data distribution and the magnitude of standard deviation while selecting important features for disease detection. Features with a wide data distribution and higher standard deviation generally contribute more to better discrimination between normal and diseased samples. Conversely, features with a narrow data distribution and lower standard deviation may not contribute effectively to disease detection.
Therefore, it is crucial to pay sufficient attention to the range of data distribution and the magnitude of standard deviation when selecting relevant features for disease detection)).
((If we consider the standard deviation of feature data (stdj) as the standard deviation related to the feature data, maxj as the maximum value present among the data of feature j, and minj as the minimum value present among the data of feature j, then we have:
$stdj \geq \alpha \times maxj \,/\, minj$)) (1)
The value of the threshold parameter (α) in the standard deviation method controls the number of features that satisfy the condition. Larger values of α result in fewer features satisfying the condition and fewer computations for classification. In this paper, threshold value α is set one.

3.2.2. Training Using Classification Methods

Supervised classifiers are one of the machine learning techniques used for disease detection and object identification based on relevant previous data. This technique is commonly applied to labeled data. In supervised classification, the machine learning model learns automatically from labeled previous data to make decisions about new incoming data.

- **K-Nearest Neighbors (KNN)**

k-Nearest Neighbors (kNN) [17] is a machine learning algorithm used for classification and regression tasks. Its operation is as follows: given a training dataset and a new sample, it first calculates the distance between the new sample and each of the existing samples in the training set. ((Then, the K nearest samples to the new sample are selected. Here, K is a parameter that determines the number of samples chosen as neighbors. Next, based on the labels of the K nearest samples, the label of the new sample is determined)).

- **J48 Decision Tree**

((J48 Decision Tree [18] is a machine learning algorithm used for classification and regression tasks. It automatically learns from the training data and creates a decision tree that can be used to predict new data. In classification problems, the goal of this algorithm is to create a decision tree using the features of the data, so that it can predict the label of a new data point based on its features. In regression problems, the goal is to create a regression model using the data features to predict the output value for a new data point)).

Random Forest Algorithm

((The Random Forest [19] algorithm was introduced in 1995 and evolved from the random subspace method. It is a tree-based learning model used for various applications such as classification and prediction. Random Forest has gained popularity due to its ability to handle non-linear classification tasks efficiently, particularly with large datasets. It is known for being

Print ISSN 2710-0952 Electronic ISSN2790-1254

effective in dealing with class imbalance and can be used for multi-class and multi-label problems)).

Artificial Neural Network (ANN)

((Artificial Neural Networks (ANNs) [20] are inspired by biological neurons and are used to solve various problems in computer science and artificial intelligence. ANNs consist of a collection of connections between neurons and different layers of neurons that can perform parallel computational operations. Unlike following explicit rules and constraints, artificial neural networks utilize machine learning algorithms to identify hidden patterns and rules within data for prediction and decision-making. Typically, these networks include an input layer to receive inputs, multiple hidden layers to process information and extract important features, and an output layer to produce outputs. By using the relationships between processing units in the network, complex and nonlinear connections can be modeled, leading to higher accuracy in prediction and decision-making)).

Support vector machine (SVM)

((SVM [21] is a supervised machine learning technique that is widely used in various fields, including computer science, artificial intelligence, data science, and industry, due to its high accuracy and applicability in classification and regression problems. In industrial equipment Prognostics and Health Management (PdM), SVM is extensively used for identifying specific conditions based on acquired signals, such as equipment failure or performance degradation. Additionally, in the development of predictive maintenance measurements for two types of rail paths, straight and curved sections, both SVM and ANN (Artificial Neural Networks) machine learning algorithms are employed)).

4. Results and Discussion

After training the data using classification methods, new samples can be predicted using the predictive model. For this purpose, first, the feature selection method must be applied to investigate the patterns in which new samples are examined based on the selected features in previous stage. Then, evaluate accuracy of proposed method in predicting new samples, the new data is predicted using the trained predictive model and compared with the true labels of these samples.

4.1. Evaluation Metrics

“To evaluate the proposed method, confusion matrix with four parameters TP, FP, TN, and FN can be used”.

This matrix is defined as follows:

- TP (True Positive): The number of samples that are correctly identified and actually belong to the positive class.
- FP (False Positive): The number of samples that are incorrectly identified and actually belong to the negative class.
- TN (True Negative): The number of samples that are correctly identified and actually belong to the negative class.
- FN (False Negative): The number of samples that are incorrectly identified and actually belong to the positive class.

Print ISSN 2710-0952 Electronic ISSN2790-1254

Using this matrix, metrics such as accuracy, sensitivity, positive predictive value (precision), and recall can be computed to evaluate the performance of the proposed method.
"Since the confusion matrix is a standard method for measuring the classification performance in two-class data, the evaluation indices derived from the confusion matrix include accuracy, sensitivity (recall), specificity, and F1 score, which are defined in the equations below":

$$Accurace = \frac{\mathrm{TP} + \mathrm{TN}}{TP+TN+FP+FN} \tag{2}$$

$$Recall = \frac{\mathrm{TP}}{TP+FN} \tag{3}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{4}$$

$$F - measure = \frac{2*Precision*Recall}{Precision+Recall} \tag{5}$$

Evaluation metrics derived from the variables of To evaluate the performance of the proposed method and compare it with other methods, we use the confusion matrix.

4.2. Implementation of The Proposed Method

A model for building a heart disease prediction system has been presented. This model initially focuses on dimensionality reduction and selecting suitable features for the available dataset. The selected features are then used for a Convolutional Neural Network (CNN) to calculate patterns for heart disease prediction based on these features. The results indicate an acceptable accuracy for feature subset selection and heart disease classification.
The dataset used in this study consists of heart disease data from the UCI database, comprising 2800 samples. Each sample contains 76 features that can determine the class changes related to the samples. The aim of this proposal is to find potentially impactful features on class changes in the samples. Ultimately, four features from the dataset have been used in this paper, and the results demonstrate that these features are suitable for heart disease prediction. Table 1 presents the detailed findings.
Table 1, the features in the prediction data set of heart patients

المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN2790-1254

F#	F Name	F#_1	F Name_2	F#_3	F Name_4	F#_5	F Name_6	F#_7	F Name_8	F#_9	F Name_10	F#_11	F Name_12
1	id	12	chol	23	Dig	34	tpeakbps	45	restckm	56	cday	67	rcaprox
2	ccf	13	smoke	24	Prop	35	tpeakbpd	46	exerckm	57	cyr	68	rcadist
3	age	14	cigs	25	Nitr	36	dummy	47	restef	58	num	69	lvx1
4	sex	15	years	26	Pro	37	trestbpd	48	restwm	59	lmt	70	lvx2
5	painloc	16	fbs	27	diuretic	38	exang	49	exeref	60	ladprox	71	lvx3
6	painexer	17	dm	28	Proto	39	xhypo	50	exerwm	61	laddist	72	lvx4
7	relrest	18	famhist	29	Thaldur	40	oldpeak	51	Thal	62	diag	73	lvf
8	pncaden	19	restecg	30	thaltime	41	slope	52	thalsev	63	cxmain	74	cathef
9	cp	20	ekgmo	31	Met	42	rldv5	53	thalpul	64	ramus	75	junk
10	trestbps	21	ekgday	32	Thalach	43	rldv5e	54	earlobe	65	om1	76	name
11	htn	22	ekgyr	33	Thalrest	44	Ca	55	Cmo	66	om2	-	-

In this paper, the selection of appropriate features for heart disease prediction has been investigated, and this selection can have a significant impact on prediction accuracy. Feature selection can be done manually based on prior knowledge in the field of heart disease, but this method may be time-consuming and costly, and in some cases, important features may be missed due to technical knowledge limitations. Therefore, using filter-based approaches for feature selection can be a fast and effective solution. In this method, features with the highest correlation with the target classes are selected using the criterion of standard deviation. Overall, selecting appropriate features for heart disease prediction is an important and complex task that requires the combination of various methods. Proper feature selection can improve prediction accuracy, reduce data volume, and increase processing time efficiency.

4.2.1. Preprocessing

In this paper, the required dataset was collected from the UCI data repository, consisting of 76 features as presented in Table 1. However, not all of these features can be considered effective indicators for heart disease detection. Therefore, only useful features need to be selected. By using appropriate features, in addition to reducing the dimensionality of the data, the accuracy of heart disease detection increases.

In the proposed method for feature selection, criteria such as filters and standard deviation have been employed. The standard deviation includes statistical features that describe the distribution of data in the features. In the proposed method, features with a standard deviation of less than 1 are considered redundant and should be removed from the data. The remaining features are used as representative features for the original dataset in the supervised learning method.

Table 2 shows the selected features in the proposed method.

Print ISSN 2710-0952 Electronic ISSN2790-1254

Number_1	Feature name_1	value of the standard deviation	Number_2	Feature name_2	value of the standard deviation	Number_3	Feature name_3	value of the standard deviation
1	age	1.4693	10	diuretic	1.2408	19	restwm	1.3079
2	sex	1.3224	11	met	1.0705	20	thal	1.307
3	pncaden	1.1186	12	thalrest	1.094	21	cxmain	1.2548
4	trestbps	1.1095	13	trestbpd	1.1572	22	om1	20.4832
5	smoke	1.1943	14	thal	1.0188	23	lvx2	20.4756
6	cigs	1.1201	15	rldv5e	1.2142	24	lvx4	21.4832
7	fbs	1.1172	16	slope	1.3019	25	cathef	35.6817
8	dig	1.1298	17	tpeakbpd	1.4066	26	num	21.4756
9	nitr	1.2341	18	ca	1.2478	27	rcaprox	32.9491

In Table 2, only 27 features with a standard deviation higher than the threshold value have been selected for use in the chosen model. The standard deviation values indicate how much data is distributed within each feature within a range. Therefore, features with higher standard deviations will contain the most useful information about the separation of samples from the two classes.

Furthermore, in Figure 3, the bars that extend to the right of the zero axis indicate that the data distribution for that specific feature is different for the two classes. Hence, such features can be considered informative.

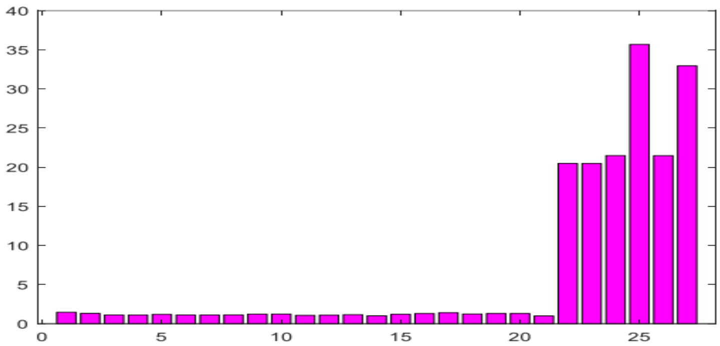

Figure 3 Standard deviation chart of selected features

Yes, in Figure 3, it can be observed that some features have higher standard deviations, indicating that the data within those features is different for the two classes. Among the 76 features available in the dataset, only 27 features with standard deviations higher than the selected threshold value have been chosen and used as inputs for the classification model. This proposed feature selection method can significantly improve the accuracy of heart disease diagnosis. In the following chapter, these selected features will be utilized for classifying patients with heart failure.

4.3. The evaluation of the proposed method

Yes, in evaluation of proposed method for diagnosing heart disease, certain metrics are used to assess the classification performance. These metrics include the confusion matrix, accuracy, sensitivity, and positive predictive value. These metrics are calculated based on comparing the predicted labels by the classifier with the true labels of the data. Among these metrics, the confusion matrix is one of the most important for evaluating the classification performance. It shows how many samples are correctly classified and how many are misclassified. By using this matrix, accuracy, sensitivity, and positive predictive value can be calculated. Additionally, accuracy is another significant metric for evaluating the classification performance, indicating what percentage of total samples are correctly classified. By considering these different metrics, the performance of the proposed method can be compared with other existing methods for diagnosing heart disease. Based on these comparisons, a decision can be made regarding the use of this method for heart disease diagnosis. In Table 3, the values of evaluation metrics for various classification methods applied in this paper on the selected subset of features based on the standard deviation are presented.

Table 3 valuesof evaluation criteria for classification methods

	RF	KNN	DT	SVM	NN
Accuracy	0.9951	0.9699	0.9552	0.9562	0.9561
Recall	0.9983	0.9856	0.9747	0.9736	0.974
Precision	0.9966	0.9833	0.979	0.9813	0.9808
F-measure	0.9974	0.9844	0.9769	0.9774	0.9774

Yes, in general, using a random forest for data classification can lead to better results compared to traditional methods due to its high learning capability and ability to extract complex features. This is particularly true for tasks related to image processing, text processing, and data classification. The appropriate selection of features can be crucial when utilizing the powerful feature learning capabilities of random forest models. It can help reduce the time and costs associated with data preprocessing. In the case of your proposed method, combining feature selection based on the standard deviation with classification techniques can significantly improve the classification results. By using this method, one can select important and effective features more accurately in diagnosing heart disease and optimally relate them to each other using the random forest. As a result, better results can be achieved in evaluation metrics, leading to higher accuracy in diagnosing heart disease. In Figure 4, the bar chart for the average accuracy of the classification methods is shown.

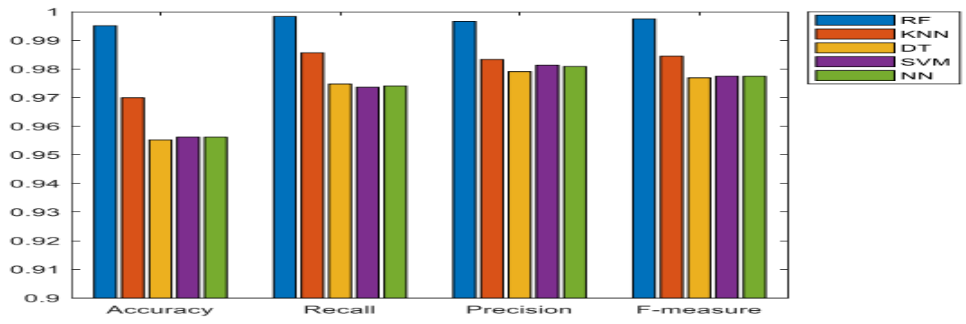

Figure 4- Bar chart Of The Average Values of Accuracy Criteria in Different Classification Methods

As seen In Figure 4, the Random Forest classification method in combination with feature selection based on standard deviation has a high capability in detecting heart diseases. Now, let's investigate each of the evaluation metrics during the process of increasing the test data.

The accuracy metric in classification tasks is the ratio of the number of samples correctly classified to the total number of samples. In the case of diagnosing heart disease mentioned in

Print ISSN 2710-0952 Electronic ISSN2790-1254

the text, to calculate accuracy, the number of samples correctly classified as diseased and the number of samples correctly classified as healthy are divided by the total number of samples.
In Figure 5, the accuracy metric is compared based on the selected features in the proposed method with other classification methods.

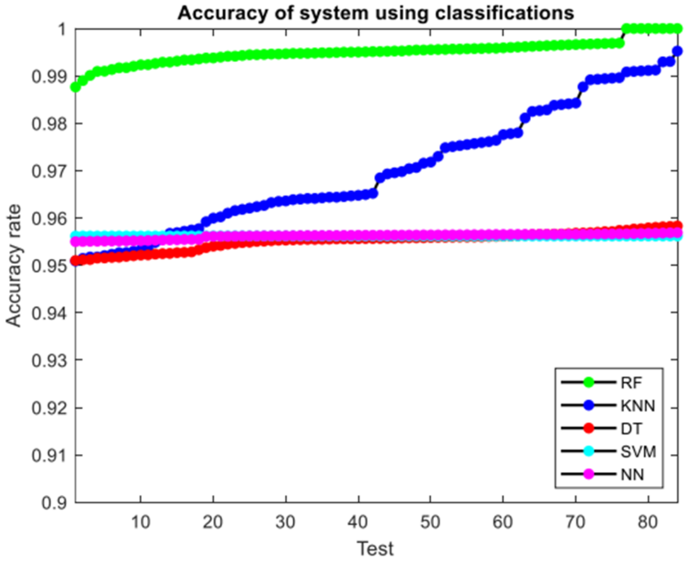

Figure 5 Comparison of accuracy of classification methods based on features with high standard deviation

based on figure 5, It Is Evident that the proposed method, using the combination of feature selection based on standard deviation and Random Forest, achieved better results in the accuracy metric compared to other methods. Therefore, using the proposed method as an effective and accurate approach in diagnosing heart disease can be considered. Regarding Figure 5, it appears that there are approximately 840 test samples, and the accuracy values from every 10 tests are shown on the graph. This arrangement was done to prevent overcrowding of data points on the accuracy and other evaluation metric graphs, enhancing the clarity of the images related to accuracy and other evaluation metrics.
The sensitivity metric, also known as the disease identification metric, is the ratio of the number of samples correctly identified as diseased to the total number of diseased samples. This metric is crucial for assessing the system's ability to identify diseased samples. In Figure 6, the sensitivity metric for different classification methods is displayed.

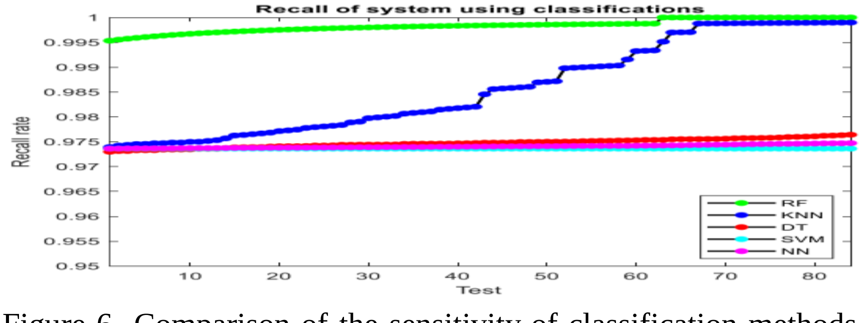

Figure 6- Comparison of the sensitivity of classification methods based on features with high standard deviation

Based on Figure 6, the proposed method, using the combination of feature selection based on standard deviation and Random Forest, achieved better results in the sensitivity metric compared to other methods. Considering the sensitivity metric, the proposed method can be used as an accurate and reliable approach in diagnosing heart disease. However, to ensure the accuracy and applicability of this method in real-world scenarios, further evaluation and more detailed analysis of results under different conditions and with a larger number of samples are needed.

The accuracy or overall accuracy metric is the ratio of the number of samples correctly classified to the total number of samples. This metric indicates how many samples are correctly classified, not only with respect to diseased and non-diseased individuals. In Figure 7, the graph related to the precision metric for different classification methods is shown.

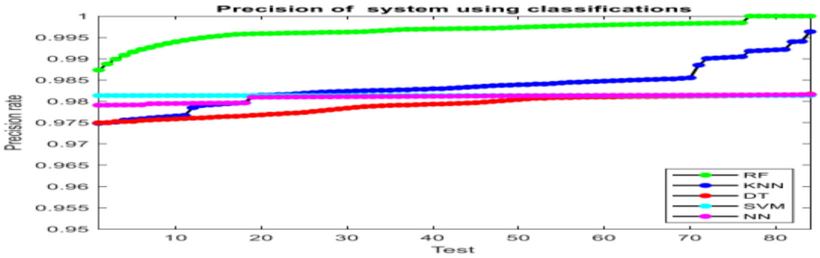

Figure 7- Comparison of precision of classification methods based on features with high standard deviation

According to this figure, the proposed method, using the combination of feature selection based on standard deviation and Random Forest, achieved better results in the accuracy metric compared to other methods. Considering the accuracy metric, the results indicate that the proposed method can be used as an accurate and reliable approach in diagnosing heart disease. However, like other metrics, further evaluation and more detailed analysis of results under different conditions and with a larger number of samples are necessary.

The F1 score is a composite metric that is the harmonic mean of precision and sensitivity. This metric shows how well the model can identify instances of heart disease while correctly classifying the ones that are identified as heart disease. In Figure 8, the graph related to the F1 score metric for different classification methods is shown.

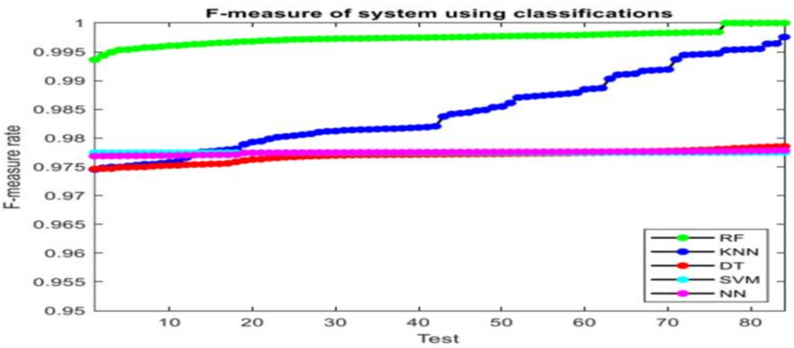

Figure 8 Comparison of the F criterion of classification methods based on features with high standard deviation

In Figure 8, the F1 score for the proposed method, using the combination of feature selection based on standard deviation and Random Forest, is displayed in comparison to other classification methods with different features. The results demonstrate that the proposed method, with the combination of feature selection based on standard deviation and Random Forest, achieves higher accuracy in diagnosing heart disease and obtains the best performance in the F1 score metric. Therefore, the results indicate that the proposed method, utilizing the combination of feature selection based on standard deviation and Random Forest, can be used as a precise and reliable approach for diagnosing heart disease and can be considered one of the best classification methods in this domain.

4.4. Comparison of the Proposed Method with Previous Approaches

In this stage, we compare the results of the proposed method with previous approaches. This comparison is performed using the test dataset and evaluating the performance of the methods with accuracy, sensitivity, specificity, and F1 score metrics. As the proposed method, we have utilized the combination of feature selection based on standard deviation and supervised classification methods for diagnosing heart disease. In comparison with the previous methods mentioned in [reference], the results of the proposed method are validated. By comparing the results of the proposed method with previous approaches, we can assess whether the proposed approach, using the combination of feature selection based on standard deviation and supervised classification methods, has provided a significant improvement in diagnosing heart disease or not. Emphasizing diversity in the test dataset for diagnosing heart disease is crucial, as patients with different characteristics and under various classifications need to be identified.
The metrics of accuracy, sensitivity, specificity, and F1 score are used to evaluate the performance of the classification methods and can serve as the best benchmarks for diagnosing heart disease. In Figure 9, bar charts display the evaluation metrics for the proposed method and previous approaches.

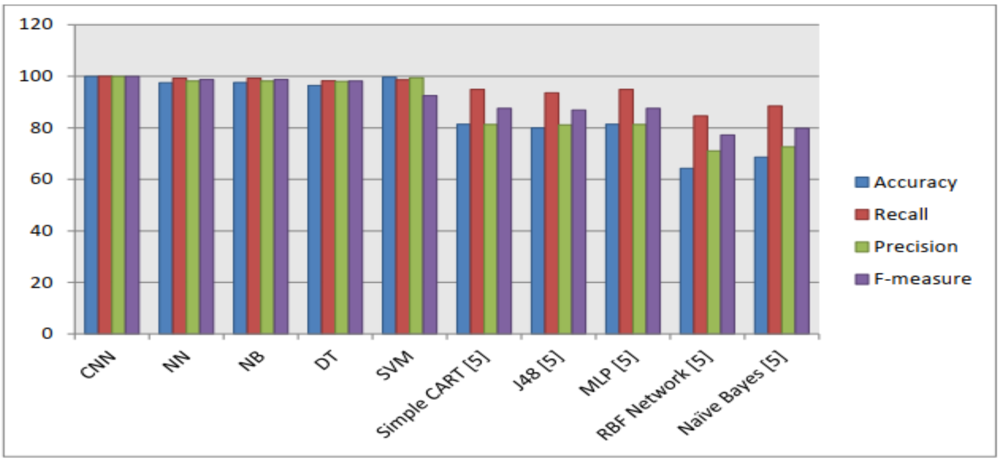

Figure 9- Comparison of the proposed method with previous methods

As shown in Figure 9, the results indicate that the proposed method, combining feature selection and classification approaches, achieved better results in diagnosing heart disease compared to previous methods. Moreover, the heterogeneous distribution of the test data is of great importance, as this diversity can guarantee an improved performance of the heart disease detection methods. Utilizing the random forest classification method based on feature selection using high standard deviation features can be considered as a suitable approach for classifying data and lead to more optimal results in data classification. Furthermore, the results of the proposed method have been compared with other classification methods, and the noticeable improvement in the proposed method compared to the previous approaches is evident.

5. Conclusion

In this paper, a disease detection system based on a combination of feature selection using high standard deviation features and supervised classification methods was proposed to diagnose heart disease in IoT-based healthcare systems. The proposed method utilized health records of patients from the PhysioNet database to create a recommender system for heart disease detection and prediction. The proposed method trained classification models on patient-specific heart disease symptoms using historical data. The implementation of the proposed method demonstrated the effectiveness of the random forest classifier based on high standard deviation features in accurately classifying existing patients and predicting new patients with heart disease. The proposed method achieved high performance scores in standard evaluation metrics compared to the random forest classifier and other implemented classification methods. The results obtained from the proposed method were compared with other existing disease detection methods, and the improvement achieved by the proposed method over previous approaches was evident. The proposed method's superior performance in comparison to previous methods in disease diagnosis is clearly observable and significant. Future research can focus on leveraging crowdsourced data collection to enhance the accuracy of healthcare recommendations, as well as integrating metaheuristic-based feature selection with deep learning to improve predictive performance for new patients.

References

[1]. Majumdar, Namrata, Shipra Shukla, and Anisha Bhatnagar. "Survey on applications of internet of things using machine learning." In 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 562-566. IEEE, 2019

[2]. Zolfagharipour L, Kadhim MH, Mandeel TH. Enhance the Security of Access to IoT-based Equipment in Fog. In2023 Al-Sadiq International Conference on Communication and Information Technology (AICCIT) 2023 Jul 4 (pp. 142-146). IEEE.

[3]. Bharadwaj, Hemantha Krishna, Aayush Agarwal, Vinay Chamola, Naga Rajiv Lakkaniga, Vikas Hassija, Mohsen Guizani, and Biplab Sikdar. "A review on the role of machine learning in enabling IoT based healthcare applications." IEEE Access 9 (2021): 38859- 38890.

[4].Al-Turjman, Fadi, ed. Artificial intelligence in IoT. Springer, 2019.

[5].Sivakumar, Nithya Rekha, and Faten Khalid Diaaldin Karim. "An IoT based big data 47 framework using equidistant heuristic and duplex deep neural network for diabetic disease prediction." Journal of Ambient Intelligence and Humanized Computing (2021): 1-11.

[6]. Amalia, R., A. Bustamam, and D. Sarwinda. "Detection and description generation of diabetic retinopathy using convolutional neural network and long short-term memory." In journal of physics: conference series, vol. 1722, no. 1, p. 012010. IOP Publishing, 2021.

[7]. Alfian, Ganjar, Muhammad Syafrudin, Norma Latif Fitriyani, Muhammad Anshari, Pavel Stasa, Jiri Svub, and Jongtae Rhee. "Deep neural network for predicting diabetic retinopathy from risk factors." Mathematics 8, no. 9 (2020): 1620.

[8]. Zhou, Huaping, Raushan Myrzashova, and Rui Zheng. "Diabetes prediction model based on an enhanced deep neural network." EURASIP Journal on Wireless Communications and Networking 2020, no. 1 (2020): 1-13.

[9]. Raja, M. J. B., & Pandian, S. C. “PSO-FCM based data mining model to predict diabetic disease”. Computer Methods and Programs in Biomedicine, 105659.47.(2020)

[10]. N Bagrecha, J., Chaithra, G S., and Jeevitha S. "Diabetes Disease Prediction using Neural Network." International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 7 Issue IV, Apr 2019.

Print ISSN 2710-0952 Electronic ISSN2790-1254

[11]. Mortajez, s., and amshidinezhad, A. "An Artificial Neural Network Model to Diagnosis of Type II Diabetes." Journal of Research in Medical and Dental Science, 2019, Volume 7, Issue 1, Page No: 66-70

[12]. M. S. Hossain and G. Muhammad, ‘‘Deep learning based pathology detection for smart connected healthcare,’’ IEEE Netw., vol. 34, no. ,6 pp. 120–125, Nov. .2020

[13]. M. Z. Uddin, ‘‘A wearable sensor-based activity prediction system to facilitate edge computing in smart healthcare system,’’ J. Parallel Distrib. Comput., vol. 123, pp. 46–53, Jan. .2019

[14]. S. S. Ram, B. Apduhan, and N. Shiratori, ‘‘A machine learning frame- work for edge computing to improve prediction accuracy in mobile health monitoring,’’ in Proc. Int. Conf. Comput. Sci. Appl. Cham, Switzerland: Springer, Jul. 2019, pp. 417–431.

[15]. P. Queralta, T. N. Gia, H. Tenhunen, and T. Westerlund, ‘‘Edge-AI in LoRa-based health monitoring: Fall detection system with fog computing and LSTM recurrent neural networks,’’ in Proc. 42nd Int. Conf. Telecommun. Signal Process. (TSP), Jul. 2019, pp. 601– 604.

[16]. G. Muhammad, M. S. Hossain, and N. Kumar, ‘‘EEG-based pathology detection for home health monitoring,’’ IEEE J. Sel. Areas Commun., vol. 39, no. 2, pp. 603–610, Feb. 2021, doi: 10.1109/JSAC.2020.3020654.

[17] Halder RK, Uddin MN, Uddin MA, Aryal S, Khraisat A. Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. Journal of Big Data. 2024 Aug 11;11(1):113.

[18] Mohammadi-Pirouz Z, Hajian-Tilaki K, Sadeghi Haddat-Zavareh M, Amoozadeh A, Bahrami S. Development of decision tree classification algorithms in predicting mortality of COVID-19 patients. International Journal of Emergency Medicine. 2024 Sep 27;17(1):126.

[19] Mutale B, Withanage NC, Mishra PK, Shen J, Abdelrahman K, Fnais MS. A performance evaluation of random forest, artificial neural network, and support vector machine learning algorithms to predict spatio-temporal land use-land cover dynamics: a case from Lusaka and Colombo. Frontiers in Environmental Science. 2024 Sep 12;12:1431645.

[20] Schmidgall S, Ziaei R, Achterberg J, Kirsch L, Hajiseyedrazi S, Eshraghian J. Brain-inspired learning in artificial neural networks: a review. APL Machine Learning. 2024 Jun 1;2(2).

[21] Wang J, Ouyang H, Jiao R, Cheng S, Zhang H, Shang Z, Jia Y, Yan W, Wu L, Liu W. The application of machine learning techniques in posttraumatic stress disorder: a systematic review and meta-analysis. NPJ Digital Medicine. 2024 May 9;7(1):121.