IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Block Chain-Based Secure Healthcare Application For Diabetic-Cardio Disease Prediction In Fog Computing

[1]P.J Mercy MCA, M.Phil., [2]N.Antony Aswathi

ABSTRACT:

Fog computing is a modern computing model which offers geographically dispersed end-users with latency-aware and highly scalable services. It is comparatively safer than cloud computing, due to information being rapidly stored and evaluated closer to data sources on local fog nodes. The advent of Blockchain (BC) technology has become a remarkable, most revolutionary, and growing development in recent years. BT's open platform stresses data protection and anonymity. It also guarantees data is protected and valid through the consensus process. BC is mainly used in money-related exchanges; now it will be used in many domains, including healthcare; This paper proposes efficient Blockchain-based secure healthcare services for disease prediction in fog computing. Diabetes and cardio diseases are considered for prediction. Initially, the patient health information is collected from Fog Nodes and stored on a Blockchain. The novel rule-based clustering algorithm was initially applied to cluster patient health records. Finally, diabetic and cardio diseases are predicted using feature selection based on an adaptive neuro-fuzzy inference system (FS-ANFIS). To evaluate the performance of the proposed work, an extensive experiment and analysis were conducted on data from real world healthcare. Purity and NMI metrics are used to analyze the performance of the rule-based clustering and accuracy is used for prediction performance. The experimental results show that the proposed work efficiently predicts the disease. The proposed work more than 81% prediction accuracy compared to the other neural network algorithms.

INTRODUCTION

Long-term technological developments offer considerable opportunities for biomedical innovation and cost reductions, but they also present a barrier to the integration of developing technologies into medical therapy. A significant amount of work is focused on smart healthcare to address traditional healthcare limits and meet rising premium healthcare expectations. With conventional healthcare, biosensors, smart healthcare could be designed and developed as a spectrum of devices, tools, software, facilities, and organizations.

A. MOTIVATION

Fog computing is one of the key technologies that contributes significantly to the promotion of IoT health-care and surveillance applications because these systems are latency-sensitive and real-time tracking, data processing, and decision-making are critical criteria in healthcare applications such as home nursing, heart care, diabetes, and other diseases. Because it contains crucial and personal information, health data is a hot topic. The goal of fog computing is to enable patients to manage their own personal health data on a local level. Fog nodes, such as smart phones or smart automobiles, save the safety information. Fog computing has a lot of benefits for fog-based applications that are slow. Mobile Fog was developed by Hongetal.

B. STATEMENT OF THE PROBLEM

Large volumes of information in a variety of formats, such as records, economic papers, clinical test outcomes, imaging tests, and vital sign evaluations, are all produced by health care contributors. The comprehensive database developed in care environments is quickly developing, with healthcare information grappling with a variety of issues, including data access and how information can be received outside of the healthcare ability. The ability to improve the data's authenticity and legitimacy is provided by blockchain. It also aids in the dissemination of data inside the network or services. Apps like these have an impact on the cost, data quality, and importance of providing health care inside the system. Blockchain is a decentralised, transparent network that eliminates the need for an intermediary. There is no need for multiple verifications in blockchain healthcare networks.

One of the most pressing real-world issues in the healthcare field is disease prediction. Many classification algorithms are used to accurately forecast diseases. One of the categorization algorithms is the artificial neural network (ANN). Because of its vast parallel structure, ANN is a massively computational parallel model with self-adaptive and self-learning capabilities that takes longer to forecast the outcome. ANN is not ideal for dealing with challenges like ambiguous and imprecise data, which might cause problems of uncertainty at any stage throughout the classification process.

One of the most pressing real-world issues in the healthcare industry is disease prediction. To accurately forecast diseases, many categorization methods are used. One type of categorization algorithm is the artificial neural network (ANN). Because of its huge parallel structure, ANN is a massively computational parallel model with self-adaptive and self-learning capabilities, but it takes longer to forecast the outcome. ANN is not suitable for dealing with challenges such as ambiguous and imprecise data, where problems of uncertainty can arise at any stage of the classification process. An Inference System (ANFIS) is a hybrid model that combines ANN and fuzzy logic.

C. CONTRIBUTIONS

The goal of this paper is to use feature selection and ANFIS to create an illness prediction model. One of the pre-processing techniques that minimises the amount of the dataset's dimensionality is feature selection. Cronbach's alpha is used in this paper to identify the best features.

The following are some of the paper's key findings:

- To ensure safe and effective data storage and sharing, a semi-centralized Blockchain-based digital healthcare network for the protection and sharing of patient data is introduced.
- The diabetes and cardiac disease patient records are grouped using a rule-based clustering technique.
- Using Feature selection based ANFIS, diabetes and cardiac disease are predicted after this clustering.
- Finally, the model is developed to assess the proposed work's performance.

II.BACKGROUND

A.FOG COMPUTING

It's a distributed computing platform that brings the cloud infrastructure of the network to the network's edge. It helps with data centre and end-user processing, networking, and storage operations and configurations. In both cloud and edge applications, fog computing refers to software specifications that operate between sensors and the cloud, such as smart access points, routers, or advanced fog devices. Fog computing combines agility, processing power, networking protocols, interface flexibility, cloud convergence, and distributed data analytics to satisfy the needs of applications that demand low latency and a vast geographical footprint.

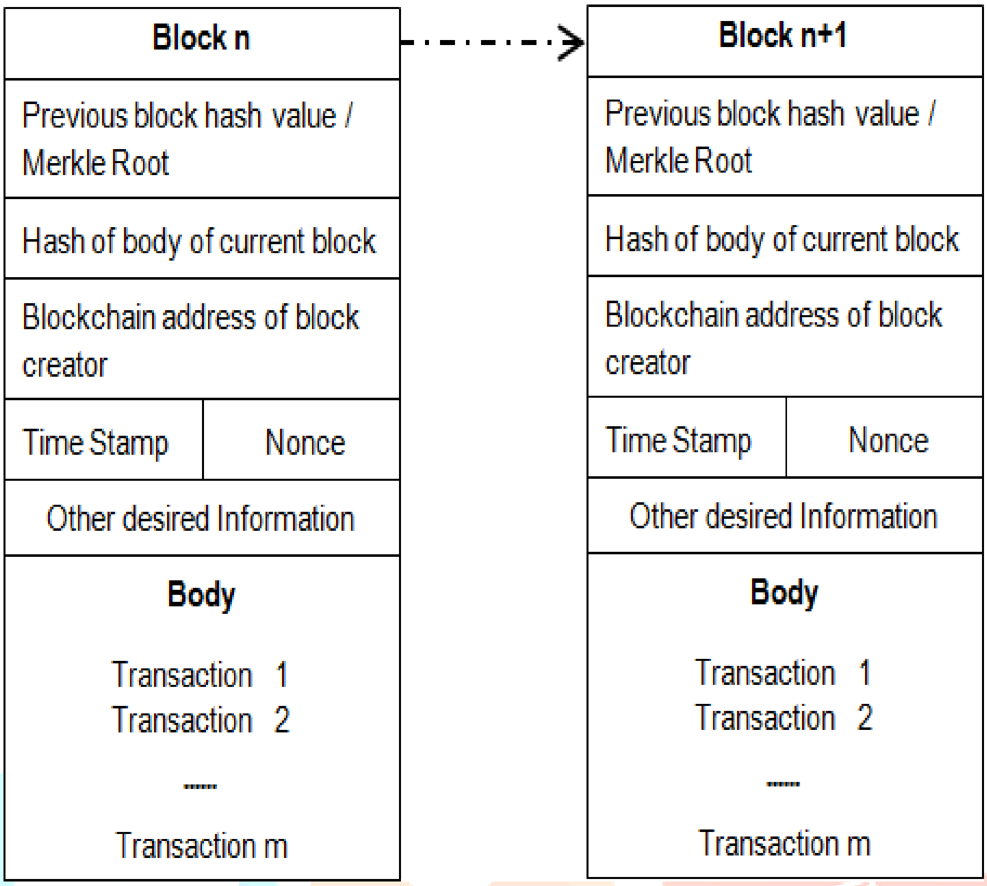
The term "fog computing" was invented by Cisco. Fog computing, according to the Open Fog Consortium, is "a horizontal system-level architecture that distributes processing, storage, and communication."

B. BLOCKCHAIN

Blockchain is a cutting-edge technology and a digital wallet that keeps account of all network transactions and events, and whose integrity is assured by a peer-to-peer computing network rather than a centralised body, removing the risk of a single central point. It's made up of structured documents organised in a block structure, with transaction batches and previous key hashes included. The data on the Blockchain network is unchallengeable because each block is chronologically linked.

In a blockchain network, any user has individual access rights to allow transactions that are amended throughout the framework, which is known as the consensus protocol. A blockchain employs the SHA256 hash for transaction insertion. The National Security Agency (NSA) develops that, which is 64 characters long. All transactions are recorded in a block chain network without altering or manipulating the public ledger; both transfers are distributed to various users across the network to transfer and update data; a block chain network may be duplicated to a different venue, such as within the same ability or healthcare distribution network, or as part of a regional or global data exchange system.

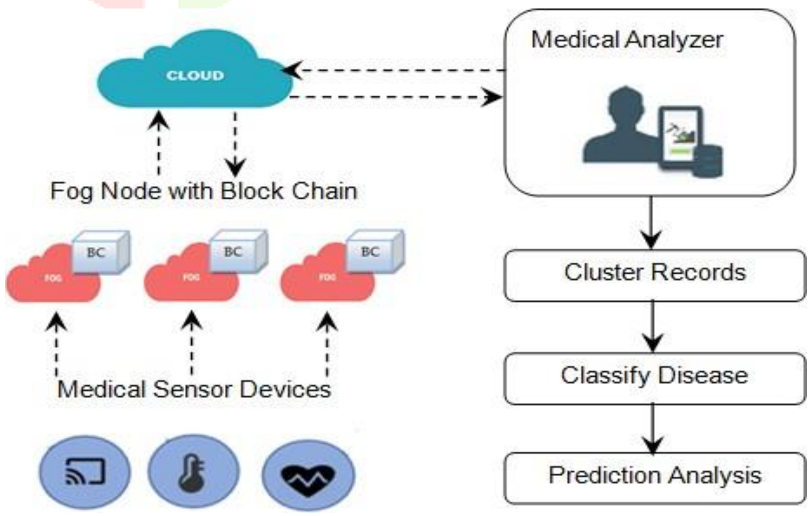
The data structure of the Block chain is a hierarchical set of blocks, as depicted in Figure 3. Blocks are connected in the form of a tuple , with the current block's header storing values such the previous block's hash, Block chain address, and so on. Every block is made up of two parts: a header and a body. Block number, previous block hash value to maintain chain stability, current block body hash to ensure transaction data integrity, timestamp, nonce, blockchain block creator address, and other requested detail are all included in the header. One or more transactions can be found in the block bodies.



MODEL OF THE SYSTEM

The proposed system model and nota- tions employed in this model are explained in this section. The IoT medical sensors are employed in this concept to collect patient health data.

These data are collected by fog nodes and sent to a medical analyzer for analysis and prediction of disease. The system model is depicted in Figure 4. It consists of five different entities.



A. SENSOR DEVICES FOR MEDICAL APPLICATIONS

Sensor devices, whether wearable systems or embedded sensors, can track a variety of human health characteristics. Because of their limited computational and storage capacities, these devices capture a variety of health-related data and transfer it to fog nodes that will be well-managed.

B. NODE OF FOG

It's a basic fog computing platform that can be a network computer that maintains underlying devices with processing resources, dedicated servers, or computational servers. It captures data from medical sensor equipment and saves it in a Blockchain-based distributed ledger.

C. BLOCKCHAIN

It's a cooperative network that keeps track of patient health and activity data. Without authorisation, no one can access the network. This is made up of a chain of blocks, each of which contains the preceding hash block, status user health.

D. CLOUD

It's utilised for storing things. It holds encrypted patient health information, which can be accessed by an authenticated medical analyst for further processing.

E.MEDICAL ANALYZER

A individual who has been granted access to patient health information. The information can be divided into two categories by the analyzer: normal patient and impacted patient. The analyzer can also tell if the patienThas diabetes or cardiovascular disease.

The notations used in the clustering and classification processes are listed in Table 1.

Notation	Description
D	Dataset
Fi	ith feature of D
DRi	ith data record in D
Aij	Attribute value of ith data record and jthfeatures.
RS	Rule set
Fi	Ithrule in RS
Freq<R,C>	Frequent rule set(R=RULE,C=COUNT)
Rthr	Rule threshold
L+R=C	Left, right and class part(rule)
Cand+	Positive candidate rule
Cand-	Negative candidate rule
CLS+	Positiveclustered data
CLS-	Negative clustered data
Cα	Cronbachs alpha

LITERATURE SURVEY

In this paper, we propose an electroencephalogram (EEG)-based seizure detection system in the IoT framework which uses the discrete wavelet transform (DWT), Hjorth parameters (HPs), statistical features, and a machine learning classifier. Seizure detection is done in two stages. In the first stage, EEG signals are decomposed by the DWT into sub-bands and features (activity, signal complexity, and standard deviation) were extracted from each of these sub-bands. In the second stage, a deep neural network (DNN) classifier is used to classify the EEG data. A prototype of the proposed neuro-detect was implemented using the hardware-in-the-loop approach. The results demonstrate a hiiisignificant difference in HP values between interictal and ictal EEG with ictal EEG being less complex than interictal EEG. In this approach, we report an accuracy of 100% for a classification of normal versus ictal EEG and 98.6% for normal and interictal versus ictal EEG.View less.

Electronic Health Records (EHRs) allows patients to control, share, and manage their health records among family members, friends, and healthcare service providers using an open channel, i.e., Internet. Thus, privacy, confidentiality, and data consistency are major challenges in such an environment. Although, cloud-based EHRs addresses the aforementioned discussions, but these are prone to various malicious attacks, trust management, and non-repudiation among servers. Hence, blockchain-based EHR systems are most popular to create the trust, security, and privacy among healthcare users. Motivated from the aforementioned discussions, we proposes a framework called as Blockchain-Based Deep Learning as-a-Service (BinDaaS). It integrates blockchain and deep-learning techniques for sharing the EHR records among multiple healthcare users and operates in two phases. In the first phase, an authentication and signature scheme is proposed based on lattices-based cryptography to resist collusion attacks among N-1 healthcare authorities from N. In the second phase, Deep Learning as-a-Service (DaaS) is used on stored EHR datasets to predict future diseases based on current indicators and features of patient. The obtained results are compared using various parameters such as accuracy, end-to-end latency, mining time, and computation and communication costs in comparison to the existing state-of-the-art proposals. From the results obtained, it is inferred that BinDaaS outperforms the other existing proposals with respect to the aforementioned parameters.View less.

The advancement of information technology in coming years will bring significant changes to the way healthcare data is processed. Technologies such as cloud computing, fog computing, and the Internet of things (IoT) will offer healthcare providers and consumers opportunities to obtain effective and efficient services via real-time data exchange. However, as with any computer system, these services are not without risks. There is the possibility that systems might be infiltrated by malicious users and, as a result, data could be corrupted, which is a cause for concern. Once an attacker damages a set of data items, the damage can spread through the database. When valid transactions read corrupted data, they can update other data items based on the value read. Given the sensitive nature of healthcare data and the critical need to provide real-time access for decision-making, it is vital that any damage done by a malicious transaction and spread by valid transactions must be corrected immediately and accurately. Here, we present two models for using fog computing in healthcare: an architecture using fog modules with heterogeneous data, and another using fog modules with homogeneous data. We propose a unique approach for each module to assess the damage caused by malicious transactions, so that original data may be recovered and affected transactions may be identified for future investigations.

Edge computing paradigm has attracted many interests in the last few years as a valid alternative to the standard cloud-based approaches to reduce the interaction timing and the huge amount of data coming from Internet of Things (IoT) devices toward the Internet. In the next future, Edge-based approaches will be essential to support time-dependent applications in the Industry 4.0 context; thus, the paper proposes BodyEdge, a novel architecture well suited for human-centric applications, in the context of the emerging healthcare industry. It consists of a tiny mobile client module and a performing edge gateway supporting multiradio and multitechnology communication to collect and locally process data coming from different scenarios; moreover, it also exploits the facilities made available from both private and public cloud platforms to guarantee a high flexibility, robustness, and adaptive service level. The advantages of the designed software platform have been evaluated in terms of reduced transmitted data and processing time through a real implementation on different hardware platforms. The conducted study also highlighted the network conditions (data load and processing delay) in which BodyEdge is a valid and inexpensive solution for healthcare application scenarios.

In this paper, we propose GuardHealth: an efficient, secure and decentralized Blockchain system for data privacy preserving and sharing. GuardHealth manages confidentiality, authentication, data preserving and data sharing when handling sensitive information. We exploit consortium Blockchain and smart contract to achieve secure data storage and sharing, which prevents data sharing without permission. A trust model is utilized for precisely managing trust of users with the implementation of the state-of-art Graph Neural Network (GNN) for malicious node detection. Security analysis and experiment results show that the proposed scheme is applicable for smart healthcare system. linksopenoverlaypanelAntonHasselgren[a]KatinaKralevska[b]DaniloGligoroski[b]Sindre,A.Pedersen[c]ArildFaxvaag[a],Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, NTNU-Norwegian University of Science and Technology, Trondheim, Norway,Department of Information Security and Communication Technology, Faculty of Information Technology and Electrical Engineering, NTNU-Norwegian University of Science and Technology, Trondheim, Norway,Library Section for Medicine and Health Sciences,

Blockchain can be described as an immutable ledger, logging data entries in a decentralized manner. This new technology has been suggested to disrupt a wide range of data-driven domains, including the health domain.

In this paper, a Linguistic Neuro-Fuzzy with Feature Extraction (LNF-FE) model is utilized for the analysis of medical data for disease classification. Initially, this model uses a linguistic *fuzzification* process to generate membership values that handle the uncertainty problems. These membership values may not significantly contribute to the model, but it will increase the dimensions, for which more time will be required to train the model. To address this issue, Feature Extraction (FE) algorithms are hybridized in the Neuro-Fuzzy (NF) model to extract only those features (a reduced feature set) that are significantly contributing to the network. These reduced features are again passed to the *Artificial Neural Network* (ANN) model for classification. This proposed model is tested and validated through eight benchmark datasets, and the performance is compared with other models. The obtained results were tested using statistical techniques such as Friedman and Holm-Bonferroni for the proof of correctness. This experimental analysis shows that our proposed model outperforms better as compared to other models for solving real-world problems.

Cloud computing provides resources over the Internet and allows a plethora of applications to be deployed to provide services for different industries. The major bottleneck being faced currently in these cloud frameworks is their limited scalability and hence inability to cater to the requirements of centralized Internet of Things (IoT) based compute environments. The main reason for this is that latency-sensitive applications like health monitoring and surveillance systems now require computation over large amounts of data (Big Data) transferred to centralized database and from database to cloud data centers which leads to drop in performance of such systems. The new paradigms of fog and edge computing provide innovative solutions by bringing resources closer to the user and provide low latency and energy efficient solutions for data processing compared to cloud domains. Still, the current fog models have many limitations and focus from a limited perspective on either accuracy of results or reduced response time but not both. We proposed a novel framework called HealthFog for integrating ensemble deep learning in Edge computing devices and deployed it for a real-life application of automatic Heart Disease analysis. HealthFog delivers healthcare as a fog service using IoT devices and efficiently manages the data of heart patients, which comes as user requests. Fog-enabled cloud framework, FogBus is used to deploy and test the performance of the proposed model in terms of power consumption, network bandwidth, latency, jitter, accuracy and execution time. HealthFog is configurable to various operation modes which provide the best Quality of Service or prediction accuracy, as required, in diverse fog computation scenarios and for different user requirements.

I Machine learning involves artificial intelligence, and it is used in solving many problems in data science. One common application of machine learning is the prediction of an outcome based upon existing data. The machine learns patterns from the existing dataset, and then applies them to an unknown dataset in order to predict the outcome. Classification is a powerful machine learning technique that is commonly used for prediction. Some classification algorithms predict with satisfactory accuracy, whereas others exhibit a limited accuracy. This paper investigates a method termed ensemble classification, which is used for improving the accuracy of weak algorithms by combining multiple classifiers. Experiments with this tool were performed using a heart disease dataset. A comparative analytical approach was done to determine how the ensemble technique can be applied for improving prediction accuracy in heart disease. The focus of this paper is not only on increasing the accuracy of weak classification algorithms, but also on the implementation of the algorithm with a medical dataset, to show its utility to predict disease at an early stage. The results of the study indicate that ensemble techniques, such as bagging and boosting, are effective in improving the prediction accuracy of weak classifiers, and exhibit satisfactory performance in identifying risk of heart disease. A maximum increase of 7% accuracy for weak classifiers was achieved with the help of ensemble classification. The performance of the process was further enhanced with a feature selection implementation, and the results showed significant improvement in prediction accuracy.

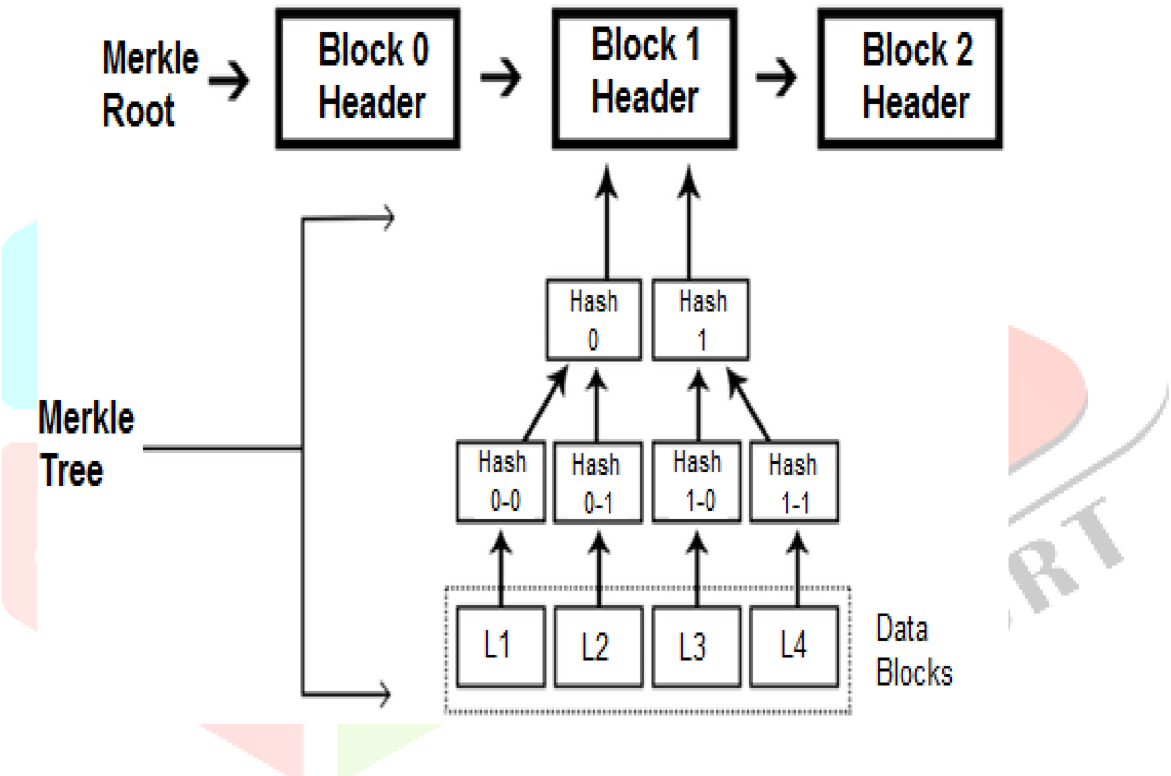
Disease gene detection is an important stage in the understanding disease processes and treatment. Some candidate disease genes are identified using many machine learning methods Although there are some differences in these methods including feature vector of genes, the method used to selecting reliable negative data (non-disease genes), and the classification method, the lack of negative data is the most significant challenge of them. Recently, candidate disease genes are identified by semi-supervised learning methods based on positive and unlabeled data. These methods are reasonably accurate and achieved more desirable results versus preceding methods. In this article, we propose a novel Positive Unlabeled (PU) learning technique based upon clustering and One-Class classification algorithm. In this regard, unlike existing methods, we make a more Reliable Negative (RN) set in three steps: (1) Clustering positive data, (2) Learning One-Class classifier models using the clusters, and (3) Selecting intersection set of negative data as the Reliable Negative set. Next, we attempt to identify and rank the candidate disease genes using a binary classifier based on support vector machine (SVM) algorithm. Experimental results indicate that the proposed method yields to the best results, that is 92.8, 93.6, and 93.1 in terms of precision, recall, and F-measure respectively. Compared to the existing methods, the increase of performances of our proposed method is 11.7 percent better than the best method in terms of F-measure. Also, results show about 6% increase in the prioritization results.

V. PROPOSED METHODOLOGY

The suggested Blockchain-based healthcare disease prediction with clustering and classification is described in this section.

STORAGE ON THE BLOCKCHAIN

Control of access, authenticity, data confidentiality, and integration are all important in the medical sector when it comes to protecting a patient's identification and sharing data with other organisations within the healthcare environment. The traditional method of gaining access control frequently necessitates trust between the data owner and the entities that store it. These organisations are also given sole responsibility for identifying and implementing access management policies. Interoperability refers to the ability of disparate information systems, software, or frameworks to synchronise data across stakeholders within and across organisational boundaries in order to promote individual safety. Data provenance refers to the data's historical record.



The consortium kind of Blockchain, often known as semi-decentralized Blockchain, is the subject of this study. A consor- tium blockchain is not given to a single entity as a private blockchain; instead, it is given to a group of approved entities. A blockchain consortium is also a collection of pre-defined nodes on the network. As a result, consortium blockchain provides security that is inherited from public blockchain. This provides the network a large degree. Consortium blockchains are most usually associated with commercial application, in which a group of companies collaborates to use blockchain technology to promote their businesses. This type of Blockchain, on the other hand, may allow select members of a group to access or use a hybrid form of access. It's possible that the root hash and its Application Program Interface (API) are the same thing.

The authorised medical analyzer gathers information about the patient and determines whether the patient has diabetes or heart disease.

CLUSTERING OF DISEASE

Clustering is an unsupervised data mining approach for discovering groupings in unlabeled data. It's used to divide a set of data into different clusters, with things in one cluster strongly connected to objects in another. Pattern identification, image processing, and consumer transaction pattern analysis all use clustering technology. It's crucial during data analysis discovery and evaluation, when researchers are looking for fundamental aspects in the data that don't surface when they don't know anything about it. However, when it comes to choosing the right people, it's important to keep in mind that there are a lot of factors to consider

Cluster classification is a structured, codified approach for data discovery and identifying clinically related groupings in the medical sector. Costly health-care services are becoming more competitive as a result of efficient clustering technologies. It aids doctors in dealing with the influx of information and can help with strategic planning by providing better facilities. Clustering results are used to investigate patient independence or association, as well as to gain a deeper understanding of evidence from medical surveys. All of these benefits prompted the researcher to develop clustering models for medical data.Clustering health data creates a slew of new issues.

Overabundance of data — Advances in medical technology, paired with high processing capacities, are increasing the amount of data created and processed in the healthcare industry. Knowledge discovery and retrieval from these massive databases is challenging and excessively expensive.There are far too many risk indicators that are both necessary and heterogeneous for decision-making.

Consumer awareness of medical treatments is increasing, and life expectancy is increasing, resulting in a growing need for better health services. Overworked and inexperienced doctors, demanding working environments, and other factors contribute to misdiagnosis and imprecise care techniques.

Choosing an appropriate clustering approach and a sufficient number of clusters in health care data can be difficult and time-consuming.

For the efficient cluster, an unique rule-based clustering technique is proposed to address this difficulty. This is a two-step algorithm: the rules are formed based on patient information in the first stage, and the clusters are generated based on the rules in the second stage.

The following is the pseudo-code for the rule generating algorithm.

This approach works well with numerical data. Numerical values are first converted to discrete values (Low, Medium, and High) . Candidate rules are developed based on these data for further processing.The creation of rules in this study is based on frequency and thresholds.

The candidate rules are retrieved based on the requirements.Take the fasting blood sugar levels of 15 patients: 120, 90, 70, 45, 100, 130, 50, 35, 138, 82, 90, 50, 120, 58, 140.The example is shown in Table 3.

Convert each of the dataset's feature values. Count how many times each record has been played. Consider the record as a candidate rule if the record frequency is more than Rthr (starting with 5–10 depending on the conditions). Clustering is the following phase. The clustering algorithm's pseudo-code is presented below.

Left, right, and a class variable make up the candidate rules (L R C). Cand and cand rules are generated using the C (class variable). Based on the candidate rules, positive and negative clusters are produced. Any record that does not match the candidate rules will be classified as an outlier.

ALGORITHM1: RULE GENERATION

Input: D

Output: RS

```
1:   RS=∅
2:   foreach Fi ∈Featuredo
3:   distFi=get distinct value(Fi)
4:   Sort(distFi)
5:   GroupdistFivaluesintoLow,MediumandHigh
6:   endfor
7:   foreachDRi ∈DataRecord do
8:   foreach Fj ∈Featuredo
9:   newAij=convertAij intoLow,Medium,HighbasedonStep5
10:  endfor
11:  endfor
12:  generatenewDRbasedonnewAij
13:  Freq<R,C>=FindandCountSimilarRecords
14:  candidate Freq<R,C>c>Rthr
15:  Ifcandidate/=∅
16:  RS=candidate
17:  else
18:  RS=∅
19:  endif
```

TABLE 3: Data conversion example.

STEPS	VALUE
Input Feature	120,90,70,45,100,130,50,35,138,82,90 ,50,50,58,140
Distinct value	120,90,70,45,100,130,50,35,138,82,58 ,140
Sort value	35,45,50,58,70,82,90,100,120,130,138 ,140
Group values	(35,45,50,58)=low (70,82,90,100)=medium (120,130,138,149)=high
convert	High,medium,medium ,low,medium,high,
Feature	Low,low,high,medium,medium,low, High,low,hgh

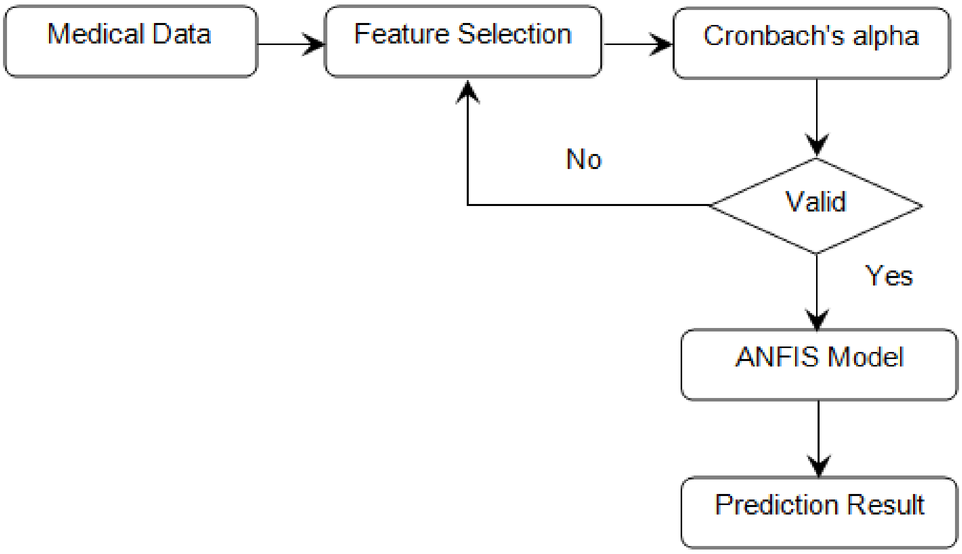
PREDICTION OF DISEASE

Medical data processing is a crucial area that requires accuracy for disease prevention, diagnosis, and processing. A crucial scientific aim has been to keep track of medical records. To ensure professional therapy, patient data containing particular disease-related characteristics and symptoms will be obtained with extreme caution. Medical data is inefficient because it can contain incomplete and redundant information maintained in medical repositories. It is critical to contain effective data planning and reduction until data mining algorithms are implemented because this can affect mining performance. If the data is accurate, dependable, and devoid of noise, disease diagnosis becomes faster and easier.In data mining, selecting a feature is an effective pre-processing strategy for reducing data dimensionality.

ALGORITHMS2: CLUSTERING

```
Input: D, RS
Output:Cls+,Cls−
1: cand+=∅ , cand−=∅
2: foreach Ri ∈ RS do
3:    Split Ri into three parts (L   R   C)
4:   cand+ = L   R   C (rule with positive patients)
5:   cand− = L   R   C (rule with negative patients)
6: end for
7: foreach rec € newDR
8:   if (rec match with cand+) then
9:      Cls+.add(rec)
10:  else (rec match with cand-) then
11.Cls-.add(rec)
12.else
13.out.add(rec)
14.end if
15.end for
```

In medical diagnostics, identifying the most severe disease-related risk factors is critical. Specific feature recognition aids in the removal of undesirable, unneeded features from the disease dataset, resulting in a more straightforward and superior output. Classification and prediction is a data mining process in which training data is used to develop a training model, which is then applied to test data to get predictive results. For the management of diabetes and cardiovascular disease, many recognition methods have been applied to disease data sets. This research proposes a Feature Selection and Adaptive Neuro-Fuzzy Inference System for disease prediction that uses the characteristics of ANN and Fuzzy Logic. The procedure of the prediction model in workflow.



WORKFLOW PREDICTION

Feature selection is a data pre-processing strategy that is often used in data mining to decrease data by removing irrelevant and redundant features from the dataset. This strategy also improves data interpretation, information analysis, learning algorithm training times, and prediction efficiency.Various feature collection methods have been applied to healthcare datasets in order to capture more usable knowledge. To predict various diseases, feature selection approaches are applied to clinical datasets. When there are more significant and non-redundant attributes in the details, different learning algorithms work better and produce more trustworthy results. Given the large amount of redundant and useless features in medical datasets, a good feature extraction approach is needed to extract interesting disease-specific properties.

Cronbach's alpha is used in this research to suggest an optimal feature selection strategy. The Cronbach alpha measures a test's internal consistency, or the consistency of its features. It can be assessed using the following criteria:

$$C\alpha = \frac{|F| \cdot CV_{avg}}{V_{avg} + (|F| - 1) \cdot CV_{avg}}$$

Where F is the number of features, CVavg is the average covariance, and Vavg is the average variance.

The following is the pseudo-code for the feature selection algorithm.

ALGORITHM3: FEATURE SELECTION

Input:D

Output:SF(SelectedFeatures)

1:pc 10,global_Cα=0,maxIter 100
2:fori =1topcdo

3: $pop^j = Random\ \{0,1\}, j \in F_j$

4:Cα_i=0

```
endfor
foriter=1tomaxIterdo
    fori=1topc
    computeCronbach'salpha(Ca)using(1)
if(Ca>Cαi)then
    Cαi=Ca
    endif
endfor
  maxCa=max(Cαi)
if(maxCa>global_Cα)then
            global_Cα        maxCa
      SF        popi(index of maxCa)
endif
ReplacethepopwhichcontainlowestCa
endfor
```

Using the random function, generate the population at random and set alpha to zero. To choose the best features, an iterative method is employed. The maximum number of iterations is 100. Calculate Cronbach's alpha for each population chosen at random. Set the specified features as population if the maximum alpha value is greater than the global alpha value. Change the population to the one with the lowest alpha and repeat the steps until the maximum number of iterations is reached. The ANFIS model uses the selected features to predict the disease.

EXPERIMENTAL RESULT

The performance of the proposed work was evaluated in this area. The suggested work was written in Java (version 1.8), and the trials were run on a Windows 7 32-bit operating system on an Intel(R) Pentium computer with a speed of 2.13 GHz and 4.0 GB RAM.

A. DATA COLLECTION

The experimental results are based on two datasets: diabetes and heart disease. There are 768 cases in the diabetes data collection, each with eight numeric characteristics. The data set information is shown in Table 4.

Name	Description	Range
Pregencies	No of pregencies	0-17
Count	Pregencies	
Glucose Level	Plasma glucose level	44-199
BP level	Diastolic hypertension	24-122
Skin	The thickness of	7-99
Thickness	Triceps skinfolds	
Insulin	Insulin serum for 2hours	14-846
BMI	Body mass index	18.2-67.1
Pedigree	A pedigree functio	0.078-2.42
Function	Of diabetes	
Age	Age in years	21-81
Class Label	The patient has diabetes or not	0 or 1

CONCLUSION

The usage of Blockchain in the present healthcare system is critical. It can lead to automated processes for gathering and verifying data, correcting and aggregating information from various sources that are undeniable, resistant to manipulation, and offering protected data with reduced cybercrime risks and system redundancy. In fog computing, this study provides efficient Blockchain-based secure healthcare services for disease prediction. Diabetes and cardiovascular illnesses are taken into account while making predictions. In comparison to other methods, the proposed work efficiently clusters and predicts disease. Security and privacy for accessing patient medical data, as well as a hybrid clustering and classification algorithm, could be introduced in the future to improve the prediction results' performance

REFERENCE

1.)M. A. Sayeed, S. P. Mohanty, E. Kougianos and H. P. Zaveri, "Neuro-Detect: A Machine learning-Based Fast and Accurate Seizure Detection System in the IoMT," in IEEE Transactions on Consumer Electronics, vol. 65, no. 3, pp. 359-368, Aug. 2019, doi: 10.1109/TCE.2019.2917895.

2.) P. Bhattacharya, S. Tanwar, U. Bodkhe, S. Tyagi and N. Kumar, "BinDaaS: Blockchain-Based Deep-Learning as-a-Service in Healthcare 4.0 Applications," in *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1242-1255, 1 April-June 2021, doi: 10.1109/TNSE.2019.2961932.

3.) DOI:10.1007/978-3-030-24907-6_6,Corpus ID: 195877753,"Ensuring Data Integrity in Fog Computing Based Health-Care Systems",Abdulwahab Alazeb, B. Panda,Published in SpaCCS 14 July 2019,Computer SciencScience

4.) P. Pace, G. Aloi, R. Gravina, G. Caliciuri, G. Fortino and A. Liotta, "An Edge-Based Architecture to Support Efficient Applications for Healthcare Industry 4.0," in IEEE Transactions on Industrial Informatics, vol. 15, no. 1, pp. 481-489, Jan. 2019, doi: 10.1109/TII.2018.2843169.

5.) Ziyu Wang, Nanqing Luo, Pan Zhou,GuardHealth: Blockchain empowered secure data management and Graph Convolutional Network enabled anomaly detection in smart healthcare,Journal of Parallel and

Distributed Computing,Volume 142,2020,Pages 1-12,ISSN 0743-7315,Ziyu Wang, Nanqiages 1-12,ISSN 0743-7315, Hubei Engineering Research Center on Big Data Security, School of Cyber Science and Engineering, Huazhong University of Science & Technology, Wuhan, 430074, ChinaReceived 1 December 2019, Revised 10 March 2020, Accepted 14 March 2020, Available online 14 April 2020, Version of Record 20 April 2020.

6.) Blockchain in healthcare and health sciences—A scoping review,NTNU University Library, NTNU-Norwegian University of Science and Technology, Trondheim, Norway
Received 13 May 2019, Revised 25 November 2019, Accepted 2 December 2019, Available online 11 December 2019, Version of Record 19 December 2019.

7.) Medical disease analysis using Neuro-Fuzzy with Feature Extraction Model for classificatclassification,HimansuDasa, BighnarajNaikb,Department of Information Technology, Veer Surendra Sai University of Technology, Burla, Sambalpur, 768018, Odisha, India,Department of Computer Application, Veer Surendra Sai University of Technology, Burla, Sambalpur, 768018, Odisha, India

Received 22 September 2019, Revised 22 December 2019, Accepted 28 December 2019, Available online 30 December 2019, Version of Record 10 January 2020.

8.) Shreshth Tuli, Nipam Basumatary, Sukhpal Singh Gill, Mohsen Kahani, 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, Volume 104,2020,Pages 187-200,ISSN 0167-739X,

9.) Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques,Author links open overlay panelC. Beulah ChristalinLathaS. CarolinJeeva,Karunya Institute of Technology and Sciences, India, Received 26 November 2018, Revised 30 June 2019, Accepted 1 July 2019, Available online 2 July 2019, Version of Record 9 July 2019., Volume 16,2019,100203,ISSN 2352-9148

10.)Vasighizaker A, Jalili S. C-PUGP: A cluster-based positive unlabeled learning method for disease gene prediction and prioritization. Comput Biol Chem. 2018 Oct;76:23-31. doi: 10.1016/j.compbiolchem.2018.05.022. Epub 2018 Jun 1. PMID: 29890338.