iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43299

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Blockchain-based Secure Healthcare for Cardio Disease Prediction of Arrhythmia

Arbaaz Bebal[1], Nomit Bhatnagar[2], Ankita Jagtap[3], Pratiksha Kamthe[4], Gajanan Arsalwad[5]
Department of Information Technology, Trinity College of Engineering and Research, Pune.

Abstract: Heart disease is the leading cause of death worldwide. According to a recent study by the Indian Council of Medical Research (ICMR), roughly 25% of the deaths among people aged 252 to 69 are caused by various heart- related issues. The most common diseases are cardiovascular disorders. Due to a shortage of professionals and a significant number of incorrectly diagnosed cases, a quick and efficient detection method is required. So we should have always leaped on vigilance and care approaches and methods to avoid folks working extra due to the guts attack. Machine learning techniques are frequently used to make disease predictions. Blockchain technology has the ability to prevent data leaks and fraud. It has the potential to improve patient-hospital coordination . The suggested method improves data security while reducing the cost, time and resources needed to maintain a patient's data and outcomes.
Keywords: Blockchain, Healthcare, Machine Learning, Arrhythmia, Convolution Neural Network, Inter-Planetary File System.

I. INTRODUCTION

Lately in the upgoing years it has been recorded that heart diseases are the major cause for death across the world. Among the deaths that are caused by heart diseases, few of them are due to natural clinical reasons while the rest are due to the delayed diagnosis of heart diseases. ECG test are the basic fundamental test to be carried out for the diagnosis of any heart disease. Diagnosis of a heart disease at an early stage can reduce the threat of death to a patient at a major extent. The shortage of specialists and high wrongly diagnosed cases have necessitated the necessity to develop a quick and efficient detection system. By applying machine learning technique, the prediction of the disease is often done. Blockchain technology has the potential to avoid fraud and data leakage. It can make better coordination between patient and hospital. The proposed system increases data security and removes the cost, time, and resources required to manage the patient's data and results.

II. MOTIVATION

Cardiovascular disease is the leading global cause of death. It is very difficult for a doctor to read an ECG report with bare eyes. At times, there is high chance to miss out any abnormality in the ECG report as the change in the ECG wave shape is hardly noticeable. Here we are developing a scheme that can analyse the ECG data of patient for predicting the type of arrhythmia.

III. LITERATURE SURVEY

1) *2021)* This project uses fog computing to create effective Blockchain-based safe health-care services for disease prediction. When developing projections, diabetes and cardiovascular diseases are taken into account. Fog Nodes are used to collect patient health data, which is then stored on a Blockchain. The new rule-based clustering technique is used to cluster the patient health records initially. Finally, diabetes and cardiovascular disorders are predicted using a feature selection-based adaptive neuro-fuzzy inference method (FS-ANFIS). Security and privacy for accessing patient medical data, as well as numerous hybrid clustering and classification methods, can be employed to increase the performance of the prediction findings.
2) (2021) Machine learning is an important and effective tool for analysing extremely complex medical data. Massive amounts of medical data are being generated, and it is critical to use this data effectively to improve the medical and health-care sectors around the world. A complete literature analysis of many machine learning algorithms used for a variety of medical applications that have recently been published in highly reliable venues is included in this survey study. Only recent research can be utilised to survey existing machine learning and deep learning algorithms for medical data. This research shows a distinct shift in the use of artificial intelligence techniques in the medical field, with deep learning methods taking precedence over previous methods.The Isolation Forest approach is used to find outliers. Isolation Forest is a decision tree-based unsupervised machine learning approach. It discovers outliers by picking one of the characteristics at random and then splitting the feature's minimum and maximum values arbitrarily. This method of random feature subdivision, which creates smaller paths

in the tree structure, distinguishes the anomalies from the regular data. We used this technique because it detects irregularities quickly and consumes less memory than other outlier identification algorithms. We can increase the accuracy of heart disease risk prediction by using ensemble classification algorithms.

3) (2021) The study of computer algorithms for experience-based automation is known as machine learning (ML). Artificial intelligence (AI) is a subset of machine learning (ML) that allows computers to accomplish tasks that would normally need human intelligence. While effective healthcare communication is critical for effectively translating and disseminating information to support and educate patients and the general public, machine learning has proven to be useful in healthcare due to its ability to manage complex discussions and conversational flexibility.
4) (2021) To merge ensemble deep learning with Edge computing devices, Health Fog is being used in real-world applications such as autonomous heart disease analysis. It uses IoT devices to provide healthcare as a fog service and efficiently organises cardiac patient data that comes in per user. More intelligent ensemble models can be used to improve accuracy even more. Furthermore, the proposed architecture might be made robust and ubiquitous for a wide range of corporate fog computing applications, including agricultural, healthcare, weather forecasting, traffic management, and smart cities.
5) (2020) Medical data is analysed using a Linguistic Neuro-Fuzzy with Feature Extraction (LNF-FE) model for disease categorization. To begin, this method uses a linguistic fuzzification technique to obtain membership values that handle uncertainty issues. These membership values may not have a significant impact on the model, but they will expand the dimensions, necessitating extra training time. According to this experimental inquiry, our proposed approach beats existing approaches in tackling real-world situations. Medical disease classification using machine learning algorithms is difficult due to the nature of data, which can contain partial, confusing, and imprecise information. The existence of such information in the dataset affects the classification model's performance. Neuro fuzzy logic likewise relies totally on human understanding and expertise. The rules of a Fuzzy Logic control system must be updated often. These platforms do not recognise machine learning or neural networks.
6) (2020) Edge-based privacy-preserving cryptosystems are one of the forthcoming aspects of cloud-based secure remote healthcare monitoring systems. A cloud-based healthcare system will typically accept remote patient data via a sensor layer and provide continuous monitoring and diagnosis utilising various decision support system prediction methodologies. Traditional healthcare systems struggle to detect and handle real-time patient medical data while maintaining privacy and security. While edge computing improves security by reducing the amount of data that must be stored in data centres, it also introduces security concerns at each edge network point. Furthermore, particular data is more vulnerable to breaches since not every edge device has the same built-in authentication and security capabilities. Edge devices may require additional hardware and software to provide optimal performance and meet local storage demands.
7) (2019) A machine learning-based automatic seizure detection approach has been created in the IoT framework, which uses Hjorth parameters, statistical characteristics, and DWT-based feature extraction. The system was validated using a hardware-in-loop simulation technique. The results of the experiments show that the proposed method is particularly effective in interpreting complex EEG oscillations, resulting in greater classification accuracy than existing methods. The recommended Neuro-Detect was prototyped using a hardware-in-the-loop simulation technique. The simulator was set up using a hardware support package provided by the vendor. The actual board was constructed, and the Neuro-Detect programme was then run on it. On the 'Neuro-Detect' channel, the user's EEG data was continuously saved in the IoT cloud storage, which increased product expenses.
8) (2019) The authors of this paper discuss machine learning and data mining, as well as how they could be used to the health-care system. The healthcare industry is a data mining application field since it has large data resources that are difficult to manage manually. In developed countries, heart disease has been identified as one of the leading causes of death. One of the reasons why people die from heart disease is that the risks are either not recognised or are detected later.

IV. DIFFERENT APPROACHES

A. Support Vector Machine (SVM)

SVM stands for Support Vector Machine and is one of the most widely used Supervised Learning algorithms for Classification and Regression issues. However, it is mostly utilized in Machine Learning for Classification difficulties. The SVM algorithm's purpose is to find the optimum line or decision boundary for categorizing n-dimensional space into classes so that additional data points can be readily placed in the correct category in the future.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

B. K-Nearest Neighbors (KNN)

K-Nearest Neighbor is a Supervised Learning-based Machine Learning algorithm that is one of the most basic. The K-NN algorithm assumes that the new case/data and existing cases are similar and places the new case in the category that is most similar to the existing categories. The K-NN method maintains all of the available data and classifies a new data point based on its resemblance to the existing cheval.

C. Convolution Neural Network (CNN)

Convolution Neural Network Traditional feature learning methods rely on semantic labels of images as supervision. They usually assume that the tags are evenly exclusive and thus do not pointing out towards the complication of labels. The learned features endow explicit semantic relations with words. We also develop a novel cross-modal feature that can both represent visual and textual contents. CNN is a method of categorizing the images as a part of deep learning. In which we apply a single neural network to the full image. The steps in CNN are: convolution, subsampling, activation and full connectedness.

D. Random Forest

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression. One of the most important features of the Random Forest Algorithm is that it can handle the data set containing continuous variables as in the case of regression and categorical variables as in the case of classification.

E. Inter Planetary File System (IPFS)

IPFS is a decentralized file sharing platform that identifies files through their content. When a file is uploaded to IPFS, it is split into chunks, each containing at most 256 kilobytes of data and/or links to other chunks. Every chunk is identified by a cryptographic hash, also named content identifier that is computed from its content.

V. PROPOSED SYSTEM

A. System

Add the patient data like Name, Age, Gender, Phone number and upload the file on the system.

B. Central Server

The CNN algorithm is applied on the uploaded patient file. In this the given data is first distributed according to the features required, the obtained features are then trained and tested. After this process the tested data is modulated for the prediction purpose.

C. Analysis

The modulated data when sent to the analysis module, the prediction takes place whether the given data shows the traits of arrhythmia or not, i.e. the given data is Normal or Abnormal in nature.

D. Blockchain

The result obtained and the details of the patient are stored on the Blockchain where the hash code of the patient is generated respectively.

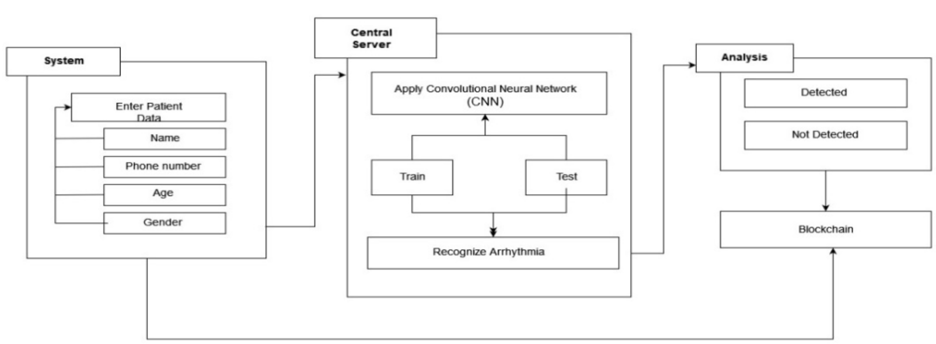


Fig. 5.1. Proposed Architectural System

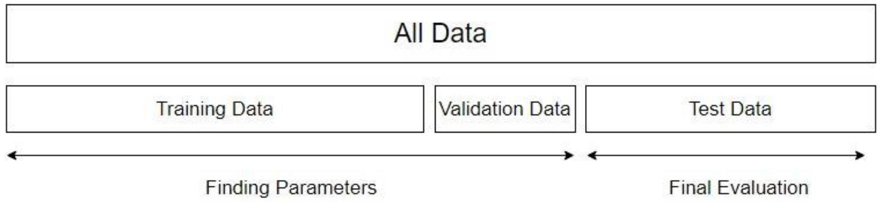


Fig. 5.2. Processing of Data

VI. IMPLEMENTATION

A. Enter the Patient Details

Name, Phone number, Age , Gender and the ECG file report in .csv format. This is designed using Java swing.

B. Reading the Dataset

This signal generator then sends the csv file to the CNN algorithm for testing. We will read out dataset using pandas.

C. Exploratory Data Analysis

We will pre-process our dataset and extract the parameters from the column as the model will understand numerical value. Dataset has 110k records each record is 187 values(which are recorded at 125hz sampling rate means 125 samples for 1 second of heart beat Testing (800 samples of each class => 4k for testing), Training (112k-4k=108k training)

Layer (type)	Output Shape	Param#	Connected to
input_1 InputLayer)	(None, 187, 1)	0	
conv1d_1 (Conv1D)	(None, 183, 32)	192	input_1[0][0]
conv1d_2 (Conv1D)	(None, 183, 32)	5152	conv1d_1[0][0]
activation_1 (Activation)	(None, 183, 32)	0	conv1d_2[0][0]

Applying convolution layer 1,2,3 and so on applying 1d convolution 32 filters of 5x1 size each.

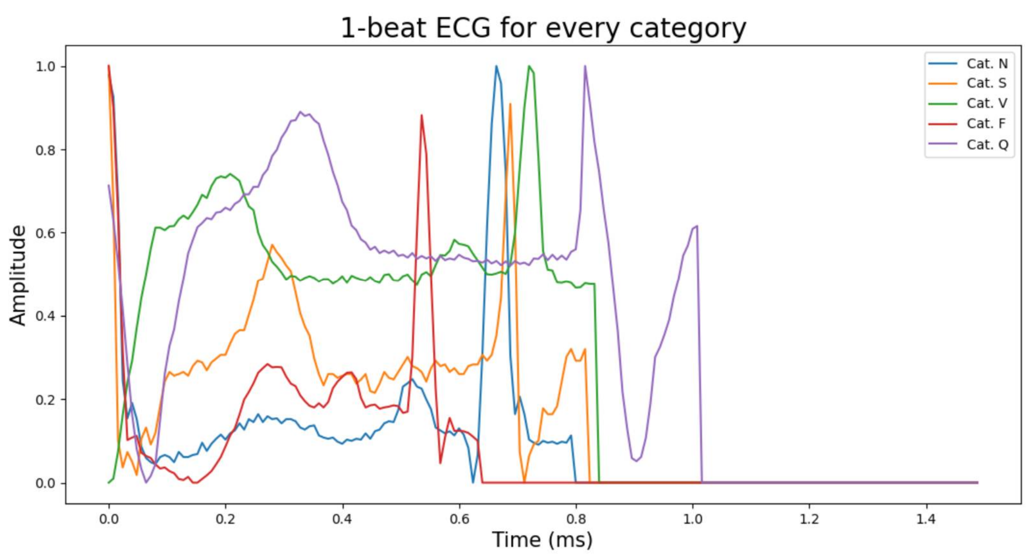


Fig. 6.1. Graph generation after training datasets

D. Handling Categorical Data

Output of CNN will be 5 values

for class 1 [1 0 0 0 0]
for class 2 [0 1 0 0 0]
for class 3 [0 0 1 0 0]
for class 4 [0 0 0 1 0]
for class 5 [0 0 0 0 1]

once the model is trained save it in model format, test the model using test data of 4k records.

E. Test Data: Performing EDA and Feature Engineering

Printing the classification report -

	Class1	Class2	Class3	Class4	Class5
Class1	1	0	0	0	0
Class2	0	1	0	0	0
Class3	0	0	1	0	0
Class4	0	0	0	1	0
Class5	0	0	0	0	1

Printing the confusion matrix -
Confusion matrix consists 4 params derived from classification report for each class,
TP Interpretation: You predicted positive and it's true.
TN Interpretation: You predicted negative and it's true.
FP Interpretation: You predicted positive and it's false.
FN Interpretation: You predicted negative and it's false.

Class1

Predicted	No	Yes
Actual		
No	TP	FP
Yes	FN	TN

Once the prediction is done by CNN the data is sent back to the java signal generator file generating signal according to data predicted that is the patient normal or abnormal and displays the result in details in the form of p, q, r, s intervals.

The predicted data is sent to the IPFS.

VII. RESULT

In our experimental setup, as shown in the figure 4.1, when entered the details of the patient for the login purpose a hash code is generated simultaneously.
The hash code her depicts the unique identity of the patient, that no external third party can alter it. Parallelly when we upload the ECG file after login, the content from it is trained using the CNN algorithm and a corresponding graph is generated of the reports. The IPFS algorithm splits the file into chunks and these chunks are identified as a cryptographic hash. In continuation the result is displayed whether the given patient is affected with arrhythmia or not.
Below depicted is an example of a unique hash code generated after login at the backend in Fig. 7.1.

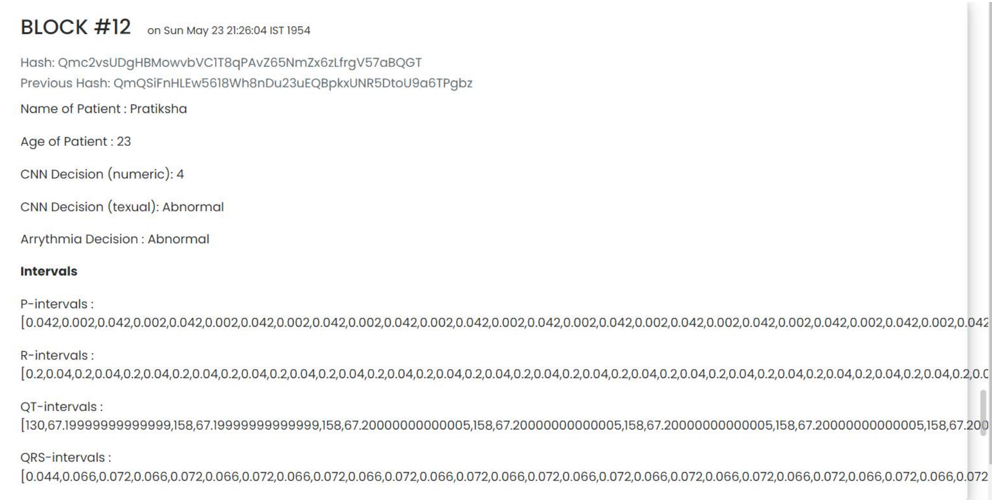
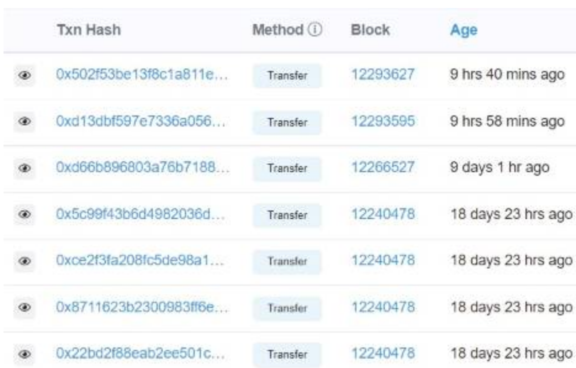


Fig. 7.1 Block History of a patient after prediction

In another scenario for multiple patient login the corresponding generation of hash code and the age of the block is mentioned in Fig. 7.2.



Txn Hash	Method	Block	Age
0x502f53be13f8c1a811e...	Transfer	12293627	9 hrs 40 mins ago
0xd13dbf597e7336a056...	Transfer	12293595	9 hrs 58 mins ago
0xd66b896803a76b7188...	Transfer	12266527	9 days 1 hr ago
0x5c99f43b6d4982036d...	Transfer	12240478	18 days 23 hrs ago
0xce2f3fa208fc5de98a1...	Transfer	12240478	18 days 23 hrs ago
0x8711623b2300983ff6e...	Transfer	12240478	18 days 23 hrs ago
0x22bd2f88eab2ee501c...	Transfer	12240478	18 days 23 hrs ago

Fig. 7.2. Blocks generation of patients

Sr No.	Block	Arrhythmia
1	12293627	Normal
2	12293595	Abnormal
3	12266527	Normal

Fig. 7.3. Prediction of Arrhythmia

A. *Accuracy*

In our proposed system, we are using the IPFS and CNN algorithm for the disease detection. The accuracy rate as compared to other models is mentioned below.

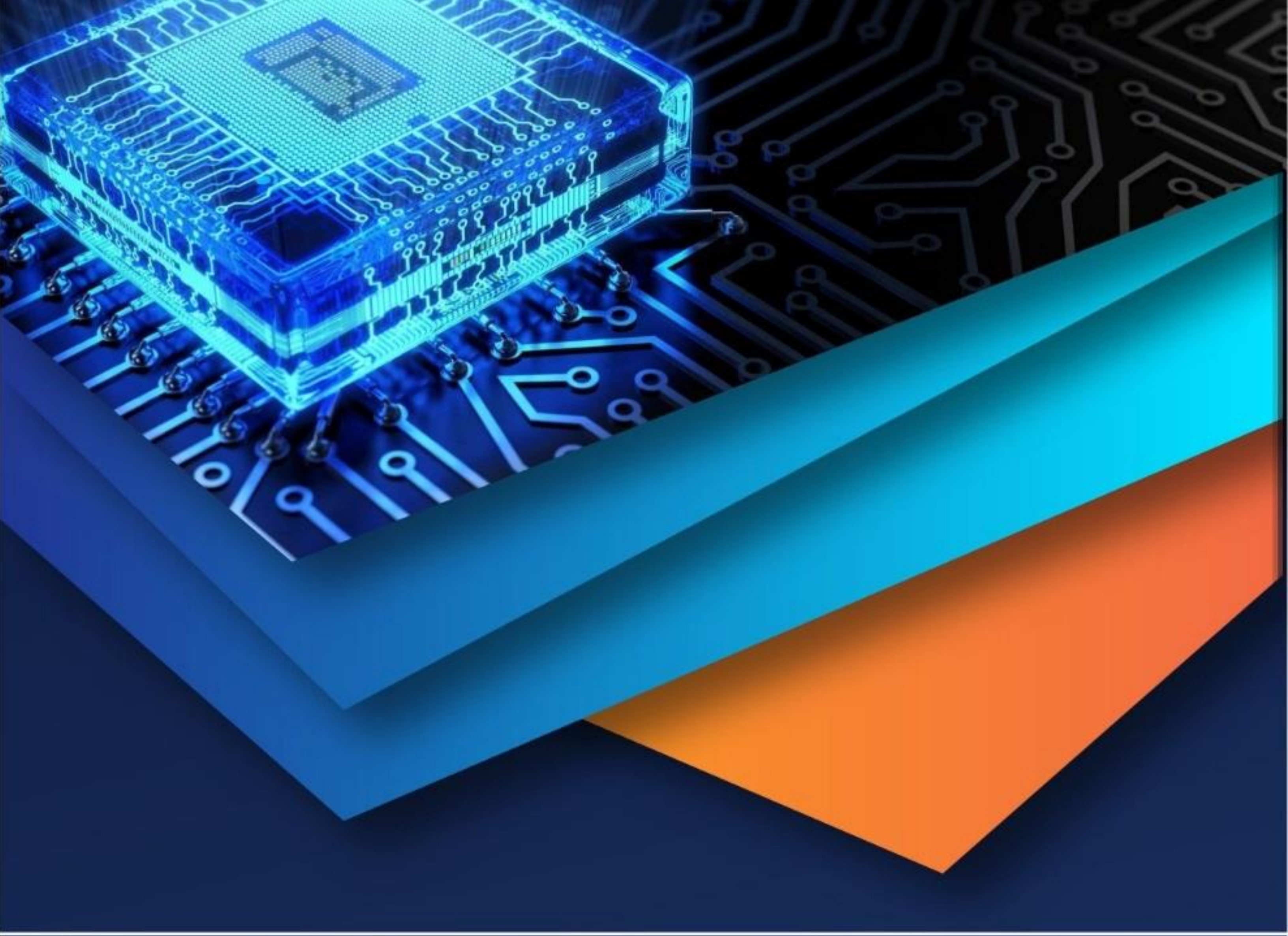
Algorithm	Accuracy Rate
SVM	83
KNN	82
Random Forest	78
CNN	84

VIII. CONCLUSION

Blockchain technology has the potential to dramatically improve and, in the long run, change how patients and physicians treat and use clinical records, as well as improve healthcare services. The employment of BC plays a crucial part in the present healthcare system. It could lead to automated processes for collecting and verifying data, correcting and aggregating information from various sources that are indisputable, resistant to manipulation, and provide protected data, as well as reduced cybercrime risks and support disseminated information with system redundancy. This research presents safe BC-based healthcare services for disease prediction in machine learning. Arrhythmia is taken into account for prediction. In comparison to other methods, the suggested work efficiently clusters and predicts the disease.

REFERENCES

[1] P. G. Shynu , Varun G. Menon: “Blockchain based secure healthcare application for Diabetic – cardio disease prediction in fog computing” (10 March 2021).
[2] E. Hope Weissler, Tristan Naumann, Tomas Andersson, Rajesh Ranganath, Olivier Elemento, Yuan Luo, Daniel F. Freitag, James Benoit, Michael C. Hughes, Faisal Khan, Paul Slater, Khader Shameer, Matthew Roe, Emmette Hutchison, Scott H. Kollins, Uli Broedl, Zhaoling Meng, Jennifer L. Wong, Lesley Curtis, Erich Huang, Marzyeh Ghassemi: The role of machine learning in clinical research: transforming the future of evidence generation (2021).
[3] Mercedeh J. Rezaei, John R. Woodward, Julia Ramírez and Patricia Munroe: A Novel Two-Stage Heart Arrhythmia Ensemble Classifier (2021).
[4] Sarkar Siddique and James C. L. Chow: Machine Learning in Healthcare Communication (2021).
[5] H. Das, B. Naik, and H. S. Behera, ‘‘Medical disease analysis using neuro-fuzzy with feature extraction model for classification,’’ Informat. Med. Unlocked, vol. 18, (2020), Art. no. 100288.
[6] R. Jayaram and S. Prabakaran, ‘‘Onboard disease prediction and rehabilitation monitoring on secure edge-cloud integrated privacy preserving healthcare system,’’ Egyptian Informat. J., to be published, doi: 10.1016/j.eij.2020.12.003.
[7] 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,’’ IEEE Trans. Consum. Electron., vol. 65, no. 3,pp. 359–368, (Aug. 2019).
[8] C. Beulah Christalin Latha, S. CarolinJeeva: Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques (2019).
[9] Arbaaz Bebal, Nomit Bhatnagar, Ankita Jagtap, Pratiksha Kamthe: Blockchain based Secure healthcare for cardio disease prediction, (2021)
[10] Abid Haleem, Mohd Javaid, Ravi Pratap Singh, Rajiv Suman, Shanay Rab: Blockchain technology applications in healthcare: An overview (2021)
[11] Ibrar Yaqoob, Khaled Salah, Raja Jayaraman & Yousof Al-Hammadi: Blockchain for healthcare data management: opportunities, challenges, and future recommendations(2021)
[12] Sarkar Siddique and James C. L. Chow: Machine Learning in Healthcare Communication (2021)
[13] R. Kumar, R. Tripathi, Scalable and secure access control policy for healthcare system using Blockchain and enhanced Bell–LaPadula model, Journal of Ambient Intelligence and Humanized Computing 12 (2) (2021 Feb) 2321–2338
[14] S. Balasubramanian, V. Shukla, J.S. Sethi, N. Islam, R. Saloum, A readiness
[15] assessment framework for Blockchain adoption: a healthcare case study, Technol.Forecast. Soc. Change 165 (2021 Apr 1)
[16] Z. Wang, N. Lou, and P. Zhou: “Guard Health: Blockchain empower ed secure data management and graph convolutions network enabled anomaly detection in smart healthcare”.
[17] Electronic ISSN: 2169- 3536, IEEE healthcare-based BLOCKCHAIN (SRHB) approach, Mobile Network. Appl. 25 (4) (2020 Aug) 1330–1337.
[18] R.B. Fekih, M. Lahami, Application of blockchain technology in healthcare: a comprehensive study, in: International Conference on Smart Homes and Health Telematics, Springer, Cham, (2020 Jun 24), pp. 268–276. 120536.
[19] N. Tariq, A. Qamar, M. Asim, F.A. Khan, Blockchain and smart healthcare security: a survey, Procedia Computer Science 175 (2020 Jan 1) 615–620.
[20] Seyednima Khezr , Md Moniruzzaman, Abdulsalam Yassine and Rachid Benlamri: Blockchain Technology in Healthcare: A Comprehensive Review and Directions for Future Research (2019).
[21] A. Shahnaz, U. Qamar, A. Khalid, Using blockchain for electronic health records, IEEE Access 7 (2019 Oct 9) 147782–147795.
[22] A.A. Siyal, A.Z. Junejo, M. Zawish, K. Ahmed, A. Khalil, G. Soursou, Applications of blockchain technology in medicine and healthcare: challenges and future perspectives, Cryptography 3 (1) (2019 Mar)
[23] R. Jayaraman, K. Salah, N. King, Improving opportunities in healthcare supply chain processes via the internet of things and blockchain technology, in: International Journal of Healthcare Information Systems and Informatics (IJHISI), vol. 14, (2019 Apr 1), pp. 49–65, 2.
[24] S. Chakraborty, S. Aich, H.C. Kim, A secure healthcare system design framework using blockchain technology, in: In 2019 21st International Conference on Advanced Communication Technology (ICACT), IEEE, (2019 Feb 17), pp. 260–264
[25] K.M. Hossein, M.E. Esmaeili, T. Dargahi, Blockchain-based privacy-preserving healthcare architecture, in: In2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), IEEE,(2019 May 5), pp. 1–4.
[26] T. Mikula, R.H. Jacobsen, Identity and access management with Blockchain in electronic healthcare records, in: In2018 21st Euromicro Conference on Digital System Design (DSD), IEEE, (2018 Aug 29), pp. 699–706.
[27] W.J. Gordon, C. Catalini, Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability, Comput. Struct. Biotechnol. J. 16 (2018 Jan 1) 224–230.
[28] Griggs KN, Ossipova O, Kohlios CP, Baccarini AN, Howson EA, Hayajneh T (2018) Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. J Med Syst 42(7):130
[29] X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li, ‘‘Integrating blockchain for data sharing and collaboration in mobile healthcare applications,’’ in Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC), (Oct. 2017), pp. 1–5.
[30] J. Daniel, A. Sargolzaei, M. Abdelghani, S. Sargolzaei, B. Amaba, Blockchain technology, cognitive computing, and healthcare innovations, J. Adv. Inf. Technol. 8 (3) (2017 Aug)
[31] Mengji Chen, Taj Malook Ateeq Ur Rehman, Yar Muhammad, Mohammad Dahman Alshehri, Aamir Akbar, Muhammad Bilal, Muazzam A. Khan- “Blockchain – Enabled healthcare system for detection of diabetes”(2021).
[32] Ahmed Farouk, Amal Alahmadi, Shohini Ghose, Atefeh Mashatan – “Blockchain platform for industrial healthcare: Vision and future opportunities. ” (2020)



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY