



IJCRR

Section: Healthcare

ISI Impact Factor
(2019-20): 1.628

IC Value (2019): 90.81

SJIF (2020) = 7.893



Copyright@IJCRR

Machine Learning and Data Analytics based Analysis for Heart Disease Prediction

Patil Rina S.¹, Gangwar Mohit²

¹PhD Research Scholar, Department of Computer Science and Engineering, Bhabha University, Bhopal, Madhya Pradesh, India; ²Dean-Engineering, Bhabha University, Bhopal, Madhya Pradesh, India.

ABSTRACT

The heart is the subsequent major organ connecting to the brain, which has stronger precedence in the human evidence. It elevates the blood and accumulations to all devices of the entire body. Forecast of circumstances of heart conditions in the medical field is important to work. Data analytics is beneficial for divining more knowledge, and it helps the medical centre predict various conditions. A huge number of patient-related data is prepared each month. The collected data can be beneficial for the source of predicting the emergence of future weaknesses. Unusual data mining and machine learning procedures have been used to indicate heart disease. This research proposed heart disease prediction using various modified Recurrent Neural Network (mRNN) deep learning algorithms. Numerous feature extraction and selection methods have been used to get important features and data collection using custom-generated IoT environments. The system effectively provides heart risk scores with the highest accuracy in a runtime environment.

Key Words: Internet of Things, Deep Learning, Feature selection, Feature extraction, Optimization, Classification, Supervised learning

INTRODUCTION

Health data mining provides immense promise to discover the secret trends in the scientific domain names data sets. For psychiatric disorders, these structures can be used. The providing raw patient records are therefore widely dispersed, homogenous, and weighty. It is useful to acquire such data in a structured manner. The health monitoring system, data gathered can then be incorporated. A user-oriented methodology to new and secret trends in the information is generated by business intelligence. It has been estimated from the WHO that 12 million people die internationally each year due to heart failure.

Related to cardio-related disease, half the deaths in the USA and other developing nations occur. It is also the primary cause of casualties in many developed countries. Most of all, it is considered the main reason for deaths worldwide. The word High Cholesterol includes the different diseases that cause the condition. The main cause of fatalities in various countries such as Sri Lanka was breast cancer. In the Western World, heart illness kills one adult every 34 seconds. Oth-

er types of heart disorders include coronary heart disease, cardiomyopathy and cardiovascular disease. A wide variety of conditions influencing the cardiovascular framework and how air is filtered and distributed through the body are included in the word 'cardiovascular disease.' Several diseases, injuries and deaths arise from cardiovascular disease (CVD). A critical and complicated role in medication is the diagnosis of disorders.

Psychological care is an essential but complex task that remains to be improved precisely and effectively. This program's automation would be highly beneficial. Regrettably, in any sub-speciality, all physicians do not have the experience. Besides, there is a lack of staff at some sites. Therefore, putting all of them together, and an automated medical diagnosis scheme would possibly be extremely helpful. Reasonable judgment on computer-based knowledge Help networks will support the achievement of clinical studies at a lower cost. A comparative analysis of the different technologies applied includes effective and precise implementations of the electronic machine. This paper aims to analyze the

Corresponding Author:

Gangwar Mohit, Dean-Engineering, Bhabha University, Bhopal, Madhya Pradesh, India.

Email: mohitgangwar@gmail.com

ISSN: 2231-2196 (Print)

ISSN: 0975-5241 (Online)

Received: 04.01.2021

Revised: 02.04.2021

Accepted: 09.05.2021

Published: 11.10.2021

various statistical and machine learning strategies for evaluating machine learning introduced in the past months.

REVIEW OF LITERATURE

This section focused on the study of various existing systems, and the work has done by various existing research in IDS

A new deep learning model for sex and age identification using a standard inertial sensor is suggested by Sun et al. (2019).¹ This approach is assessed with data obtained from more than 700 individuals using the highest sensor-based training repository value. Ten experiments of Monte-Carlo test sets were given to confirm the robustness and efficacy of the presented method. In protection applications, such as cell phone authorization, biometrics has been commonly embraced. The reliability of hard biometrics-standard physiological properties like face, including fingerprint and behavioural factors such as visual features and voice-is also the subject of these implementations. A supervised neural technique uses a single inertial sensor attached to the lower back of gender and age recognition issues. Machine learning-based techniques have been commonly used throughout vision-based gender detection. Still, they haven't been used for electromagnetic sensor-based gender acknowledgement to our best understanding.

According to Lo et al.², a point fulfilment network utilizes a specific dimensional image taken from every suitable angle of view to perform 3d model of objects. The suggested technique has resolved many primary problems in vision-based disease diagnosis relative to previous methods, such as view deformation and scale uncertainty. For testing this method, various experiments were carried out. Still, the feasibility of SVM and NB, the algorithm was demonstrated in the robust prediction. Tested objects have asked to remember the complete description of food products eaten over the last 24 hours and the corresponding portion sizes to measure the calorie consumption. The size of both the food component relies mainly on the entity's personal opinions, which is often too unreliable. Under the circumstances of view occlusion, this approach demonstrates the effectiveness and accuracy of food volume estimation. The finished pixel value of the obstructed food products can be accessed by using the point fulfilment network.

According to Uddin and Hassan³, a comprehensive approach to healthcare monitoring activity recognition using body receptors and a complex Convolutional Neural Network called CNN. They analyze signals from various body sensors in healthcare services, like ECG, wearable sensors, and control system sensors. A CNN Based operation is equipped based on the features after removing essential characteristics from the wearable sensors based on PCA. Eventually, the deeply

educated CNN was used for the identification of data testing events. A publicly accessible standard dataset is applied to the defined solution and then contrasted with other traditional methods. The experimental of these system findings suggests that the counsel made is superior to others, showing both approaches' robustness to be implemented in the environment sensor-based intelligent health systems for behavioural assisting. As a growing field of study, esteemed solutions to implementing a comprehensive smart healthcare delivery system for individuals to extend their independent lives must now be explored. The detectors were positioned on the neck, right wrist, or left ankle of the subject, separately, and connected using elastic bands. Several sensors help us quantify the movement experienced by different body parts, including momentum, turn frequency, and direction of the gravitational flux, thereby better measuring the body's nature.

According to Ganesan and Sivakumar,⁴ a fresh Cloud and IoT-based Healthcare application were developed to monitor and diagnose serious diseases. In the training phase, the classifier was trained using the data from the benchmark dataset. During the testing phase, the actual patient data to identify disease was used to identify the presence of disease.

According to Majumder et al.,⁵ proposed a multiple sensors system using a smart IoT that gives an early warning of disease risk, constantly collects the data from the user, and sends it to the smartphone via Bluetooth using Body Area Sensor (BAS) system. All the processing and data analysis took place in the application to view real-time user plots of future cardiac arrest. An IoT system with a low power consumption communication model developed regularly collects body temperatures and heart rate using a smartphone. Here ML and signal processing techniques were used to analyze sensor data and predict high accuracy cardiac arrests. A wearable device was implemented based on a smartphone for heart rate detection. It used a combination of ECG and body temperature. A heart rate analysis is done on the Android platform where users can view body temperature and plots of real-time ECG signals.

According to Yu et al.⁶ an adversarial training approach is proposed to multitask learning to estimate multi-type Cardiac Indices in MRI and CT. By using multitask learning networks, these task dependencies are shared and learned. Lastly, they transferred parameters learned from MRI to CT. A series of experiments were performed. The authors first optimized system efficiency over 2900 cardiac MRI images through ten-fold cross-validation. Then the network was run on an independent data set with 2360 cardiac CT images. The results of all experiments on the proposed reverse mapping indicate excellent performance in estimating multiple cardiac indices.

Ali et al.⁷ implemented an optimally designed OCI- DBN, a deep belief network to resolve these problems and boost

system performance. They used Ruzzo-approach Tompa's to delete features that do not contribute enough to boost processing speed. They stacked evolutionary algorithms stacking two genetic algorithms to have an optimally designed DBN to obtain maximal configuration settings. To provide an insight into how the system operates, an RBM and DBN study was conducted.

According to Kumar et al.⁸ cloud-based and IoT-based m-healthcare applications have been produced and updated to observe and diagnose the genuine level of severity. To gather medical data in remote areas, such instruments are used. The approximate estimate can be collected as vital information, connected to the human body, gained from IoT devices. They were using the UCI Repository dataset and clinical sensors to foresee the average person with a major influence on non-communicable diseases medicinal knowledge is produced. The resulting information is securely stored by implementing a new federal storage method in five different steps, such as data storage, data recovery, data gathering, database division, and file merging.

According to Park et al.⁹ a smart wheelchair method was installed that monitors but visualizes the user's location through a mobile app that attempts to address the inconsistent position of the user. They used strain gauges and shift sensing, using IoT and Wireless techniques to communicate with low consumption. It's an extension of Arduino that detects various user stances. This integration complements the user by offering real-time interactive and visualized data for mobile apps to sit properly and understand their present incarnation. With pressure displayed in red, yellow, pink and green triangles. This is a great example of Information systems.

Kumar and Devi¹⁰ introduced IoT's three-tier architecture with early detection using a supervised learning algorithm to detect cardiac disease. This system also suggested multi-tier architectures that stored and processed massive quantities of data from wearable IoT environments. First-tier focuses on data collection from numerous analogue sensors; Tier two demonstrates Apache Hadoop that stores vast quantities of data into the cloud storage, while Tier three uses Apache-based Mahout to build a logistics learning base detection as well as a prediction model. In conclusion, ROC demonstrates to get entire system analysis of heart disease prediction.

Nashif et al.¹¹ designed a disease prediction system based on the cloud. A real-time monitoring system sensing health parameters like blood pressure, temperature, heartbeat and humidity was developed using an Arduino microcontroller. The proposed system can detect heart disease using ML techniques as recorded data transmitted to a central server is updated every 10 seconds.

Santhi and Renuka¹² implemented an optimized prediction model using a genetic algorithm. They studied various prediction models and important feature selection algorithms.

Its performance is better than other traditional prediction models. The various prediction models were retested with heart disease data sets and validated with real-time data sets. The n-cross-validation methodology is applied to produce balanced training and testing data set.

Alotaibi¹³ used the Rapid miner tool and various ML approaches to improve the previous accuracy score and to predict heart disease. UCI heart disease dataset was tested. The proposed work improved the previous accuracy score.

Proposed works

a. System Architecture

According to many factors, the medical remote management forecast period demand has expanded significantly. When it is very common today in developing countries that older adults typically live separately in their own homes, the elderly population is growing. Besides, the Internet of Things (IoT) makes these remote patient monitoring systems theoretically feasible (IoT as the idea of a capable and modifiable environment of surveillance in which controls or actuators to human and non-living particles) and makes it financially viable because of the even lower expense of sensors. Because of the advent of smart mobile technology, this is also understood that consumers are now prepared to embrace these types of solutions that capture private and sensitive data from people in real-time, such as temperatures, blood glucose, breathing, mass spectrometry sensor, to name a few. For example, personal health observers such as smart beds instantly notify who is utilizing them. More, they are sure to enlighten different patients' physical levels, making real smart home drug dispensers automatically warn when medication is not taken, for example. Several virtual care monitoring devices use various technologies inside facilities and their residences to monitor and/or track clinicians or diagnostic applications. Regrettably, most of these systems are not versatile when new sensors are introduced during runtime, as far as we know. Neither has it allowed regular users with the smart technologies added to generate ad-hoc notifications automatically.

Distributed data processing is one of the processes involved that can resolve a number of the protection, distribution, integration, and management challenges of aggressive data innovation. For a genuine analysis and diagnosis system that results in incredible medicinal implementations, the primary requirement for continuous and limitless access to patient details from anywhere and from any computerized gadget is necessary. This propels current research towards thinking about pervasive clinical frameworks focused on the internet. The Predictive Analytic System hardware implementation predicts that disease relative to body temperature, pulse rate, and individual tension. The sensors are mounted on the human chest to monitor the physician's metabolic rate, heart

rate, cholesterol levels. It sends data from the information it predicts and diagnoses the disease to the Arduino microcontroller. This information is hosted in the cloud via the end receiver. It can be monitored and anywhere whenever through the internet. The condition of the physician will be registered on the database and retained. A smartphone framework for machine learning is designed for viewing data from sensors and sending emergency alerts. Healthcare practitioners can track, anticipate, diagnose, and inform their clients at any moment by using this method.

The conceptual system can control several predictor variables on wearable technology, integrated with such a handheld device. Sensor networks work together again to obtain the necessary data, except encountered some problems. It operates by reducing less human involvement to preserve its precision. Using cardiovascular sensors is a legitimate diagnostic device for remote areas heart-prone clinicians that monitor heart rate, cholesterol levels, metabolic rate, and many other metrics. N number of measurements obtained can be preserved and viewed subsequently by practitioners to correct acute and chronic treatment as depicts in Figure 1.

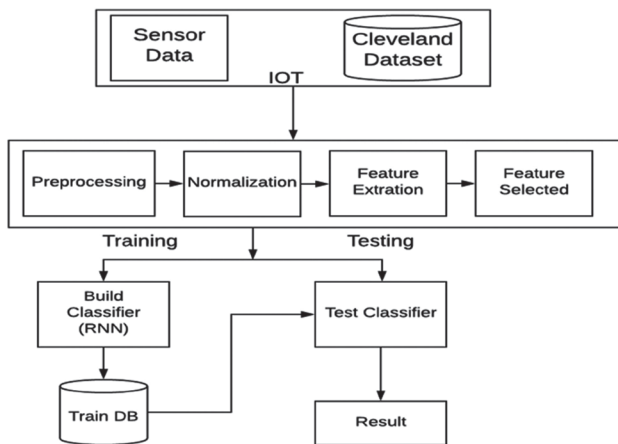


Figure 1: Proposed System architecture.

The methodology of computer vision will sensitize and evaluate the teaching context to inculcate heart performance. Two major physician and patient interfaces would allow data to be transmitted to each other. In actual environments, this system recognizes fundamental heart problems for the sake of the health of the individual. It produces alarm in the sense of Text messaging, email, etc. based between high and low input parameters set a right for a physician and any close relative registered liable. It'll also enrich that system with a sense of focus and consideration that adds value to the condition of the individual.

Algorithm Design

Algorithm 1: Proposed modified Recurrent Neural Network Algorithm (mRNN)

Input: Train_Feature set {} which having values of numeric or string of training dataset, Test_Feature set {} which having values of numeric or string of test dataset, Threshold denominator Th, Collection List cL.

Output: classified all instances with the desired weight.

Step 1: Read all features from the Testing dataset using the below function

$$\text{Test_Feature} = \sum_{j=1}^n (T[j])$$

Step 2: Read all features from the training dataset using the below function

$$\text{Train_Feature} = \sum_{k=1}^m (T[k])$$

Step 3: Read all features from Trainset using below

Step 4: Generate the weight of both feature set

$$\text{Weight} = \text{classifyInstance}(\text{Train_Feature}, \text{Test_Feature})$$

Step 5: Verify with Th

$$\text{Selected_Instance} = \text{result} = \text{Weight} > \text{Th} ? 1 : 0;$$

Add each selected_instance into cL, when n = null

Step 6: Return cL

RESULTS

The proposed research focuses on the soft computing approach and classification based recognition, fundamentally this approach consuming the respectable exposure rate, nevertheless generating sometimes more false-positive ratio. Figure 2 below displays the 3-fold classification technique used on all parameters and explains all implementation consistency.

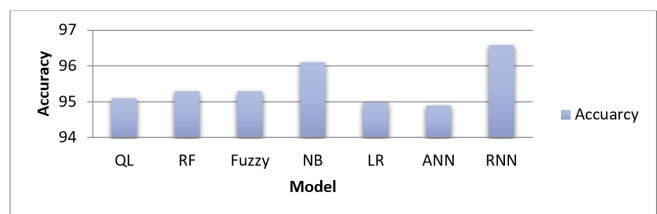


Figure 2: Classification accuracy of the system with various machine learning algorithms.

The above Figure 2 improves the importance of different experimental research focusing on various statistical tests with seven distinct algorithms: Fuzzy Logic, Q-Learning, Naïve Bayes, Linear Regression, ANN and Random Forest with Recommended Perceptron Algorithm.

DISCUSSIONS

The proposed system makes sense to interrelate quantitative assessment and clinical scores and could potentially address many challenges in decision-making. The generated training repository was applied with six machine learning models to establish patterns of common, questionable and dangerous activities. To test and rate the machine learning strategies, the 3-fold, 5-fold and 10-fold cross-validation model was employed using the behaviour classification, training database. For data management, the mRNN classification algorithm was used during the classifier. The neural network was shown and debated for each model. Both uncertainty metrics show the system accuracy of properly classifying, wrongly classifying, recession, and device recall.

CONCLUSIONS

This study introduced an ideal framework for prediction models in real-time. Users with the coronary disease can use that. Unlike many other systems, it is capable of tracking and predicting both. The system's diagnostic method will predict cardiovascular disease utilizing ML algorithms. The predictive conclusions are based on the dataset example of heart disease. On the other hand, the device is very economical; we used an enthused pulse sensor and sent the data via the Arduino suite microcontroller to mobile devices. To check the variances and raise the alarm if the patient's heart rate increases above the usual heart rate. To prove the system's efficacy, we performed tests with both the tracking and diagnostic method. We conducted experiments with supervised machine classification methods such as QL, Linear Regression, Random Forest, Naive Bayes, ANN and Fuzzy Logic, etc. The procedure was done with the validation set test, and 89% efficacy of the proposed method was obtained for mRNN. To extract various feature extraction and selection with embedded deep learning techniques will be an interesting future work for the system.

ACKNOWLEDGEMENT

The authors acknowledge the immense help received from the scholars whose articles are cited and included in references of this manuscript. The authors are also grateful to authors/editors/publishers of all those articles, journals and books from where the literature for this article has been reviewed and discussed

Source of Funding: Self.

Conflict of Interest: None.

Authors' Contribution: Ms. Patil performed the data analysis and developed the overall paper under the supervision of Dr. Gangwar

REFERENCES

1. Sun Y, Lo FP, Lo B. A deep learning approach on gender and age recognition using a single inertial sensor. In 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN) IEEE. 2019; 19:1-4.
2. Lo FP, Sun Y, Qiu J, Lo B. A novel vision-based approach for the dietary assessment using deep learning view synthesis. In 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN) IEEE. 2019 May 19 :1-4.
3. Uddin MZ, Hassan MM. Activity recognition for cognitive assistance using body sensors data and deep convolutional neural network. IEEE Sensors Journal. 2018 Sep 20;19(19):8413-9.
4. Ganesan M, Sivakumar N. IoT based heart disease prediction and diagnosis model for healthcare using machine learning models. In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN) 2019 Mar 29 (pp. 1-5). IEEE.
5. Majumder AK, ElSaadany YA, Young R, Ucci DR. An energy-efficient wearable smart IoT system to predict cardiac arrest. Advances in Human-Computer Interaction. 2019;2(12):2019.
6. Yu C, Gao Z, Zhang W, Yang G, Zhao S, Zhang H, Zhang Y, Li S. Multitask learning for estimating multitype cardiac indices in MRI and CT based on adversarial reverse mapping. IEEE transactions on neural networks and learning systems. 2020 Apr 16.
7. Ali SA, Raza B, Malik AK, Shahid AR, Faheem M, Alquhayz H, Kumar YJ. An Optimally Configured and Improved Deep Belief Network (OCI-DBN) Approach for Heart Disease Prediction Based on Ruzzo-Tompa and Stacked Genetic Algorithm. IEEE Access. 2020 Apr 3;8:65947-58.
8. Kumar PM, Lokesh S, Varatharajan R, Babu GC, Parthasarathy P. Cloud and IoT based disease prediction and diagnosis system for healthcare using Fuzzy neural classifier. Fut Gen Compt Sys. 2018 Sep 1;86:527-34.
9. Park M, Song Y, Lee J, Paek J. Design and implementation of a smart chair system for iot. In 2016 International Conference on Information and Communication Technology Convergence (ICTC). IEEE. 2016;19:1200-1203.
10. Kumar PM, Gandhi UD. A novel three-tier Internet of Things architecture with a machine learning algorithm for early detection of heart diseases. Compt Elect Engg. 2018 Jan 1;65:222-35.
11. Nashif S, Raihan MR, Islam MR, Imam MH. Heart disease detection by using machine learning algorithms and a real-time cardiovascular health monitoring system. World J Engg Tech. 2018 Sep 12;6(4):854-73.
12. Salem T. Study and analysis of prediction model for heart disease: an optimization approach using genetic algorithm. Int J Pure Appl Mathem. 2018;119(16):5323-36.
13. Alotaibi FS. Implementation of a machine learning model to predict heart failure disease. Int J Adv Compt SciAppl. 2019;10(6):261-8.