



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 1, January 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Diagnosis Disease Prediction for Smart Healthcare Systems using Machine Learning Algorithm

¹Shambhu Kumar, ²Prof. Santosh Nagar, ³Prof. Anurag Shrivastava

M. Tech. Scholar, Department of Computer Science and Engineering, NIIST, Bhopal, India¹

Assistant Professor, Department of Computer Science and Engineering, NIIST, Bhopal, India²

Head of Dept., Department of Computer Science and Engineering, NIIST, Bhopal, India³

ABSTRACT: - In today's world, one of the major threats to human health is Diabetes and Heart disease. Diabetes is a metabolic disease where a person suffers from increased sugar level, is because of either the pancreas does not produce sufficient insulin for the body or the cells do not respond to the insulin. The persistent diabetes leads to malfunction, injury and failure of different organs such as kidneys, eyes, nerves, blood vessels and heart. The criticalities of Heart diseases are more crucial and can even lead to vulnerable consequences if it is not detected at an earlier stage. The techniques such as electronic health records, body area networks are emerged to continuously monitor and diagnose patient's health conditions through the projection of medical sensors and wearable devices across human bodies. Machine learning provides a reliable and excellent support for prediction of a Diabetes and Heart disease with correct case of training and testing. In this paper studied of Diagnosis of diabetes and heart disease desires great support of different classifiers algorithm to detect diabetes and heart disease in early stage, since it cannot be cured which brings great complication to our health system.

KEYWORDS: -Machine Learning, Diabetes Disease, Heart Disease

I. INTRODUCTION

The term healthcare refers to a system that entails the enhancement of health-related services in order to fulfill the clinical requests of the individuals. In healthcare services, patients, doctors, clinicians, researchers and medical industries are all making an effort to maintain and restore health records. In recent years, with the remarkable development of technologies, data is continuously increasing day by day in every sector including healthcare which in turn demands more and more data mining applications. However, due to digitization of healthcare system, the medical organizations are generating large amount of healthcare data [1]. In general, healthcare data [2] consists of all records related to health stored in a digital form. It may contain detailed information about patients' medical history, doctors prescribed notes, clinical reports etc. All these data are voluminous, high dimensional and diversified in nature. Healthy decision making is a challenge in the modern era due to this rising complexity of healthcare data [1,2]. Machine learning, data mining, and statistical techniques are the key fields of study that enhance the ability of individuals to make the right decisions in order to maximize the outcome of any working domain [3]. Human data analytical capability rate is much smaller when compared to the amount of data that is stored [2]. This becomes even more critical when it comes to healthcare domain as the number of available experts for healthcare data analysis is comparatively less. Therefore, there is a requirement for computerized (semi-automatic) Medical Disease Diagnosis Systems (MDDSs) that can empower doctors and medical professionals to make informed (better) healthcare decisions for any individuals. This will enhance the quality of care, improve the diagnosis of diseases and ultimately reduce the healthcare costs. The present study is mainly focused on classification/prediction problems of healthcare data based on machine learning (supervised) approaches. There are several learning algorithms which can be used with data mining techniques to resolve classification problems in healthcare domain so as to increase the diagnostic speed, accuracy and reliability.

II. RELATED WORK

Romany Fouad Mansour et al. [1], have new headways in Internet of Things (IoT), distributed computing, and Artificial Intelligence (AI) changed the customary medical services framework into savvy medical care. By

consolidating key innovations, for example, IoT and AI, clinical benefits can be moved along. The union of IoT and AI offers various open doors in medical services area. In this view, the ebb and flow research article presents another AI and IoT union based illness analysis model for brilliant medical care framework. The significant objective of this article is to plan a sickness analysis model for coronary illness and diabetes utilizing AI and IoT union strategies. The introduced model incorporates various stages to be specific, information procurement, preprocessing, characterization, and boundary tuning. IoT gadgets, for example, wearables and sensors license consistent information assortment while AI strategies use the information in sickness conclusion. The utilization of CSO helps in significant improvement in the symptomatic results of CLSTM model. The presentation of CSO-LSTM model was approved utilizing medical services information. During the trial and error, the introduced CSO-LSTM model achieved the greatest correctnesses of 96.16% and 97.26% in diagnosing coronary illness and diabetes individually. Consequently, the proposed CSO-LSTM model can be utilized as a suitable illness finding device for brilliant medical services frameworks.

G. Muhammad et al. [2], pathology recognition framework is proposed in this review. The framework utilizes a profound convolutional network comprising of 1D and 2D convolutions. Highlights from various convolutional layers are combined utilizing a combination organization. Different sorts of organizations are examined; the sorts incorporate a multi-facet perceptron (MLP) with a fluctuating number of stowed away layers, and an autoencoder. Tests are finished utilizing an openly accessible EEG signal data set that contains two classes: ordinary and strange. The exploratory outcomes show that the proposed framework accomplishes more noteworthy than 89% precision utilizing the convolutional network followed by the MLP with two secret layers. The proposed framework is additionally assessed in a cloud-based structure, and its exhibition is viewed as practically identical with the presentation acquired utilizing just a nearby server.

A. A. Mutlag et al. [3], medical services applications, various sensors and gadgets produce huge measures of information which are the focal point of basic errands. Their administration at the edge of the organization should be possible by Fog figuring execution. Mist Nodes could perform only few assignments. A tough choice worries which undertakings will perform locally by Fog Nodes. Every hub should choose such assignments painstakingly founded on the current context oriented data, for instance, errands' need, asset burden, and asset accessibility. We recommend in this paper a Multi-Agent Fog Computing model for medical services basic errands the executives. The fundamental job of the multi-specialist framework is planning between three choice tables to improve booking the basic undertakings by appointing errands with their need, load in the organization, and organization asset accessibility. The initial step is to conclude whether a basic undertaking can be handled locally; in any case, the subsequent advance includes the modern choice of the most reasonable neighbor Fog Node to assign it. Assuming no Fog Node is equipped for handling the errand all through the organization, it is then shipped off the Cloud confronting the most noteworthy inertness.

M. S. Hossain et al. [4], new age correspondence advances and progressed profound learning models present a gigantic chance to foster quick, precise, and consistent conveyed frameworks in various areas including the medical services area. in this article, we recommend a savvy medical care structure comprising of a pathology location framework, which is created utilizing profound learning. The pathology can be identified from electroencephalogram signs of a subject. in the system, a savvy EEG headset catches EEG flags and sends them to a versatile edge figuring server. The server preprocesses the signs and communicates them to a cloud server. The cloud server does the principle handling utilizing profound learning and settles on whether or not the subject has pathology. Clients and partners of the structure are associated by means of a confirmation chief situated in the cloud server. Analyze results on a freely accessible data set affirm the fittingness of the proposed system.

M. S. Hossain et al. [5], expanding interest for computerized, remote, shrewd, and constant medical care administrations in savvy urban areas, brilliant medical care observing is important to give improved and complete consideration to inhabitants. In this checking, wellbeing related media or signs gathered from brilliant gadgets/objects are sent and handled to take care of the requirement for quality consideration. Be that as it may, it is trying to make a structure or technique to deal with media-related medical services information investigation or signs (e.g., voice/sound, video, or electroglottographic (EGG) motions toward) satisfy the complex on-need medical care needs for effective brilliant city the board. To this end, this paper proposes a cloud-arranged brilliant medical care checking system that cooperates with encompassing savvy gadgets, conditions, and shrewd city partners for reasonable and available medical services. The info gadgets are associated with the Internet and the caught signals are sent to the cloud. For order, a Gaussian combination model-based methodology is utilized. Trial results show that the proposed strategy can accomplish VPD that is over 93% exact.

V Krishnapraseeda et al. [6], Diabetes mellitus is a persistent infection related with unpredictable evident levels of the sugar glucose in the blood and it's everything except a huge general clinical issue. Diabetes has turned into the 4th driving defense end in made countries. Regardless of the way that couple of systems have been made to predict this persevering contamination, there is a necessity for imaginative techniques which might help in early assumption for diabetes and its disarrays. This investigation work targets encouraging a convincing diabetes figure system by taking advantage of AI. In this examination, specific AI methods are applied to perform perceptive assessment over the diabetes educational lists. The proposed framework makes progress toward better counteraction, analysis and the executives of Type 2 diabetes and spotlights on creating AI methodologies for exactness medication. It's anything but a stage not just for early acknowledgment of Type 2 diabetes yet additionally counteraction of unsafe inconveniences. The proposed framework permits the experts to foresee all the more decisively which treatment and counteraction frameworks for diabetes will function as the best for determined patients. Consequently, the framework will significantly add to the nature of medical services and useful generally of the present local area.

Valasapalli Mounika et al. [7], type 2 diabetes mellitus (DM) is an ongoing condition whose pervasiveness has been bit by bit developing across the world. Around 30 million individuals in India have Diabetes, and are a lot more in danger. Consequently, to block diabetes and its indications identified with it, so early conclusion is required. The thought behind using various techniques for speculative assurance of Type-2 Diabetes subject to demonstrative examination the improvement of the determination time of the infection through the route toward evaluating intriguing characteristics and step by step propensities, allowing the assessing of Type 2 Diabetes without the need of clinical tests through perceptive investigation. A monstrous proportion of clinical data is available today regarding the contamination, their signs, purposes behind disorder, and their ramifications for prosperity. Utilizing different AI strategies, Since these calculations are so exact, the danger of Type-2 Diabetes can be prevised, which is essential for the clinical field.

I.K. Mujawar et al. [8], a counterfeit neural organization began to be utilized in clinical cycles and issues when it was guaranteed that these cycles and issues may be intricate in nature and that would be best taken care of by nonlinear methodologies. A fake neural organization has been utilized and applied in extraordinary numbers for clinical and pathobiological measures with colossal achievement. It is utilized in sickness cycles to address its each subspecialty. Infection finding is one of the significant and profoundly mindful biomedical regions where a counterfeit neural organization has been utilized by and large. In the proposed work, patient's information was gathered from clinic and utilized in the framework plan. The proposed work presents an engineering, plan and improvement of the online framework for a diabetes conclusion. The proposed framework is planned with an open source advancement climate which will be valuable for individuals in the Diabetes Diagnosis measure.

Cecilia Saint-Pierre et al. [9], transcendence of type 2 diabetes mellitus (T2DM) has almost increased in late numerous years and normally presents comorbidities and intricacies. T2DM is a multisystemic ailment, requiring multidisciplinary treatment given by bunches working in an arranged and synergistic manner. The utilization of casual association examination techniques in the clinical consideration region has allowed researchers to take apart relationship among specialists and their positions inside care gatherings. We analyzed whether the development of care gatherings, showed as astounding casual associations, is connected with patient development. For this, we show a data driven strategy and use existing relational association assessment estimations and estimations proposed for this investigation. We inspected course of action and HbA1c blood test result data from patients treated at three fundamental clinical benefits living spaces, tending to six interesting practices. Patients with extraordinary metabolic control during the researched period were treated by bunches that were more instinctive, synergistic and multidisciplinary, however patients with disintegrating or unstable metabolic control were treated by bunches with not such a lot of collaboration but instead more movement breakdowns. Results from the proposed estimations were dependable with the past composition and reveal critical pieces of collaboration and multidisciplinary.

Problem Formulation

Now-a-days, diabetes and heart disease is considered one of the key reasons of death among the people in the world. Diabetes and heart disease is the most common endocrine disease. The disease is characterized by metabolic abnormalities and by long-term complications involving the eyes, kidneys, nerves and blood vessels. The diagnosis of symptomatic diabetes is not difficult. When a patient presents with signs and symptoms attributable to an osmotic diuresis and is found to have hyperglycemia essentially all physicians agree that diabetes and heart disease is present. At present the diagnosis of diabetes cure does not exist for the diabetes, and then only option is to take care of the

health of people affected, maintain their glucose levels in the blood to the nearest possible normal values. The availability of extensive medical information leads to the search for proper tools to support physicians to diagnose diabetes mellitus disease accurately. This research aimed at improving the classification accuracy of diabetes mellitus and reducing diagnostic miss-classification based on the extracted significant diabetes features on machine learning classifiers.

Following are the problems which are to be considering as identify from the Base Paper:

1. In medical science diagnosis of disease data involves of a number of medical tests which are needed to diagnose a certain disease and the diagnosis are depend on the surgeon experience, if a less experience surgeon can diagnose a problem incorrectly.
2. The number of False-positives is quite high in some specific cases, which can be further reduced.
3. The issue of overcrowding is a global issue and a public health problem that affects individuals irrespective of one's social class.
4. The occurrence of this problem is attributed to certain factors, including the inflow and outflow of patients, patients' waiting time, and efficiency of healthcare providers, agglomerations, and availability of resources.

III. MACHINE LEARNING ALGORITHM

Uproarious information is available in the heap of substance that will be identified through the anomaly strategies. The information can be spatial or can be a transient method spatial connected with the geological conditions and worldly connected with the time perspectives [14, 15]. The principle point of exception identification is to deal with the loud information that is introduced in the heap of text. Different methods for recognizing abnormalities in Text are specified in below:

Learning

The main property of an ML is its capability to learn. Learning or preparing is a procedure by methods for which a neural system adjusts to a boost by making legitimate parameter modifications, bringing about the generation of wanted reaction. Learning in an ML is chiefly ordered into two classes as [16].

- Supervised learning
- Unsupervised learning

Supervised Learning

Regulated learning is two stage forms, in the initial step: a model is fabricated depicting a foreordained arrangement of information classes or ideas. The model developed by investigating database tuples portrayed by traits. Each tuple is expected to have a place with a predefined class, as dictated by one of the qualities, called to have a place with a reclassified class, as controlled by one of the traits called the class name characteristic. The information tuple are dissected to fabricate the model all things considered from the preparation dataset [17].

Unsupervised learning

It is the kind of learning in which the class mark of each preparation test isn't knows, and the number or set of classes to be scholarly may not be known ahead of time. The prerequisite for having a named reaction variable in preparing information from the administered learning system may not be fulfilled in a few circumstances.

Data mining field is a highly efficient techniques like association rule learning. Data mining performs the interesting machine-learning algorithms like inductive-rule learning with the construction of decision trees to development of large databases process. Data mining techniques are employed in large interesting organizations and data investigations. Many data mining approaches use classification related methods for identification of useful information from continuous data streams.

Nearest Neighbors Algorithm

The Nearest Neighbor (NN) rule differentiates the classification of unknown data point because of closest neighbor whose class is known. The nearest neighbor is calculated based on estimation of k that represents how many nearest neighbors are taken to characterize the data point class. It utilizes more than one closest neighbor to find out the class where the given data point belong termed as KNN. The data samples are required in memory at run time called as memory-based technique. The training points are allocated weights based on their distances from the sample data point.

However, the computational complexity and memory requirements remained key issue. For addressing the memory utilization problem, size of data gets minimized. The repeated patterns without additional data are removed from the training data set [18].

Naive Bayes Classifier

Naive Bayes Classifier technique is functioned based on Bayesian theorem. The designed technique is used when dimensionality of input is high. Bayesian Classifier is used for computing the possible output depending on the input. It is feasible to add new raw data at runtime. A Naive Bayes classifier represents presence (or absence) of a feature (attribute) of class that is unrelated to presence (or absence) of any other feature when class variable is known. Naïve Bayesian Classification Algorithm was introduced by Shinde S.B and Amrit Priyadarshi (2015) that denotes statistical method and supervised learning method for classification. Naive Bayesian Algorithm is used to predict the heart disease. Raw hospital dataset is employed. After that, the data gets preprocessed and transformed. Finally by using the designed data mining algorithm, heart disease was predicted and accuracy was computed.

Support Vector Machine

SVM are used in many applications like medical, military for classification purpose. SVM are employed for classification, regression or ranking function. SVM depends on statistical learning theory and structural risk minimization principal. SVM determines the location of decision boundaries called hyper plane for optimal separation of classes as described in figure 3. Margin maximization through creating largest distance between separating hyper plane and instances on either side are employed to minimize upper bound on expected generalization error. Classification accuracy of SVM not depends on dimension of classified entities. The data analysis in SVM is based on convex quadratic programming. It is expensive as quadratic programming methods need large matrix operations and time consuming numerical computations.

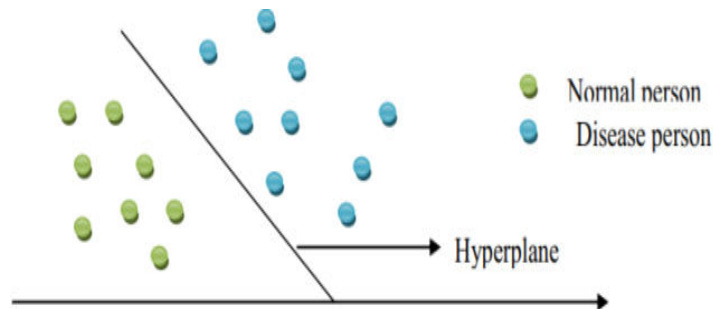


Fig. 3: Support Vector Classification

IV. SIMULATION PARAMETER

Dataset was separated into two datasets (70%/30%, preparing/testing) to keep away from any predisposition in preparing and testing. Of the information, 70% was utilized to prepare the ML model, and the excess 30% was utilized for testing the presentation of the proposed movement arrangement framework. The articulations to compute accuracy and review are given in Equations (2) and (3).

Accuracy gives a proportion of how precise your model is in anticipating the real up-sides out of the absolute up-sides anticipated by your framework. Review gives the quantity of real up-sides caught by our model by grouping these as obvious positive. F-measure can give a harmony among accuracy and review, and it is liked over precision where information is uneven.

Accordingly, F-measure was used in this review as a presentation metric to give a decent and fair measure utilizing the equation.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \times 100$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \times 100$$

$$\text{F - measure} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \times 100$$

Where,

TP—True Positive, FP—False Positive, FN—False Negative

V. CONCLUSION

During the diabetes and heart disease, numerous patients are expected to be diagnosed, treated, and monitored, thus bringing a huge burden to medical organizations. Adopting additional automated services can reduce the workload for doctors, overcrowding, and mortality rate. Moreover, in smart hospitals, ML approaches can serve as clinical decision support systems for handling issues related to the diabetes and heart disease. To the best of our knowledge, this study is the first to propose an ML model based on laboratory findings in smart hospital environments. Different ML approaches were implemented, aiming to improve the accuracy of diagnosis for diabetes and heart disease.

REFERENCES

- [1] Romany Fouad Mansour, Adnen El Amraoui, Issam Nouaouri, Vicente García Díaz, Deepak Gupta, and Sachin Kumar, "Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems", *IEEE Access* 2021.
- [2] 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. 603610, Feb. 2021,
- [3] A. A. Mutlag, M. K. A. Ghani, M. A. Mohammed, M. S. Maashi, O. Mohd, S. A. Mostafa, K. H. Abdulkareem, G. Marques, and I. de la Torre Díez, "MAFC: Multi-agent fog computing model for healthcare critical tasks management," *Sensors*, vol. 20, no. 7, p. 1853, Mar. 2020.
- [4] M. S. Hossain and G. Muhammad, "Deep learning based pathology detection for smart connected healthcare," *IEEE Netw.*, vol. 34, no. 6, pp. 120125, Nov. 2020.
- [5] M. S. Hossain G. Muhammad and A. Alamri "Smart Healthcare Monitoring: A Voice Pathology Detection Paradigm for Smart Cities" *Multimedia Systems* vol. 25 no. 5 pp. 565-75 Oct. 2019
- [6] V Krishnapraseda, M S Geetha Devasena, V Venkatesh and A Kousalya, "Predictive Analytics on Diabetes Data using Machine Learning Techniques", 7th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 458-463, IEEE 2021
- [7] Valasapalli Mounika, Devi Sree Neeli, Gorla Suma Sree, Parimi Mourya and Modala Aravind Babu, "Prediction of Type-2 Diabetes using Machine Learning Algorithms", International Conference on Artificial Intelligence and Smart Systems (ICAIS), pp. 167-173, IEEE 2021
- [8] I.K. Mujawar, B.T. Jadhav, V.B. Waghmare and R.Y. Patil, "Development of Diabetes Diagnosis System with Artificial Neural Network and Open Source Environment", International Conference on Emerging Smart Computing and Informatics (ESCI), pp. 778-784, IEEE 2021
- [9] Cecilia Saint-Pierre;Florencia Prieto;Valeria Herskovic;Marcos Sepúlveda, "Team Collaboration Networks and Multidisciplinarity in Diabetes Care: Implications for Patient Outcomes", *IEEE Journal of Biomedical and Health Informatics*, Vol. 14(1), pp. 319-329, 2020.
- [10] M. S. Hossain and G. Muhammad, "Emotion-aware connected healthcare big data towards 5G," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 23992406, Aug. 2018.
- [11] M. Pham, Y. Mengistu, H. Do, and W. Sheng, "Delivering home healthcare through a cloud-based smart home environment (CoSHE)," *Future Gener. Comput. Syst.*, vol. 81, pp. 129140, Apr. 2018.
- [12] A. Kaur and A. Jasuja, "Health monitoring based on IoT using raspberry PI," in *Proc. Int. Conf. Comput., Commun. Autom. (ICCCA)*, Greater Noida, India, May 2017, pp. 13351340.
- [13] U. Satija, B. Ramkumar, and M. Sabarimalai Manikandan, "Realtime signal quality-aware ECG telemetry system for IoT-based health care monitoring," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 815823, Jun. 2017.
- [14] O. S. Alwan and K. Prahald Rao, "Dedicated real-time monitoring system for health care using ZigBee," *Healthcare Technol. Lett.*, vol. 4, no. 4, pp. 142144, Aug. 2017.
- [15] P. Kakria N.K. Tripathi and P. Kitipawang "A Real-Time Health Monitoring System for Remote Cardiac Patients Using Smartphone and Wearable Sensors" *I. J. Telemedicine and Applications* vol. 2015 2015.
- [16] G. Villarrubia, J. Bajo, J. De Paz, and J. Corchado, "Monitoring and detection platform to prevent anomalous situations in home care," *Sensors*, vol. 14, no. 6, pp. 99009921, Jun. 2014.
- [17] M.A. Alsheikh et al. "Machine Learning in Wireless Sensor Networks: Algorithms Strategies and Applications" *IEEE Commun. Surveys Tutorials* vol. 16 no. 4 pp. 1996-2018 Fourth Quarter 2014.
- [18] F. F. Gong, X. Z. Sun, J. Lin, and X. D. Gu, "Primary exploration in establishment of China's intelligent medical treatment," (in Chinese), *Mod. Hos Manag.*, vol. 11, no. 2, pp. 28_29, 2013.
- [19] A. Krizhevsky I. Sutskever and G. E. Hinton "ImageNet Classification with Deep Convolutional Neural Networks" *Proc. 25th Int'l. Conf. Neural Info. Processing Systems* pp. 1097-1105 2012.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details