JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Smart Healthcare Prediction Using Machine Learning

Prof. Shalu Saraswat[1], Shweta Gabhane[2], Alisha Pawar[3],Suhas Pingat[4], Shreyas Patil[5]

[1,2,3,4,5,] Department of Information Technology Engineering, Pune District Education Association's College of Engineering, Pune, India, 412307

Abstract:

Depending on predictive modelling, the "Smart Health Prediction Using Machine Learning" system forecasts the disease of patients or users based on the symptoms that the user inputs to the system. There are three ways to sign in to the application: user/patient, doctor, and admin. The tool evaluates the user's or patient's symptoms as input and outputs the likelihood of the disease based on the prediction made using the algorithm. The Nave Bayes Classifier is used to make intelligent health predictions. By taking into account all of its features that were trained during the training phase, the Nave Bayes Classifier calculates the disease % probability. An accurate interpretation of disease data aids in early disease prediction for patients and users and gives them a clear understanding of the condition.

Keywords: Machine Learning, Naïve Bayes, Prediction Analysis, Symptoms.

1.Introduction:

A generative technique for creating predictive modelling using specific cases is machine learning. It's a division of AI that advances the theory that computer programmes can see patterns in data, learn from it, and make judgments with little help from humans. A programming approach called machine learning employs test data or previously gathered data to accurately optimise results.

The machine learning algorithm has two stages: planning and research. The user's or patient's signs and symptom logs are utilised to forecast the sickness and health rate. Based on user/patient experience, machine learning technology provides a robust application arena in the medical industry to address health disease prediction challenges. We employ machine learning to monitor all symptoms and illnesses. Machine learning technology aids in the quick analysis of predictive models. With the help of technology, the user/patient can decide on their own whether to consult a doctor about certain symptoms or not, leading to better patient health services. The large amount of obtained data is analysed using Gaussian Naïve Bayes classifier. We also showed how symptom data storage and data classification can help the administrative, clinical, academic, and educational elements of disease prediction from symptoms for each sub-field of disease predictions. There are numerous data collection methods.

To predict smart health, there are various resources available. However, studies have focused on serious disease, and a degree of risk has been determined. These techniques aren't typically employed for disease prediction in general disease.

By using a perfectly suited Machine Learning Algorithm technique which analyse patient symptoms and aids in the diagnosis of various illnesses. It saves time and resources.

2.Literature Survey:

There have been numerous studies done related to predicting the disease using different machine learning techniques and algorithms which can be used by medical institutions.

Naveenkumar and his colleague has proposed a system of health prediction using machine learning algorithms. In prediction of system he used different algorithms like Naive Bayes, K-Nearest Neighbor, and Decision Tree. This proposed system had an accuracy of 94%[1].Aditi Gavhane and her colleague suggested prediction for heart disease that utilizes Machine Learning. Multi-Layer Perceptron model is used in this system. This system predicts heart disease based on basic symptoms like age, sex, pulse rate, etc. The accuracy of this suggested system is 91% [2].Gupta A and his colleague proposed system for heart prediction which makes use of naive bayes algorithm.The accuracy of naive bayes algorithm is 97%.This system predicts heart diseases based on basic sysmtoms chest pain, pain of discomfort in jaw ,neck, shortness of breathing [3].D Dahiwade and his colleague designed a model for prediction of the disease using approaches of machine learning and used techniques like KNN and CNN. This paper suggests disease prediction i.e. based on patient's symptoms. The accuracy of KNN is 95% and the accuracy of CNN is 98%[4].N. Shabaz Ali and his colleague designed a model for prediction of diseases using data mining technique. This paper is made on how the data mining techniques are used along with the machine learning to predict the diseases based on the user symptoms[5].Shubham Salunke and his colleague has designed system for identifying the diseases that patients are suffering from on the basis of naive bayes algorithm[6].H. Pandey and his colleague make use of IoT and machine learning for healthcare monitoring. The Internet of Things (IoT) has enabled the invention of smart health monitoring systems. These health monitoring systems can track a person's mental and physical wellness on the basis of stress level, anxiety level and hypertension of particular patient[7].Monika Gandhi and her colleague has proposed system for predicting heart diseases using data mining techniques. . In this paper, data mining methods namely, Naive Bayes, Neural network, Decision tree algorithm are analyzed on medical data sets using algorithms[8].S. Ananth and his colleague has proposed system for health prediction using IoT which has enabled invention in healthcare. The accuracy of the system is about 82%[9].S.S and his colleague proposed model for health and disease prognosis system. For prediction of diseases, different machine learning algorithms such as Random Forest, Naive Bayes, Logistic Regression, Support Vector Machine, K-Nearest Neighbours , Decision Tree and Gradient Boosting are compared to predict in an efficacious manner with better accuracy. The best accuracy model is saved for disease prediction. This system is especially used for early prediction of disease[10].Rudra A. and his colleague has proposed system for multiple disease prediction. This system has additional appearance of consulting drugs and medication of disease expected which is the biggest drawback of the model. The accuracy of the system is about 85%[11].M. Asia and his colleague tried to find a scalable solution that can predict different disease utilizing Random Forest Algorithm. This system presents a comparison against Naïve-Bayes classifier but Random Forest gives more accurate results with accuracy 98%[12].Sneha R. and his colleague has designed system named disease prediction based of classification algorithm. The classification algorithms such as Decision Tree, Support Vector Machine(SVM), K Nearest Neighbour(KNN),Random Forest, Logistic Regression, Naive Bayes are used for building this disease prediction model. The Naive Bayes provides highest accuracy 97% and hence used for prediction of the diseases[13].Farooqui and his colleague has designed health prediction system using support vector machine and multilinear regression. The result generated by proposed system has accuracy up to 87%[14].N. Kosarkar and his colleague has proposed system for health prediction using different algorithms of machine learning like random forest, support vector machine and logistic regression. The accuracy of the proposed model is 82%[15].

3. Project Analysis :

3.1.Existing Method :

The model forecasts chronic diseases for a particular region and population. Only specific diseases are included in disease prediction. In this approach, disease risk is predicted using big data and the convolutional neural networks algorithm. For S-type data, the method uses machine learning algorithms like K-nearest neighbours and Decision Tree. For some diseases, the system has an accuracy up to 94.8 percent rate. In the prior research, we decomposed machine learning techniques to forecast chronic disease outbreaks in populations that are prone

to illness. Using actual hospital data from a few particular regions/areas, we are testing improved prediction models. We offer a brand-new multimodal illness risk prediction technique for Convolutional Neural Networks that uses structured and unstructured patient/user input.

3.2. Proposed Method:

If a person is diagnosed with some sort of disease, then he needs to consult a doctor/physician which is both time consuming and expensive. Sometimes it is also very difficult for user to visit the hospital so, it is difficult to guess the disease by its own. But if the above procedure is done using electronic software application, then it would be very effective for patient and it will save time and resource both for the user/patient. The process would run very smoothly. Smart health care prediction is a web-based program that predicts user's illness based on the symptoms of user/patient that he feels. The Data set required for Smart Health Prediction Framework have been compiled from various health-related websites. The consumer would be able to predict the outbreak of the illness on the basis of symptoms present in the web-application. The aim of this project is to create a web platform that can predict disease events based on a range of symptoms. Users can choose from a range of symptoms and find diseases with probabilistic estimates and conditions.

Efficiency comparison TABLE

Diseases	NB	LR	K*	DT
Breast cancer wise	97.30	92.90	95.70	94.90
Breast Cancer	72.70	67.70	73.70	74.20
Dermatomgaly	97.40	96.80	94.50	94.10
Echo Chambers	95.70	94.50	89.30	96.40
Liveries	54.80	68.70	66.80	65.80
Pimaricin Diabetes	75.70	77.40	70.10	74.40
Heamatidroses	75.30	74.40	73.70	72.10
Heart-c	83.30	83.70	75.10	77.10
Heart-statlog	84.80	84.00	73.80	75.50
Heart-b	83.90	84.20	77.80	80.20
Hepatitis	83.80	83.80	80.10	79.20
Lung Cancer	53.20	47.20	41.60	40.80
Lymphs	84.90	78.40	83.10	78.20
Postooerasis	68.10	61.10	61.60	69.70
Tumor	49.70	41.60	38.00	41.30
Success Ratio	**8/15**	**5/15**	**0/15**	**2/15**

NB – Naïve Bayes, LR – Linear Regression, K*- Kth Nearest & DT – Decision Tree

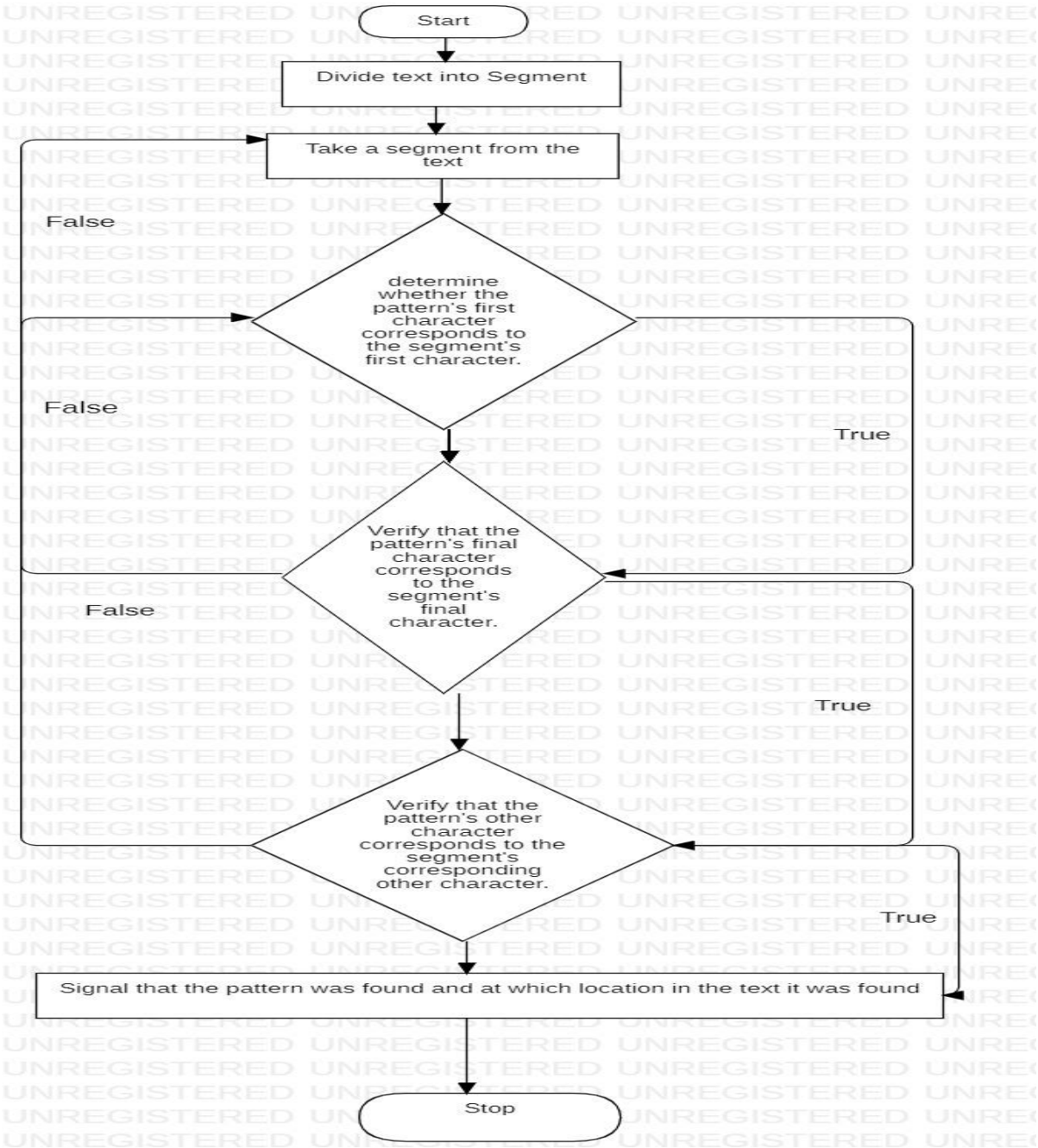
In the above table we have compared different methods in terms of accuracy and it resulted that the Naïve Bayes algorithm has higher accuracy and better timing comparatively. We used naïve bayes algorithm identify patient data because medical data are increasing at an exponential rate, requiring the processing of existing data in order to predict exact disease based on symptoms. By having the input as a user/patient record, we were able to get accurate general disease risk prediction as an output that helped us understand the degree of disease risk prediction. Because of this method, the disease prediction and risk prediction could be achieved very fast and easily with low cost.

3.Algorithm and Architecture

3.1 Naïve Bayes Algorithm

The Naïve Bayes algorithm is a simple dynamic method for creating models for assigning class labels to problem instances to find a mapping to object. Class labels are chosen from a finite set of choices. It is a family of algorithms based on a general principle, not a particular algorithm. According to this principle, the value of each function of all Naive Bayes Classifiers is independent of the value of other features. for example, if the

fruit is orange, round, and around 10cm-15cm in diameter, we might call it an orange. The Naïve Bayes algorithm also takes into account each feature to determine if the fruit is an orange. There are a n-variety of probability models, but for some of them, the Naïve Bayes algorithm performs best in supervised learning model.



3.2 Random Forest Algorithm:

Random forest is based on bagging algorithm that uses a kind of ensemble learning techniques. In random forest algorithm multiple decision trees are used and it also combines the output of the tree. In this way it can avoid overfitting problem also improve the accuracy. A forest is made up of many different types of trees, and the more trees there are, the more robust the forest will be.

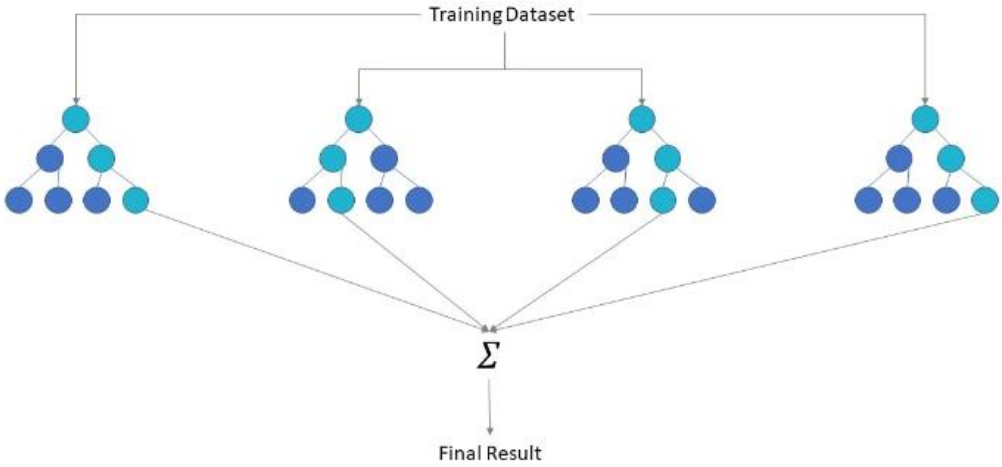


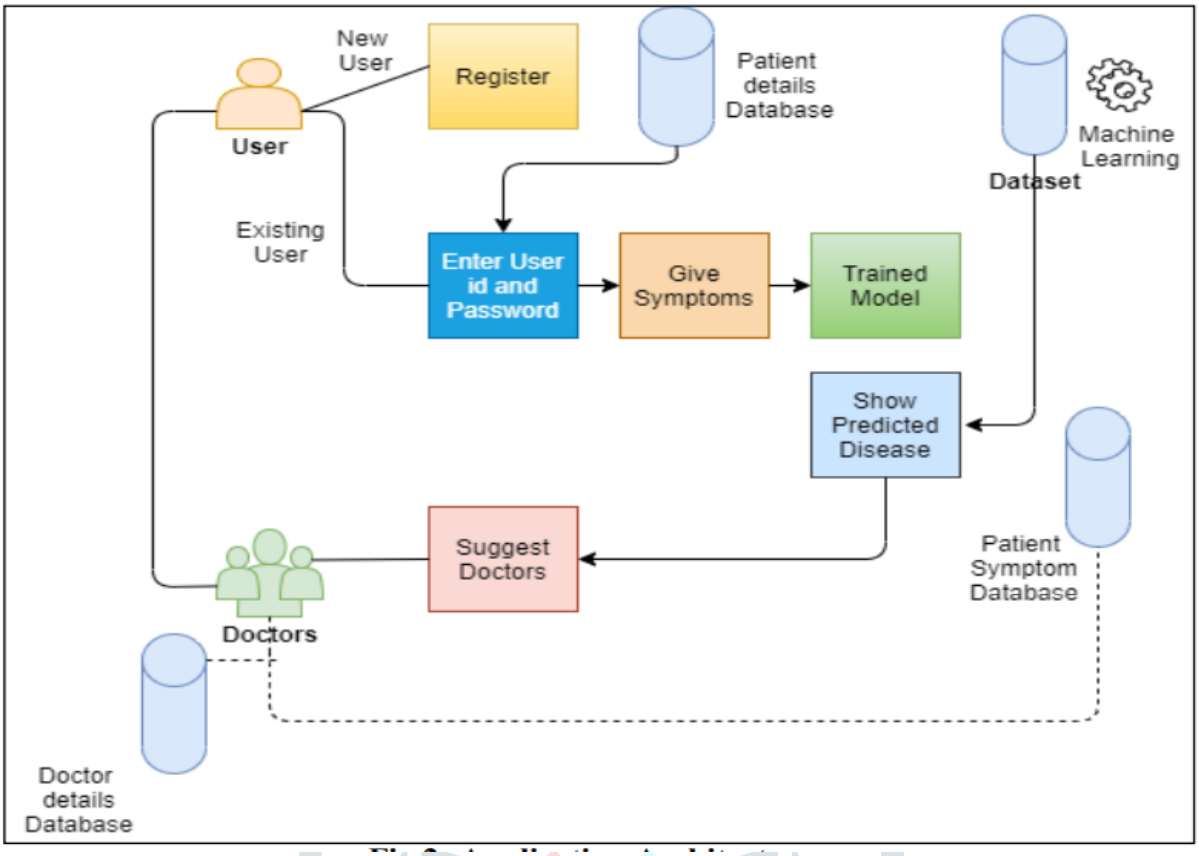
Diagram of Random Forest Classifier

3.3 Support Vector Classifier- SVC:

- One of the most well-liked supervised learning algorithms, Support Vector Machine, or SVM, is used to solve Classification and Regression problems. However, it is largely employed in Machine Learning Classification issues.

- The SVM algorithm's objective is to establish the best line or decision boundary that can divide n-dimensional space into classes, allowing us to quickly classify fresh data points in the future. A hyperplane is the name given to this optimal decision boundary.

Architecture:



4.Conclusion:

It has been a great pleasure for us to work on this exciting and challenging project. This project proved good for us as it provided practical knowledge of not only programming in Python and Sqlite web based application. It also provides knowledge about the latest technology used in developing web enabled application and client server technology that will be great demand in future. This will provide better opportunities and guidance in future in developing projects independently.

5.References:

1) Naveenkumar S1 , Kirubhakaran R2 , Jeeva G3 , Shobana M4 , Sangeetha K5," Smart Health Prediction Using Machine Learning: A survey",3 march 2021.
2) Aditi Gavhane, Gouthami Kokkula, Isha Pandya and Kailas Devadkar, "Prediction of Heart Disease Using Machine Learning" IEEE Xplore ISBN: 978-1-5386-0965-1, pp. 1275-1278, 2018.
3) Gupta A., Kumar L., Jain R., Nagrath P. (2020) Heart Disease Prediction Using Classification (Naive Bayes). In: Singh P., Pawłowski W., Tanwar S., Kumar N., Rodrigues J., Obaidat M. (eds) Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019). Lecture Notes in Networks and Systems, vol 121. Springer, Singapore.
4) D. Dahiwade, G. Patle and E. Meshram, "Designing Disease Prediction Model Using Machine Learning Approach," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 1211-1215
5) N. Shabaz Ali, G. Divya,"Prediction of Diseases in Smart Health CareSystem using Machine Learning" 5 January 2020.
6) Shubham Salunke, ShubhamRajiwade, Deepak Yadav, S.K.Sabnis, "smarthealth prediction system using machinelearning",March 2020.

7) H. Pandey and S. Prabha, "Smart Health Monitoring System using IOT and Machine Learning Techniques," 2020 Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2020, pp. 1-4.
8) Monika Gandhi and Shailendra Narayan Singh "Predictions in heart disease using techniques of data mining" 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE) pp. 520-525 2015.
9) S. Ananth, P. Sathya, P. Madhan Mohan," Smart Health Monitoring System through IOT",In 2019 International Conference on Communication and Signal Processing (ICCSP),IEEE April 2019.
10) S. S, V. S and A. R, "Health and Disease Prognosis System using Machine Learning," 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT), Visakhapatnam, India, 2021, pp. 192-196.
11) Rudra A. Godse Smita S. Gunjal Karan A. Jagtap Neha Mahamuni and Suchita Wankhade "Multiple DiseasePrediction Using Different Machine Learning Algorithms Comparatively" International Journal of Advanced Research in Computer and Communication Engineering vol. 8 no. 12 Dec 2019.
12) R. Delshi Howsalya Devi P. Sreevalli Keerthana Prathyusha and M. Asia "Prediction of Diseases Using Random Forest Classification Algorithm" Zeichen Journal vol. 6 no. 5 2020.
13) Sneha R. Monisha S. Jahnavi C. and S. Nandini "Disease Prediction Based On Symptoms Using Classification Algorithm" Journal of Xi'an University of Architecture & Technology vol. 12 no. 4 2020.
14) Farooqui, Md. Ehtisham and Ahmad, Dr. Jameel, Disease Prediction System using Support Vector Machine and Multilinear Regression (August 13, 2020). International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, Volume, 8, Issue, 4, July, 2020.
15) N. Kosarkar, P. Basuri, P. Karamore, P. Gawali, P. Badole and P. Jumle, "Disease Prediction using Machine Learning," 2022 10th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-22), Nagpur, India, 2022, pp. 1-4, doi: 10.1109/ICETET-SIP-2254415.2022.9791739.