



A Machine Learning Model for Prediction of Diabetes and Hypertension Using Ensemble Learning Approach

¹Dr. Smitha Shekar B, ²Dr. Harish G, ³Pooja A

¹Associate Professor, ²Associate Professor, ³M.Tech Student

^{1,2,3}Department of Computer Science & Engineering

^{1,2,3}Dr. Ambedkar Institute of Technology, Bengaluru, India

Mail id: ¹smithashekarb.cs@drait.edu.in, ²c_harishg.cs@drait.edu.in, ³poojaankaiah@gmail.com

Abstract: In today's world, diabetes and hypertension are the most common and non-communicable growing diseases in several countries. Researchers are working on early prediction and prevention of diabetes and hypertension. Several machine learning techniques are being utilized to predict both diseases. Most of the people suffering from these diseases, have higher complications related to heart disease, stroke, kidney failure, neuropathy, retinopathy, foot damage and pregnancy related issues. In this paper we have worked on diabetes and hypertension and proposed a disease prediction model to predict and classify the disease based on individual's risk factor. Various Machine Learning algorithms like Logistic Regression, Support Vector Machine, Naïve Bayes, Multi-Layer Perceptron, k-nearest neighbours, Decision Tree, AdaBoost, Gradient Boosting, Random Forest are used for prediction of diabetes and hypertension. The results obtained after a comparative study, showed that Logistic Regression obtained high accuracy of 83.16% on diabetes dataset and Decision Tree and Gradient Boosting obtained most effective accuracy of 100% on hypertension dataset. In the proposed work, a web application is developed to provide a practical application of the disease prediction model by gathering risk factor data and a health care Chatbot provides real time assistance to patients.

Index Terms: Diabetes, Hypertension, Prediction, Machine Learning Algorithm, Web Application, Chatbot.

I. INTRODUCTION

In this paper, machine learning method have been used to determine the diabetes and hypertension type based on individual's risk factor. Machine learning is a sub-branch of "Artificial Intelligence" (AI). It is fast growing technology in various field which enables computers to learn automatically from the previous data. Machine learning uses several algorithm for building prediction model, currently, this technique increasingly used in various field such as healthcare,

image processing, speech recognition, object detection, robotics, data mining, video games, text analysis finance and many more.

Nowadays, diabetes is a common critical disease based on age, obesity, lack of exercise, hereditary, living style, high blood pressure, bad diet, etc. can causes diabetes. Diabetes is a condition in which glucose level in your blood are too high. When we eat some of food is turned into sugar is called glucose, glucose travels to all of your body cells through bloodstream [1]. Diabetes is caused because of excessive glucose level in the body, it caused because of improper working of beta cells in pancreas and also it effect on different part of the body. Where person is bear from the extensive level of blood glucose in the body, which is either insulin production is not enough or body cells do not respond adequately to insulin. Common types of diabetes are

Type1: Is an insulin dependent diabetes also called “juvenile diabetes”. In type 1 diabetes cells from body’s immune system destroy the insulin producing beta cells of pancreas without symptoms or pain over a period of months or years, the beta cells are destroy as a result the body does not have sufficient insulin, normally insulin assist the cells of body remove glucose out of the body, without insulin glucose develop in the blood this can injure blood vessels are also causes, blindness, kidney failure and heart disease [1]. Type 1 diabetes people inject insulin several times a daily.

Type2: Is non-insulin dependent diabetes. In type 2 diabetes the pancreas is incapable to produce larger amount of insulin needed to trigger the resistant cells to take in glucose from the bloodstream but cells become insulin resistance and ignore the message to absorb glucose [2]. In addition to type 2 diabetes commonly occur in people who are elder or in those who are obese.

Gestational diabetes: It is also one type of diabetes that is diagnosed during pregnancy. During pregnancy, a temporary organ evolve attached to the mother and the fetus called placenta. The placenta provide nutrients and oxygen to the fetus, and also generates number of hormones that work to sustain pregnancy. Some of the hormones damage the action of insulin, making it ineffectual. This insulin counteracting impact generally starts at about 20- 24 weeks of pregnancy, the effect increases as the placenta develop larger and become most important in the last few months. Typically, the pancreas is able to modify by producing higher insulin but in certain cases the amount of placental hormones may become too overwhelming for the pancreas to remunerate and gestational diabetes result [2] [3]. And gestational diabetes lead to some complication of large baby, low blood sugar in newborn baby, high blood pressure in mother and in future both mother and child having diabetes.

The diabetes common signs and symptoms are frequent urination, being thirsty and intensify hungry, tiredness, blurred vision, unexpected weight loss (generally Type 1), obesity (generally Type 2), delayed healing and itching of your skin. And some of diabetes complication are heart disease, stroke, nerve damage and cardiovascular disease, damage, kidney damage, foot damage and many more.

Diabetes cannot be cured but it can be controlled by taking prescribed medication, eating and drinking well planned diet, exercising regularly which helps your body control your glucose level, manage your blood pressure and cholesterol level.

Diagnosis level of diabetes

- a) Hemoglobin A1C (HbA1C): HbA1C test is simple blood test is also called Glycated hemoglobin test to diagnose whether a patient have diabetes or prediabetes [4]. This test evaluate the glucose level in the blood over the last 2 to 3 months. Hemoglobin is a protein found in red blood cells and gives blood its “red colour”, hemoglobin carries oxygen throughout the body. Hemoglobin A1C levels may be checked every 2 to 3 months until a patient with diabetes reaches the treatment goals.
- b) Fasting Plasma Glucose (FPG): Fasting glucose test is one of the blood test is conducted to measure blood sugar level after fasted at least 8 hours.
- c) Oral Glucose Tolerance Test (OGTT): After FPG test patient drink sugary liquid then take the reading of blood sugar level again in 2 hour.

Table 1: Diagnosis levels of diabetes-ranges

Indication	HbA1C	FPG (mg/dL)	OGTT (mg/dL)
Normal	Less than 5.7%	99 or below (mg/dl)	139 or below (mg/dl)
Prediabetes	5.7% to 6.4%	100 to 125 (mg/dl)	140 to 199 (mg/dl)
Diabetes	6.5% or above	126 or above (mg/dl)	200 or above (mg/dl)

Hypertension: Hypertension is most common disease in the world and most of people are suffering from this disease. Hypertension occurs when blood pressure is high, it is essential to diagnose and manage at an initial stage. Hypertension is considered as one of major risk factor for heart disease and it is common among patient with diabetes. Blood pressure is a measurements of force of the blood as it pumps through arteries is too high. Arteries are the blood vessels that carry blood away from heart to supply tissues with oxygen and nutrients. Blood pressure is measured with an equipment called a blood pressure cuff or sphygmomanometer. Blood pressure reading it has two number namely, systolic blood pressure is a measurements of pressure in the arteries wall as our heart pumps blood, and diastolic blood pressure is measurements of pressure in the arteries wall when our heart is rest [9]. Hypertension is most common in men at the age of 64 and in women hypertension occur after the age of 65.

Blood pressure measurements fall into many categories: Blood pressure is depicted as Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP). If SBP is greater than or equal to 140 and DBP is greater than or equal to 90 is considered as “Stage-1 hypertension”. And SBP is above or equal to 160 and DBP is above or equal to 100 is considered as “Stage-2 hypertension”. Pre-hypertension would be considered as SBP and DBP is above or equal to 120 over 80 as shown in Table 2. Pre-hypertension is an earlier statement used to describe the patient with high risk of hypertension, pre-hypertension often leads to high blood pressure. Normal blood pressure would be considered as SBP and DBP is below 120 over 80 is considered as person having a normal blood pressure.

Table 2: Blood pressure value with respect to SBP and DBP

Blood pressure category	SBP (mmHg)	DBP (mmHg)
Normal	<120	and <80
Prehypertension	120-139	or 80-89
Stage-1 hypertension	140-159	or 90-99
Stage-2 hypertension	≥ 160	or ≥ 100

Peoples are obese, high cholesterol levels, water salt ratio, hormone levels, too much salt, hereditary, obesity and long term stress are also at higher risk for developing hypertension. Smoking cigarettes, taking alcohol are also causes for developing high blood pressure and diabetes. Most people with hypertension do not notice symptoms at an early. Symptoms of high blood pressure are early morning headaches, nose bleeds, visual disturbances, More severe forms may exhibit fatigue, confusion, nausea, anxiety, vomiting and, chest pain.

If not treated, hypertension can causes persistent chest pain, heart attack, heart failure, and an irregular heartbeat, which can lead to an unexpected death. Hypertension can lead to heart failure, kidney failure, stroke, sexual dysfunction, angina, microvascular disease, peripheral artery disease. Hypertension cannot be cured but it can be managed by reducing and managing mental stress, living a healthy life style, medication, increasing physical activity.

Diagnosis of hypertension

- a) Ambulatory Blood Pressure (BP) Monitoring (ABPM) this device monitor the patient blood pressure for 24-hour and provide a clear picture of blood pressure changes over an average day and night.
- b) Lab test: A urinalysis test of urine protein levels and detect kidney disease, blood test for measuring various hormone level and cholesterol test.
- c) Electrocardiogram (ECG or EKG): This test monitor the heart's electrical activity.
- d) Echocardiogram: An electrocardiogram uses sound waves to produce image of the heart.

II. LITERATURE SURVEY

Diabetes and hypertension are most common and non-communicable disease in the world. Now a days many researcher have worked on predicting diabetes and hypertension disease using several machine learning techniques. Among them, in author [5], Pradeep Kandhasamy et al. used several machine learning techniques J48 decision tree, K-Nearest neighbour, Random forest and Support vector machine to classify patients with diabetes or not. The four models used for predicting diabetes mellitus under two different cases: One is noisy data after comparison using 4 different classifier with noisy data, the decision tree J48 classifier achieves higher accuracy then the other 3 classifiers. Another one is without noisy dataset in this dataset KNN and random forest algorithm gives higher accuracy compared to other classifiers in terms of Accuracy, Sensitivity and Specificity.

In [6], author Ahmed Kareem et al. applied machine learning technique to design a prediction algorithm to classify the medical diabetes data and improve the accuracy in diabetes prediction and used three classifiers to perform a classification namely, Random Forest, Radial Basis Function Network (RBFN) and Multi-Layer Perceptron (MLP) the

best classifier to give the approximate result comparing to clinical results. Finally, RBFN and MLP has the highest specificity holds best for the analysis of diabetes data and using tenfold Cross-validation evaluation techniques the RBFN gave the best accuracy.

Hani Bani-Salameh et al., in [7] had applied Artificial Neural Network (ANNs) technique to predict diabetes and hypertension, for performance support vector machine, Multi-Layer Perceptron (MLP), and KNN algorithm used and developed a disease diagnosis tool to predict the “Diabetes and Hypertension” depend on individual risk factor. Two dataset used for prediction, the first dataset for diabetes and second for hypertension to predict the disease for both dataset multi-layer perceptron gave higher accuracy compared to other two algorithm.

Deepti Sisodia et al., in [8] applied a classification algorithm namely decision tree, SVM and naïve bayes are used and WEKA software tool used for predict the diabetes. Finally, naïve bayes algorithm achieved best accuracy result compared to other algorithm on variant performance measures like precision, accuracy, F-measure and recall and result are checked using Receiver Operating Characteristic curve.

Melin et al., designed a hybrid model in [9] based on Neural Network (NN) and fuzzy system to predict and categorize the hypertension types based upon individual risk factor and behaviour of blood pressure in a duration of 24 hour. Different technique of soft computing and three fuzzy inference system are used, first two fuzzy system used to classify heart rate level and blood pressure and third fuzzy system used to sort patient’s night profile. Finally, author concluded that neural network is best efficient technique to predict the behaviour of blood pressure of different person.

Katayama et al., in [10] designed a blood pressure measurement device that can measure blood pressure consistently using the Fiber Bragg Grating (FBG) sensor. FBG sensor was used to evaluate the pulse wave, and blood pressure was projected using both Artificial Neural Network (ANN) and Partial Least Squared Regression (PLSR) method. The result indicate that prediction accuracy using PLSR was lower because correlation coefficient was small and pulse wave shape was big and ANN method used for blood pressure prediction, ANN method obtained best result because coefficient correlation was big compared to PLSR. Finally author concluded that artificial neural network is the best technique for prediction of blood pressure compared to PLSR.

Namrata Singh et al., in [11] developed a hybrid model for diagnosing a hypertension among diabetes person. A new technique called rule based explanation method from support vector machine was added and it consider second opinion for predicting hypertension among diabetic patient. For evaluation, a new dataset related to diabetic and hypertension are applied to developed model, finally, concluded that the proposed model obtained best result compared to other hybrid model.

III. METHODOLOGY

3.1 Model diagram

The aim of this work is to analyse diabetes and hypertension type over classification techniques. This work concentrates on to predict the disease of diabetes and hypertension, hence diagnosing of disease at an early stage is very essential to save human life.

The disease dataset containing the information about individual risk factor of patients will be fed into the prediction algorithms. Then the performance of algorithm will be tested with appropriate evaluation model. After comparison of accuracy of each algorithm then develop a web application to provide the practical application of the proposed disease prediction model as shown in the Fig.1. After, live chat model developed to establish communication between doctor and patient, and also developed chatbot model to provide a real-time assistance to patients and also provide basic details about the disease.

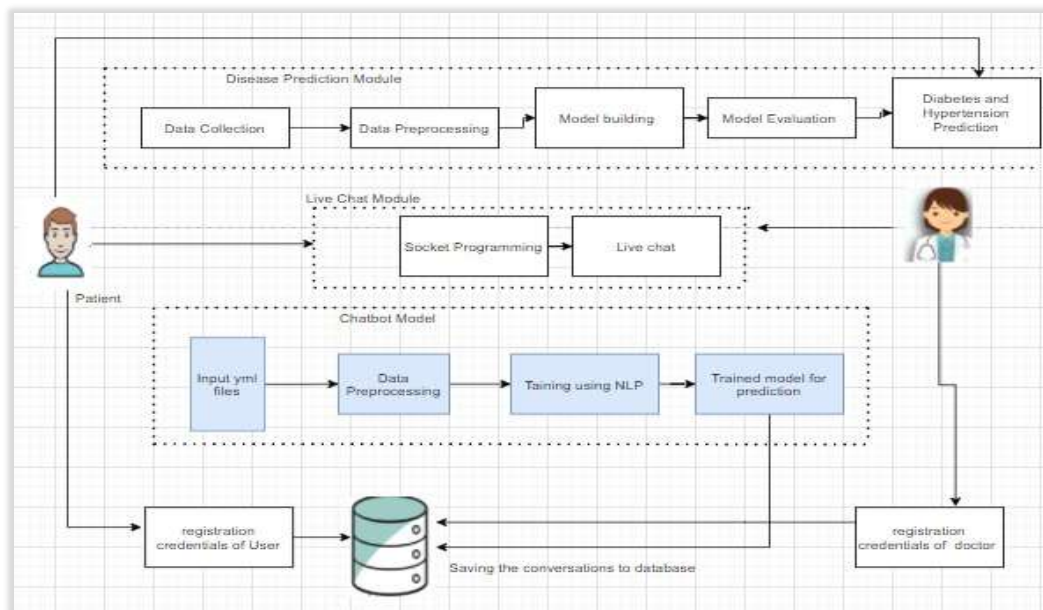


Fig.1: Block diagram of proposed model

3.2 Dataset description: For this study, we have used online dataset for analyse the disease over classification techniques, dataset download from kaggle machine learning repositories which contain a record of 768 patient with 9 attributes: pregnancies, glucose, skin thickness (mm), diastolic blood pressure (mmHg), body mass index (kg/m^2), insulin (mm U/mL), diabetes pedigree function and age (year) these features define the diabetes and outcome shows that whether person has diabetes or not [12].

The second dataset has been used hypertension dataset for prediction of “Hypertension types”. In the dataset, 8 feature define the hypertension types: sex (M/F), age (year), height (cm), weight (Kg), systolic blood pressure (mmHg), diastolic blood pressure (mmHg), heart rate (b/m), and BMI (kg/m^2) and these features define the categories of hypertension and outcome shows whether person have prehypertension, stage1 hypertension and stage2 hypertension and normal in the hypertension dataset [13].

3.3 Brief description of algorithms

- SVM:** A SVM is supervised very effective data classification and regression algorithm, the main aim of this algorithm to deal with linearly separable binary classification data, data points lying on the positive and negative hyperplane called support vector. A maximum hyper-plane that can segregate the data by their class label and at the same time avoid overfitting the data. During the training process “SVM classifier” were used to extracted

the features [14]. This algorithm applicable for diagnosis purpose and also used for image classification, text classification and face detection. SVM algorithm has two types: linear and non-linear SVM.

- b) **Decision tree:** A most popular supervised learning algorithm, used for both classification and regression analysis and it uses a divide and conquer approach for searching. Decision tree creates a binary tree for the purpose of improving prediction accuracy and easy to understand and analyse the given data [15]. Decision tree normally ability to imitate human thinking capacity while making a decision. All the output data in the form of categorical type and main advantage of decision tree is that classification can be easily understand.
- c) **Logistic Regression:** Is a supervised learning and one of the simple and most ideal algorithm used to solve a classification problem. This algorithm used for predict binary output variables using a given set of input variables [16]. The output variables outcome should be discrete, such as $y \in \{0, 1\}$. A binary output variable can have only two values such as discrete or categorical. The probability in the logistic regression is represented by the sigmoid function is also called “logistic function or the S-curve”. The sigmoid function gives an ‘S’ shaped curve. This curve has a finite limit that is Y can only be 0 or 1. The threshold value indicate the probability of 0 or 1. Logistic regression can be used in lone sanction, customer segmentation, spam filter, whether prediction and exam result prediction cases.
- d) **Random forest:** Is a supervised learning algorithm and combine multiple algorithm to obtain the better predictive result compare to other algorithm. This algorithm combines the output of randomly created several decision tree and to generate the final result. It gives high accuracy and run efficiently on large dataset, use multiple tree to reduce the risk of overfitting issues and take less time to train the data. This algorithm use in many application namely, Remote sensing, object detection and many more.
- e) **Naïve Bayes:** A simple statistical learning classification algorithm based on principle of Bayes theorem. It is power classification, high accuracy, most effectual classification algorithm suitable for large dataset and building machine learning model for fast prediction. Using Bayes theorem calculate the probability of an event this algorithm also used for spam filtration, sentimental analysis.
- f) **KNN:** The K-nearest neighbours is a simple, and easy to implement supervised learning algorithm that can be used for both classification and regression analysis. This algorithm loads all the available data and sort a new data points based on the similarity and it is a non-parametric and lazy learner algorithm because it does not learn from the training data instantly, rather it stores the data at the time of classification. It is commonly used because easy to interpretation and low calculation time.
- g) **MLP:** It is an artificial neural network that produce a set of outputs from a set of inputs. This learning algorithm is also called back propagation algorithm and multiple neuron present in different layer. MLP model is most popular network architecture used in several research field, each layer is a large connection of neuron and containing input layer. In this layer feed all the input data, hidden layer and output layer. In this model each node use activation function and back propagation method for training data and minimizes the error.
- h) **Adaboost:** Adaptive boosting is a statistical classification meta-algorithm it is was first truly successful boosting algorithm to improve the performance of any machine learning algorithm and it is developed for binary

classification. Adaboost is implemented by integrate several weak learners into single strong learner and this algorithm uses decision tree as the base learners.

- i) **Gradient boosting:** Is a greedy algorithm and can overfitting a training data fastly. It can benefit from regularization methods that correct several parts of the algorithm and usually boost the performance of the algorithm by reducing overfitting.

3.4 Web application: For creating web application of diabetes and hypertension, we used flask which is an Application Programming Interface (API) of python [17] and used for developing web application, machine learning application and end to end projects, which was developed by Armin Ronacher. Flask is a web application framework written in python and it is a third party python library used for developing web application. Flask's framework is easy to develop simple web application, because it has less code to implement and provide useful tools and features to creating simple web application. Web application developed to provide the practical application of the proposed disease prediction model and gather risk factor data and display a result

3.5 Live Chat: In Live chat the mutual interaction is done between the user and the doctor through chat and the patient takes advice from the doctor, which helps to avoid waiting hours in the crowd. The working of live chat between patient and doctor first the doctor and user have to login into the both tabs, user will check the availability of the doctor. So, if the doctor is in online he can chat with the doctor and have the session, now the user can chat with doctor by clicking on the chat button, In other side the doctor also should click on the chat button, then chat windows will be open for both doctor and patient.

3.6 Chatbot: Chatbots are interactive application that are designed to simulate human conversations [18]. Chatbots developed in many industry, mainly healthcare industry that can diagnose the disease and provide real-time assistance to patients and also provide basic details about the disease before consulting a doctor. Chatbots are software developed with ML algorithms and natural language processing, to engage in a conversation with a user.

VI. RESULT

From the below Table 3 we can conclude that logistic regression is most ideal algorithm for predicting diabetes disease, which gives an accuracy of around 83.16%. In Fig.2 shows the accuracy graph of each algorithms on diabetes data. And if people want to control diabetes, they should really keep their sugar level down with increase in age they must should follow proper diet and regular exercising. In the hypertension dataset we can conclude that decision tree is most ideal algorithm for predicting hypertension, which gives 100% accuracy on the hypertension dataset as shown in Table 4 and Fig.3 shows the accuracy graph of each algorithms on hypertension data. Accuracy can be defined using below equation, accuracy is the overall success rate of the algorithm.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Table 3: Comparison Results of Classifiers

Diabetes dataset result	
Algorithms	Accuracy
Multilayer perceptron	70.12%
K-nearest neighbor	74.02%
Adaboost	77.92%
Naïve bayes	77.92%
Decision tree	79.87%
Gradient boostings	80.51%
SVM	80.52%
Random forest	81.81%
Logistic regression	83.11%

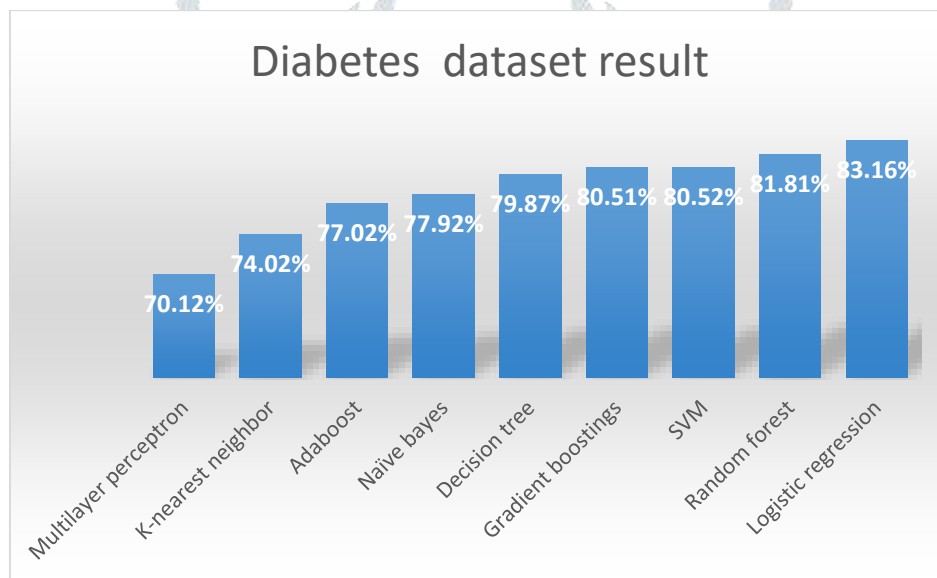


Fig.2: Accuracy graph of diabetes dataset

Table 4: Comparison Results of Classifiers

Hypertension dataset result	
Algorithms	Accuracy
AdaBoost	79.54%
K-nearest neighbor	81.81%
Multilayer perceptron	84.09%

Logistic regression	88.63%
Naïve bayes	90.90%
SVM	93.18%
Random Forest	95.45%
Gradient boosting	100%
Decision tree	100%

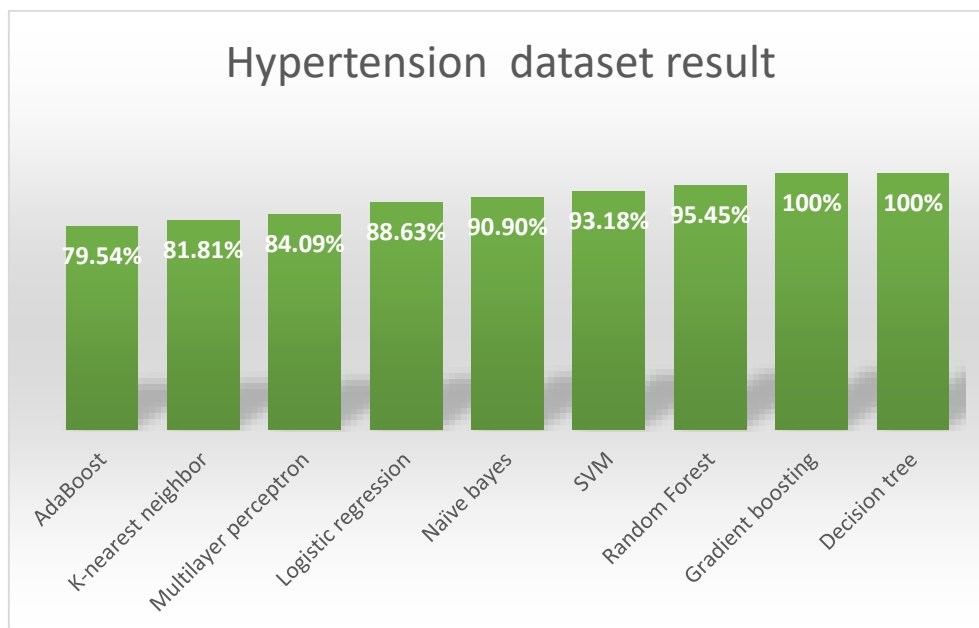


Fig.3: Accuracy graph of hypertension dataset

Finally, developed web application using flask, using this web application user interface patient or new user can able to login this application as shown in Fig.4. In Fig.5 shows the values entered on respected cell either one of the feature met the diabetes condition display a message person having diabetes or non-diabetes. And also using user interface entered blood pressure value and display a message person having pre-hypertension, stage-1 hypertension, stage-2 hypertension and normal as shown in Fig.6. In Fig.7 represent the live chat between the doctor and patient and chatbot provide a real-time assistance to patients and also provide basic details about the disease as shown in Fig.8.

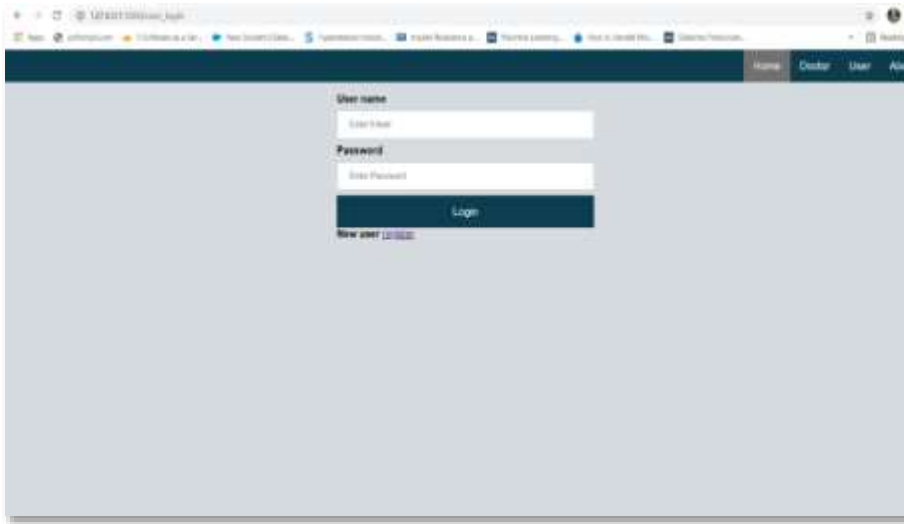


Fig.4: User login form

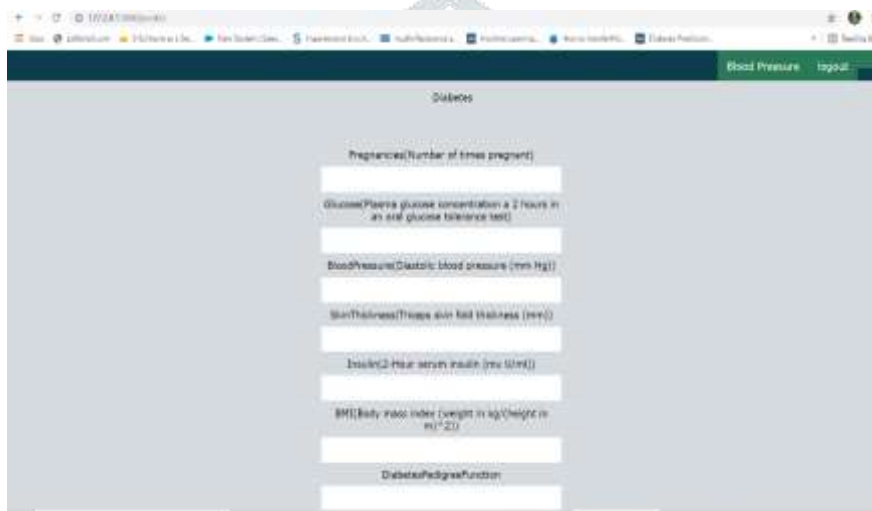


Fig.5: Diabetes prediction

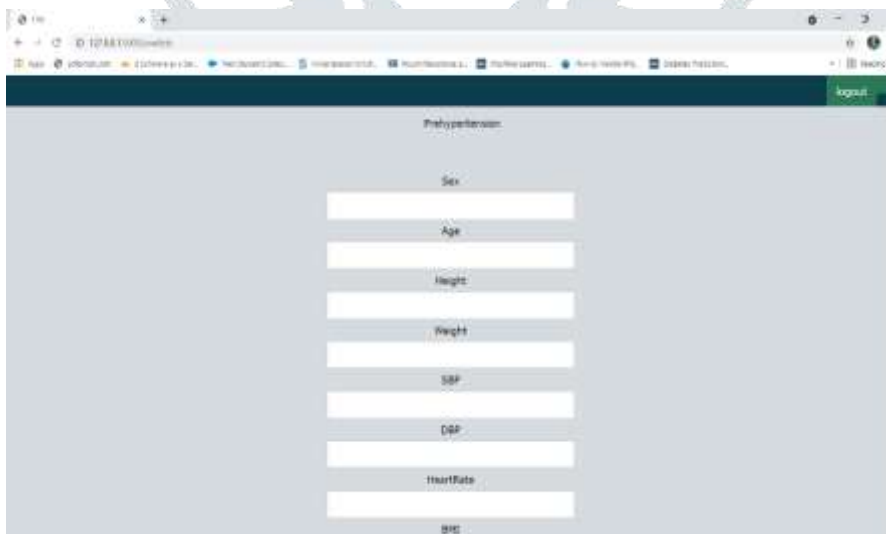


Fig.6: Hypertension prediction

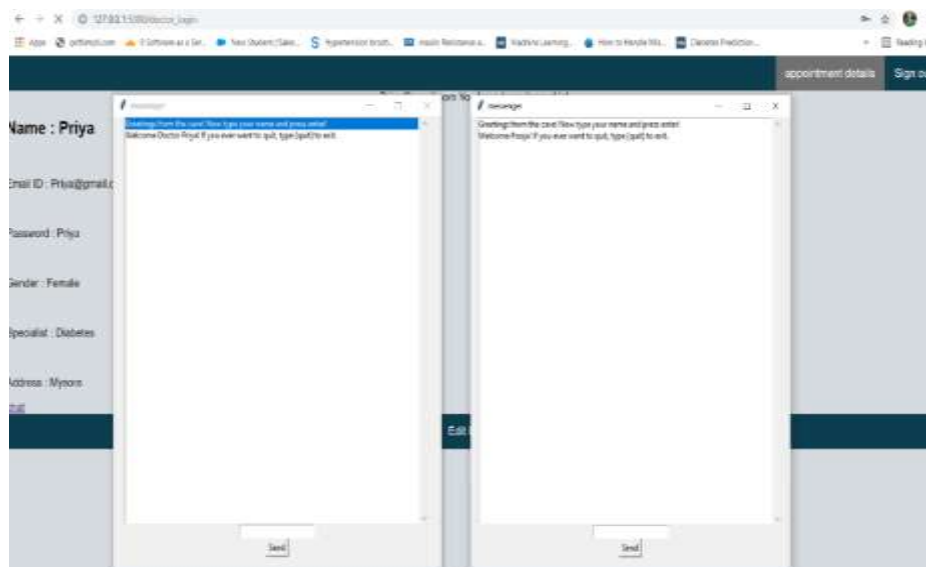


Fig.7: Live chat between doctor and patient



Fig.8: Automated healthcare Chatbot

V. CONCLUSION

In today's world most of the people are involved in hectic schedules, thereby neglecting their health care, which lead to chronic problems such as diabetes and hypertension. In this study, the proposed objective was to predict, whether the patient has been diagnosed with disease or not based on the various features. In the process data gathering along with model deployment was achieved. During model evaluation, various Machine Learning algorithms were compared on the basis of accuracy score finding the best algorithm. Finally, Web application and Chabot applications were designed and developed.

REFERENCE

- 1] Aishwarya Mujumdar, V Vaidehi, "Diabetes Prediction using Machine Learning Algorithms", Procedia Computer Science, Volume 165, 2019, Pages 292-299,ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2020.01.047>.

- 2] Sneha, N., Gangil, T. Analysis of diabetes mellitus for early prediction using optimal features selection. *J Big Data* 6, 13 (2019). <https://doi.org/10.1186/s40537-019-0175-6>.
- 3] SJ. Pradeep Kandhasamy, S. Balamurali. "Performance Analysis of Classifier Models to Predict Diabetes Mellitus".
- 4] <https://www.healthline.com/health/type-2-diabetes/ac1-test>
- 5] SJ. Pradeep Kandhasamy, S. Balamurali. "Performance Analysis of Classifier Models to Predict Diabetes Mellitus".
- 6] Ahmed Kareem, Lei Shi, Lin Wei, Yongcai Tao, "A Comparative Analysis and Risk Prediction of Diabetes at Early Stage using Machine Learning Approach".
- 7] Bani-Salameh, Hani & Alkhatib, Shadi & Abdalla, Moawyah & Banat, Ruaa & Zyod, Hala. (2020). Prediction of Diabetes and Hypertension using Multi-Layer Perceptron Neural Networks. *International Journal of Modeling, Simulation, and Scientific Computing*. 12. 10.1142/S1793962321500124.
- 8] Deepti Sisodia, Dilip Singh Sisodia, "Prediction of Diabetes using Classification Algorithm," *Procedia Computer Science*, Volume 132, 2018, Pages 1578-1585, ISSN 1877-0509,
- 9] P. Melin, I. Miramontes, and G. Prado-Arechiga, "A hybrid model based on modular neural networks and fuzzy systems for classification of blood pressure and hypertension risk diagnosis," *Expert Systems with Applications*, vol. 107, pp. 146–164, 2018.
- 10] Katayama K, Ishizawa H et al. Improvement of Blood Pressure Prediction Using Artificial Neural Network," 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rome, (2018), pp. 1-5, doi: 10.1109/MeMeA.2018.8438657.
- 11] N. Singh, P. Singh, and D. Bhagat, "A rule extraction approach from support vector machines for diagnosing hypertension among diabetics," *Expert Systems with Applications*, vol. 130, pp. 188–205, 2019.
- 12] <https://www.kaggle.com/uciml/pima-indians-diabetes-database>
- 13] https://figshare.com/articles/dataset/PPG-BP_Database_zip/5459299
- 14] B. Smitha Shekar and G. Harish, "A Machine Learning Model For Detection and Recognition Of Traffic Signs," 2021 International Conference on Intelligent Technologies (CONIT), IEEE xplore 2021, pp. 1-4, doi: 10.1109/CONIT51480.2021.9498270.
- 15] <https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm>
- 16] <https://www.javatpoint.com/logistic-regression-in-machine-learning>
- 17] <https://pymbook.readthedocs.io/en/latest/flask.html>
- 18] L. Athota, V. K. Shukla, N. Pandey and A. Rana, "Chatbot for Healthcare System Using Artificial Intelligence," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 619-622, doi: 10.1109/ICRITO48877.2020.9197833.