



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 7, July 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

Early Prediction of Diabetes Mellitus in Pregnant Women

Ankit N R¹, Arpita P M², Ashwin H V³, Divya G⁴, Prof. Raghu B R⁵

B.E. Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology,
Davanagere, Karnataka, India¹

B.E. Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology,
Davanagere, Karnataka, India²

B.E. Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology,
Davanagere, Karnataka, India³

B.E. Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology,
Davanagere, Karnataka, India⁴

Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and
Technology, Davanagere, Karnataka, India⁵

ABSTRACT: Diabetes is a chronic disease that can lead to a global health crisis. According to the International Diabetes Federation, 382 million people worldwide have diabetes. By 2035, this number will double to 592 million. Diabetes is a disease caused by elevated blood sugar levels. This high blood sugar level causes symptoms of frequent urination, increased thirst, and hunger. Diabetes is one of the leading causes of blindness, kidney failure, amputation, heart failure and stroke. When we eat, our body converts food into sugar or glucose. At this point, the pancreas should secrete insulin. Insulin acts as the key to unlocking our cells, allowing glucose to enter and use glucose for energy. However, in diabetes, this system does not work. Although type 1 and type 2 diabetes are the most common forms of the disease, there are other forms and other forms, such as gestational diabetes, that occur during pregnancy. Machine learning is a new branch of data science that studies how machines learn through experience. The purpose of this project is to develop a system that can more accurately and early predict diabetes in patients by combining the results of various machine learning methods. Algorithms such as random forest classifiers, neural networks, and decision trees are used. Each algorithm was used to calculate the accuracy of the model. The model is then adopted with good accuracy as a model for predicting diabetes.

KEYWORDS: Random Forest Classifier, Decision Tree, Artificial Neural Networks

I. INTRODUCTION

Deep Learning: Machine learning's deep learning subfield use algorithms to process data, model mental processes, or create abstractions. Layers of algorithms are used in deep learning (DL) to process data, comprehend spoken language, and visually identify objects. Each level is connected to the one before it, and the output of one level is utilised as the input of another. The input layer is the initial layer of the network, and the output layer is the final layer. The term "hidden layers" refers to all intermediate layers. Typically, each layer is a straightforward algorithm with a single type of activation function. Deep-learning architectures have been used in domains like computer science, including deep neural networks, deep belief networks, graph neural networks, recurrent neural networks, and convolutional neural networks.

Random Forest Classifier: Popular machine learning algorithm Random Forest is a part of the supervised learning methodology. It can be applied to ML issues involving both classification and regression. It is built on the idea of ensemble learning, which is a method of integrating various classifiers to address difficult issues and enhance model

performance. Random Forest, as the name implies, is a classifier that uses a number of decision trees on different subsets of the provided dataset and averages them to increase the dataset's predictive accuracy. Instead than depending on a single decision tree, the random forest uses forecasts from each tree and predicts the result based on the votes of the majority of predictions.

ANN:Artificial neural networks (ANNs) are computer networks with biological inspiration. Here, we concentrate on multilayer perceptrons (MLPs) with backpropagation learning algorithms among the many ANN variants. The most popular ANNs for a range of issues are MLPs, which have three layers: input, hidden, and output. They are based on a supervised approach. We talk about the structure, method, pre-treatment of the data, overfitting, and sensitivity analysis of MLPs. We also describe MLPs' benefits and drawbacks and suggest using them in ecological modelling. Finally, a case study illustrating the usefulness of MLP in ecological models is provided.

DecisionTree:The supervised learning algorithms family includes the decision tree algorithm. The decision tree technique, in contrast to other supervised learning methods, is capable of handling both classification and regression issues. By learning straightforward decision rules derived from previous data, a Decision Tree is used to build a training model that may be used to predict the class or value of the target variable (training data). In decision trees, we begin at the tree's root when anticipating a record's class label. We contrast the root attribute's values with that of the attribute on the record. We follow the branch that corresponds to that value and go on to the next node based on the comparison.

II. RELATED WORK

Kamrul Hasan et al. [1] Independent component analysis, correlation attribute selection, and principal component analysis characteristic selection methods were compared, and Naive Bayes, gradient boosting, decision trees, random forest, multilayer perceptron, and AdaBoost were applied. An accuracy of 78.9% was achieved using an ensemble of AdaBoost classifiers and gradient boosting.

Maniruzzaman et al [2] We worked with data sets taken from the National Health and Nutrition Survey. Objects were selected using a logistic regression method, and a random forest classifier gave the best results.

KM Jyoti Rani [3] Diabetes is a chronic condition that could lead to a global health care disaster. 382 million people worldwide have diabetes, according to the International Diabetes Federation. This will double to 592 million by 2035. Diabetes is a condition brought on by elevated blood glucose levels. The symptoms of this elevated blood sugar level include frequent urination, increased thirst, and increased hunger. One of the main causes of stroke, kidney failure, heart failure, amputations, blindness, and kidney failure is diabetes. Our bodies convert food into sugars, such as glucose, when we eat. Our pancreas is then expected to release insulin. Insulin acts as a key to unlock our cells, allowing glucose to enter and be used by us as fuel. However, this mechanism does not function in diabetes. The most prevalent forms of the disease are type 1 and type 2, but there are other varieties as well, including gestational diabetes, which develops during pregnancy. Data science has an emerging topic called machine learning that studies how machines learn from experience. The goal of this study is to create a system that, by fusing the findings of several machine learning approaches, can more accurately conduct early diabetes prediction for a patient.

MithushiSoni, Dr.Sunita Varma[4] Diabetes is a disease caused by high levels of glucose in the human body. Diabetes should not be ignored if left untreated. Diabetes can cause serious problems such as: B. Heart problems, kidney problems, blood pressure, eye damage, and it can also affect other organs in the human body. Early prediction can help control diabetes. To achieve this goal, the project will apply a variety of machine learning techniques to predict and improve the accuracy of diabetes in the human body and patients at an early stage. Machine learning techniques provide better results for prediction by building models from datasets collected from patients. Within the task, predict diabetes using classification and ensemble machine learning methods for datasets. Which K-nearest neighbor (KNN), logistic regression (LR), decision tree (DT), support vector machine (SVM), gradient boosting (GB), random forest (RF). Accuracy varies from model to model. Compare with other models. Model accuracy above the accuracy of doing the project work indicates that the model is effective in predicting diabetes. As a result, Random Forest is more accurate than other forests. Machine learning technology.

Mohebbi et al. [5] We used a conventional neural network and a perceptron multilayer neural network to detect diabetes in a dataset of 9 patients. The data set was based on a continuous blood glucose monitoring data set. Six patients were used for training and validation and the remaining three were used for testing. The existing neural network achieved the highest accuracy of 77.5%.

Pham et al. [6] The authors applied three different deep learning methods. The data set was collected by the authors at the hospital and consisted of 12,000 values. 1/6 of the data set was used for validation, 2/3 of the data set was used for

training and 1/6 for testing. The data set was pre-processed and reduced to 7191 values. The authors achieved a maximum accuracy value of 59.6%.

Tafa et al. [7] We split the data set into a 50% training set and a 50% test set. For diabetes prediction, a model was proposed using a combination of Naive Bayes and Support Vector Machines. Data sets were collected from three different locations and the proposed model was tested on these data sets. Eight attributes were present in the data set and consisted of 402 patients, of which 80 were patients with type 2 diabetes. The ensemble of naive bays and support vector machines achieved an accuracy of 97.6%, which is far superior to algorithms run individually on the dataset. In other words, the naive bay achieves an accuracy of 94.52 and the support vector machine reaches 95.52%. The author didn't mention any pre-processing methods to filter out unwanted values from the dataset.

III. METHODOLOGY

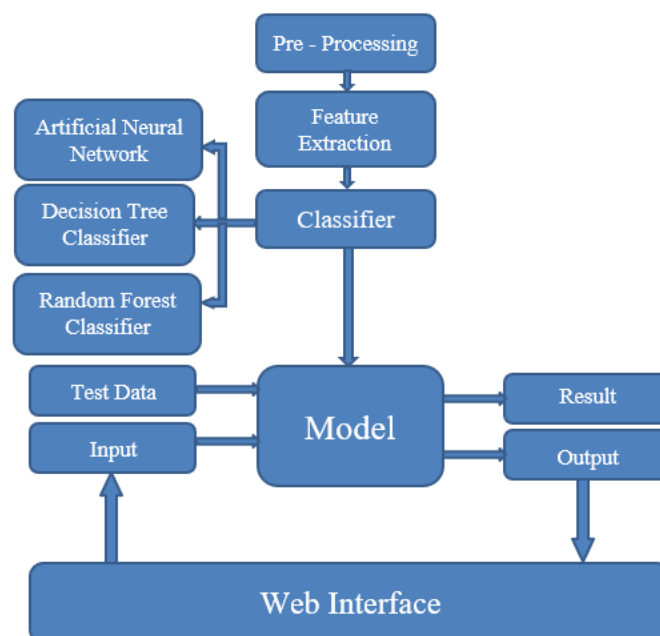


Fig 1: Methodology Diagram for Diabetes Prediction

Data collection:

Machines cannot learn without data because ML is so dependent on data. This is the most important aspect that enables algorithm training. Machine learning projects require training datasets. This is the actual dataset used to train the model to perform various actions.

Data Pre-processing:

Data Pre-processing includes the following main tasks : -

- Import the Libraries.
- Importing the dataset.
- Analyse the data.
- Taking care of Missing Data.
- Data Visualization
- Splitting Data into Train and Test.

IV. EXPERIMENTAL RESULTS

Figures show the results of predicting diabetes prediction in which the first screen contains some valuable information about the diabetes and their effects, fig 3 is the screen where the inputs are given to predict the diabetes in pregnant

women for which the output is shown in the fig 4, this example says the presence of diabetes for the provided inputs and fig 5 is the second example where for that input values the output is negative.

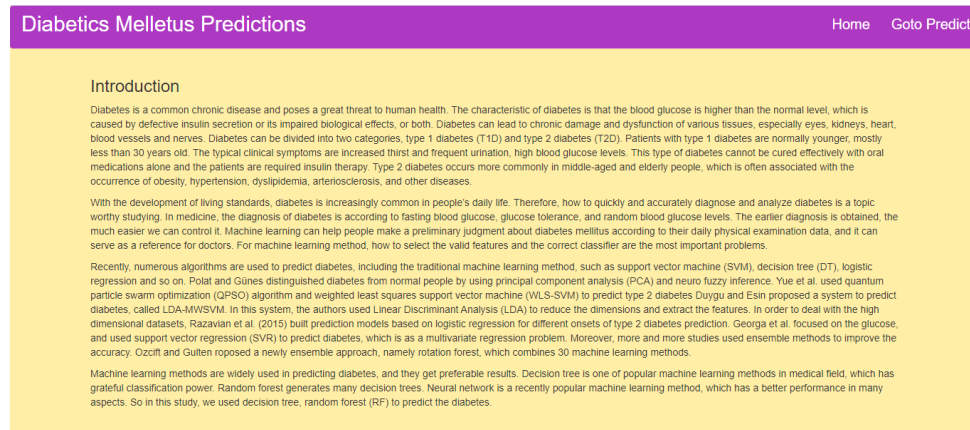


Fig 2: Homescreen of diabetes mellitus prediction in pregnant women

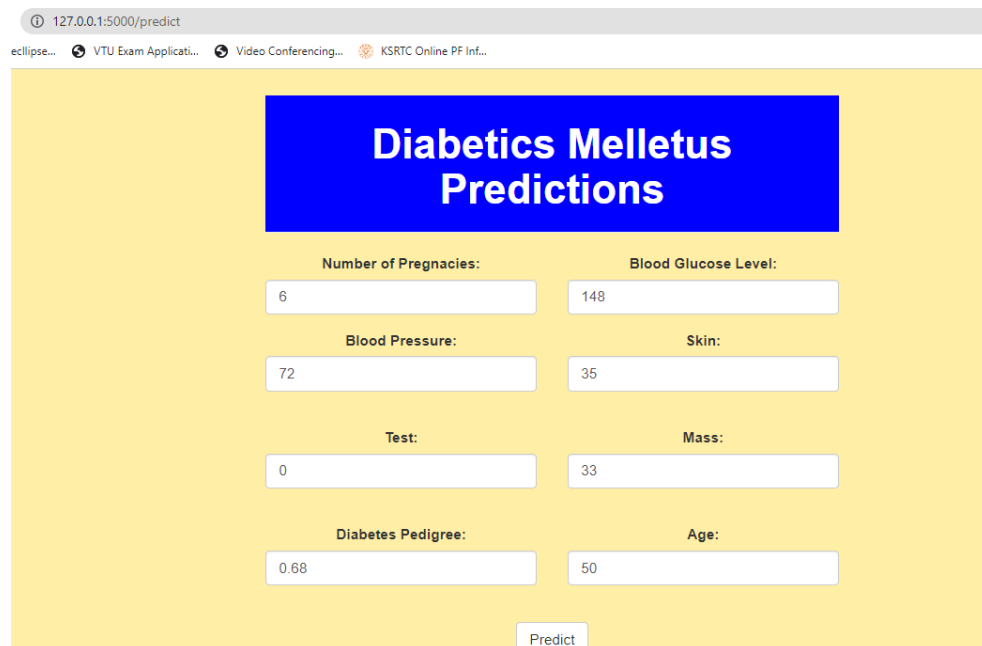


Fig 3: Providing Input Values to Predict Diabetes.

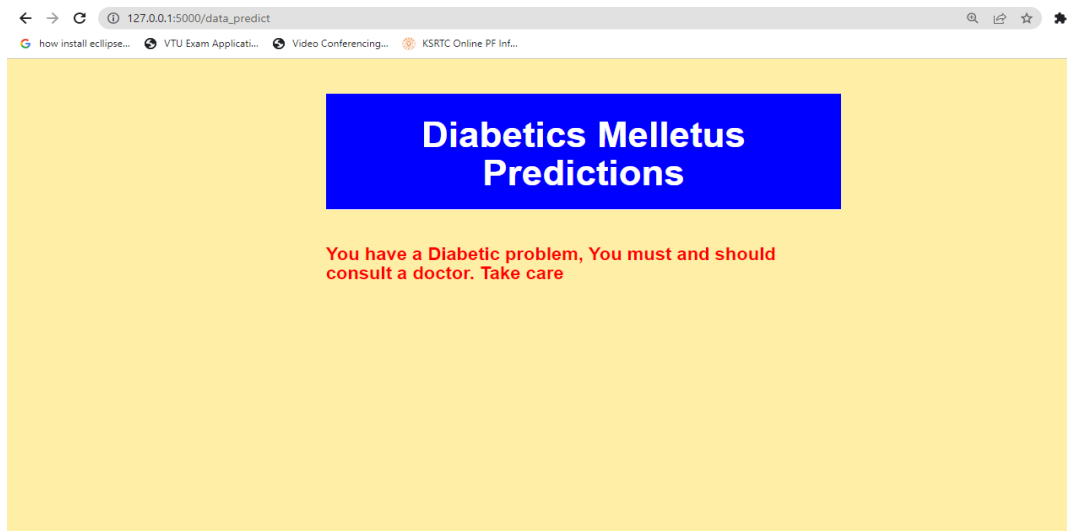


Fig 4: Output Screen with Results

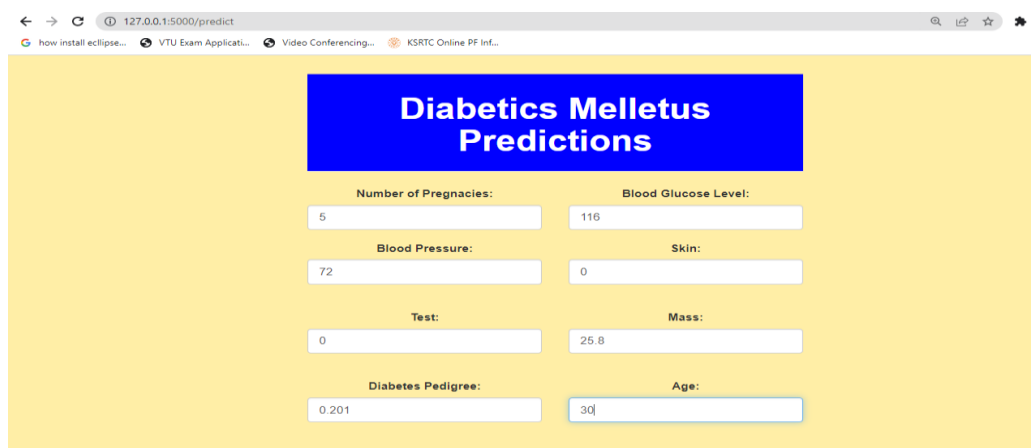


Fig 5: Inputs for a normal Patient



Fig 6: Pregnant Womendo nothaveaDiabeticProblem



V. CONCLUSION

Machine learning methods are useful for diagnosing diseases. The ability to predict diabetes early plays an important role in choosing the right treatment for the patient. Several existing classification methods for the medical diagnosis of diabetic patients based on accuracy have been discussed. A classification problem was found in the precision expression. Three machine learning methods were applied to the Pima Indians diabetes dataset and trained and validated on the test dataset. The implementation results of our model showed that the Random Forest classifier outperformed other models. Using association rule analysis, the results showed that there was a strong relationship between BMI and blood glucose levels.

REFERENCES

- [1] Md. Kamrul Hasan, Md. Ashraful Alam, D. Das, E. Hussain, and M. Hasan, "Diabetes prediction using ensembling of different machine learning classifiers," *IEEE Access*, vol. 8, pp. 76516–76531, 2020.
- [2] M. Maniruzzaman, M. J. Rahman, M. Al-Mehedi Hasan et al., "Classification and prediction of diabetes disease using machine learning paradigm," *Journal of Health Information Science and System*, vol. 8, 2020
- [3] K M Jyoti Rani , "Diabetes Prediction Using Machine Learning" July 2020, *International Journal of Scientific Research in Computer Science Engineering and Information Technology*
- [4] MitushiSoni , Dr. Sunita Varma, 2020, "Diabetes Prediction using Machine Learning Techniques", *International Journal Of Engineering Research & Technology (IJERT) Volume 09*, Issue 09 (September 2020)
- [5] Mohebbi, T. B. Aradottir, A. R. Johansen, H. Bengtsson, M. Fraccaro, and M. Morup, "A Deep Learning Approach to Adherence Detection for Type 2 Diabetics," in *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2896–2899, Korea, July 2017
- [6] T. Pham, T. Tran, D. Phung, and S. Venkatesh, "Predicting healthcare trajectories from medical records: a deep learning approach," *Journal of Biomedical Informatics*, vol. 69, pp. 218–229, 2017.
- [7] Z. Tafa, N. Pervetica, and B. Karahoda, "An Intelligent System for Diabetes Prediction," *IEEE Explore*, in *Proceedings of the 4th Mediterranean Conference on Embedded Computing (MECO)*, pp. 378–382, Budva, Montenegro, June 2015



INNO  SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details