International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 05 | May 2025 www.irjet.net p-ISSN: 2395-0072

Prediction of Diabetes and HyperTension using Machine Learning Techniques

Lokesh Naik[1], M Sakshi Rao[2], Nagendra Prasad B G[3], Navneeth C[4], Mr. Navile Nageshwara Naveen[5]

[1,2,3,4]Students, Computer Science and Engineering, Jyothy Institute of Technology, Bangalore, India
[5]Assisstant Professor, Computer Science and Engineering, Jyothy Institute of Technology, Bangalore, India

---

Abstract - *Diabetes, the worldwide health hazard, is distinguished by persistant blood sugar abnormality which precipitates acute complication if it's not diagnosed and treated. Prevalence of this disorder underlines imperative to have effective forecast tools so as to find candidates at risk to adopt measures against prevention. It considers critical variables like Age, BMI, Hyper Tension and Heart Disease while predicting the disorder. Also, this project is intended in forecasting hypertension and heart disease. We implement drill through approach by identifying presence or absence of hypertension, one of the most important parameters in identifying diabetics. Work is all about early forecasting of diabetics and one of its parameter. This forecasting system help healthcare professionals target resources to high-risk patients and enhance preventive care outcomes. Future developments will involve incorporating real-time health data and broadening to encompass genetic and environmental influences for further optimization of predictive accuracy and scalability.*

Key Words: **Diabetes, HyperTension, BMI, Logistic Regression**

1.INTRODUCTION

Diabetes is a critical illness that affects millions of people all over the world and challenges the health of the world population. It is seen as high sugar levels that emanate from lack of insulin, ineffective insulin as well as a combination of both. The traditional mechanisms used in testing for the disease such as fasting blood glucose tests as well as hemoglobin HBA1c tests offer information on sugar levels of individuals but may fail to detect the disease in its earliest stages in individuals suspected to be at risk. One of the best ways through which the risk of the disease can be determined is through the use of machine learning techniques that have the potential of understanding information on the levels of blood sugar that might not be easily detected through the use of traditional statistical approaches. It is through the use of these techniques that it is possible to train models through the use of large datasets and detect the high sugar levels as well as risk factors such as age, body mass index, physical activity as well as the presence of hypertension, thus making it easy to carry out a risk assessment of individuals. The use of wearable health monitoring devices as well as real-time data analytics can contribute to the accuracy of the estimates. It detects the presence of Hypertension which is one of the factors of detecting diabetic in individuals. It also seeks to search on the right food that can be taken by individuals (fruits and vegetables) in order to reduce the prevalence of diabetes among the people.

1.1 Motivation

Diabetics is one of the critical health issues that appears to be common among people. Traditional diagnosis of sugar level will assess the current condition of the person. Developing Developing machine learning model to detect diabetics aims to predict at early stages. The model takes into account the other key factors such as hypertension, BMI and age. Following drill through approach, it also aims in detecting hypertension as well. This initiative leverages in predicting diabetics and foods (fruits and vegetables) to reduce glucose levels.

1.2 Objective

The goal is to provide early detection of diabetics and future prediction as per key parameters such as Hypertension, BMI and age. While the traditional diagnosis assess the current health condition, Machine learning approach aims in predicting it in early days as well as possibility in future. Finally, detecting diabetics and one of its parameter hypertension, it leverages in timely assessing and analysing the health conditions upon certain variation.

1.3 Scope

The scope of this initiative is to detect and analyse diabetics and its parameter hypertension. Providing additional prediction for heart failure aims in assessing the major health issues that the world faces. Predicting diabetics and other health issues by considering parameters such as BMI, hypertension, smoking habits and age, analysis the effect of these attributes on health condition. The scope of this work also includes suggesting the appropriate in take of fruits and vegetables in the diet to reduce their glucose levels.

2. UNDERSTANDING DIABETICS

Diabetes is a severe health condition attributed to high blood sugar levels and can lead to severe complications if not managed properly. The condition occurs when the pancreas cannot produce the required insulin (Type 1 diabetes) or when the produced insulin is not effective in the body (Type 2 diabetes). Insulin is a hormone secreted by the pancreas to allow the blood sugar to enter the cells from the bloodstream, where it is used as a source of energy. Diabetics can also occur due to genetics as well. Traditional diagnosis of diabetes takes into account the glucose level before and after food in take as well as HbA1c level. Usually in type 1 diabetes, your pancreas does not make insulin or makes very little insulin. Type 1 diabetes can affect people at any age, but commonly develops in children and young adults. People having type 1 diabetes will need daily insulin injections to control their glucose levels. The primary indicator of type 2 diabetes is insulin resistance, when the body cannot fully respond to insulin. Type 2 diabetes is the most common type of diabetes. The normal range of glucose level is as follows. - *Fasting: 70-99 mg/dL; Postprandial (after meal): <140 mg/dL (1 hour), <120 mg/dL (2 hours).* The normal range for HbA1c is as follows*. Normal: <5.7% ; Prediabetes: 5.7-6.4% ; Diabetes: ≥6.5%.*

2.1 Machine Learning in Diabetes Prediction

Machine learning is a transformative technology that empowers computers to learn from dataset, extract hidden patterns, and make informed decisions without explicit programming. By implementing algorithms and statistical models, machine learning enables to predict outcomes based on feature value. ML is classified into supervised and unsupervised learning. Supervised learning is used when the target outcome is known while unsupervised learning is used when the hidden pattern is to be discovered. This initiative involves the use of supervised learning model such as logistic regression.

The diabetes prediction is made by training large dataset that contain records such as age, BMI and hypertension to yield appreciable accuracy.

2.2 Data Visualization for Diabetes Prediction

Data visualization in this work aims to find the trends in the prediction records so far. It visualises the predictions made so far and suggest the path for improvement. by transforming raw stock market data into meaningful visual insights that enhance user understanding and decision-making. The platform utilizes advanced libraries like React Charts to render real-time and historical data through interactive and dynamic visual elements. Additional tools such as zoom, drag, and tooltip features enable deeper inspection of price movements and trading patterns.

The visualization enables to visualize the rise and fall of parameter values based on prediction history. Though the interface allows visualising the trends in parameters, the relation between each of the attributes that determine the existence of diabetes is also examined.

The following shows the visualisation made in this work.

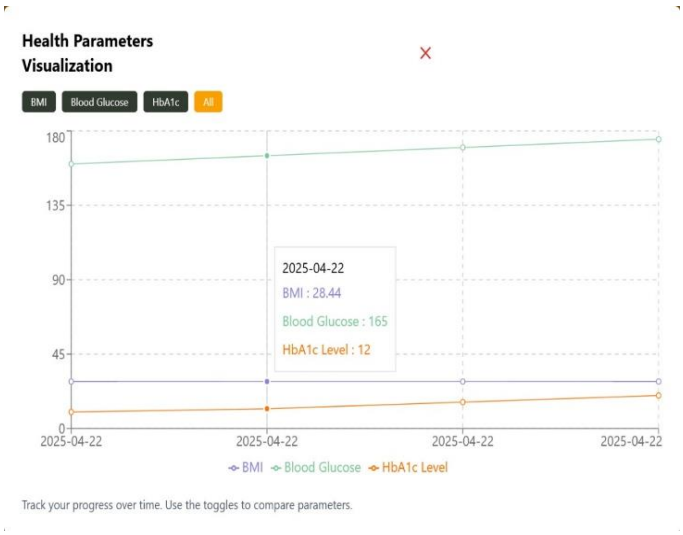

Fig 1: Trends in parameter of predicting diabetes

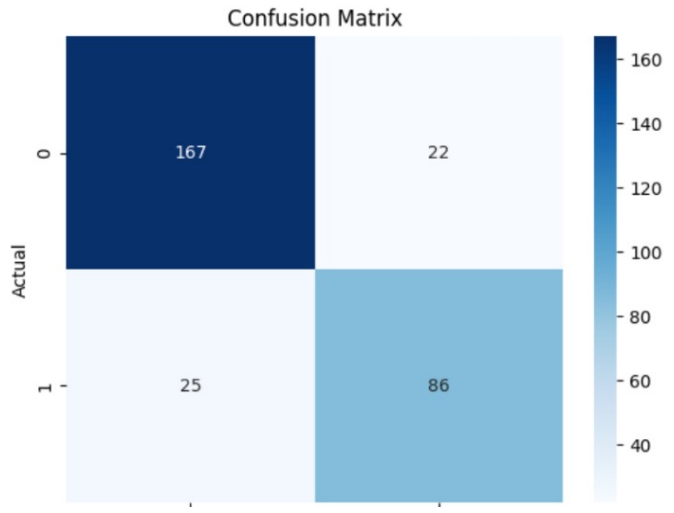

Fig 2: Confusion matrix in predicting diabetes

3. RELATED WORK

3.1 Literature Survey

Numerous studies have explored the application of machine learning and data science techniques for stock market forecasting and analysis. Montagna et al. [1] explored machine learning models for hypertension detection using data collected from World Hypertension Day testing. The study demonstrated that Random Forest

achieved a sensitivity of 81.8% and specificity of 62.9%, making it one of the better-performing models for hypertension risk classification. Similarly, Islam et al. [2] applied explainable artificial intelligence (XAI) techniques to hypertension prediction in Ethiopia using XGBoostCahn et al. [3] developed a gradient-boosted tree model to predict diabetes onset from pre diabetes using electronic medical records (EMR) data. The model achieved an AUC (Area Under Curve) value exceeding 0.86 across multiple datasets, outperforming traditional Logistic Regression models. Meanwhile, Soni and Varma (2020) conducted an extensive comparison of machine learning algorithms, including SVM, KNN, and Random Forest, for diabetes prediction using the Pima Indian Diabetes Dataset (PIDD). The study found that Random Forest outperformed other models in terms of accuracy, demonstrating the effectiveness of ensemble methods in medical data classification. [5] A recent 2024 study explored deep learning and hybrid models for diabetes detection, integrating machine learning (ML) with deep learning (DL) to improve classification accuracy. The hybrid approach outperformed traditional models, demonstrating higher classification rates and improved feature extraction capabilities. Raja Krishnamoorthi et al. [6] suggests a workflow for diabetes prediction based on their findings. Machine learning models are examined, and machine learning-based architecture for diabetes prediction is proposed and evaluated by the authors. [7] In this study, diabetes prediction is done using different ML algorithms on the dataset created by using samples from PIMA Indian Diabetes dataset. Jayroop Ramesh et al. [8] proposed a remote monitoring workflow for automated diabetes risk prediction and management, using personal health devices, smart wearables and smartphones. [9] This disease was examined by machine learning (ML) algorithms in this paper. The goal behind this research is to create an effective model with high precision to predict diabetes. In this work, K-nearest neighbor algorithm is used. Isfafuzzaman Tasin et al. [10] used the Pima Indian diabetes dataset and collected additional samples from 203 individuals from a local textile factory in Bangladesh. This work uses ML classification methods, that is, decision tree, SVM, Random Forest, Logistic Regression and KNN to determine which algorithm produces the best prediction results

4. PROPOSED MODEL

4.1 Workflow

The proposed operational framework of diabetes prediction adheres to a systematic methodology that aids predictors in analysing the outcome based on feature values. The platform provides prediction of various health issues at the user disposal. The user gives the necessary parameter values for the prediction. The model gives the outcome with certain appreciable accuracy.

The dataset is trained with the appropriate machine learning model (logistic regression in this case). The parameter values entered by the user is mapped against the trained dataset and returns with the output that aligns with the trained records. The model is expected to provide result with better accuracy.

The user selects the health condition that they want to assess. The user is then asked for the right values of the parameters which upon providing responds with the outcome that matches with the trained records.

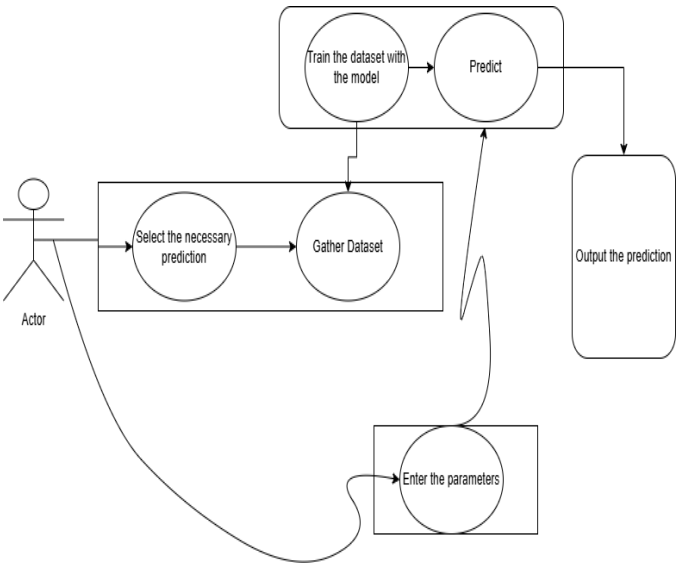

Fig 3: Workflow

4.2 Techniques Used

This task entails the use of Python to perform machine learning-related tasks. The tools used in this project include pandas for reading and cleaning the data. It also uses scikit learn for various machine learning models. The displayed output of the task is a result of using the react js frontend. The prediction dataset is recorded and stored in MongoDB.

With data saved within the database, I will show how the attributes affect the disease load and the rate at which they grow and decline. Using Machine learning to predict the likelihood of getting the disease will include preparing necessary software and hardware, creating a middleware and the use of experimental methodology in both training the model and evaluating results. I will determine a model for disease prediction, determine evaluation metrics like accuracy and deploy the model to the real life use. I will use the same configuration to detect heart failures and also spread to cancer detection it will use a linear regression machine learning model to achieve appreciable accuracy and prediction. The aim of this initiative is to assist the medical staffs or the patients for early prediction of any health issues that the project covers to resolves.

4.3 Implementation Details

This project is centered on the use of machine learning (ML) algorithms to forecast Diabetes and Heart Failure using medical information. Through patient data like age, blood pressure, glucose level, and other medical history, predictive models that will be able to detect patients who are likely to develop these diseases are intended to be created. The primary concept behind the project is using supervised learning algorithms to categorize a person as being at risk or not being at risk of having diabetes and heart failure on the basis of input features. Application of a machine learning system for prediction of Diabetes and Heart Failure encompasses numerous fundamental steps like data gathering, preprocessing, selection of model, training, testing, and deployment.

The front-end side of the application was created very effectively through React.js that offers the end-user the opportunity to get a dynamic, responsive, and well-built interface that significantly improves the interaction with the user. For illustrating the records about the forecasts, there are the charting components that are built in a very advanced way and they are developed through Chart.js and D3.js that offer the possibility to show the differences between the attributes.

The following gives the detailed work and procedure of implementation. *Data Preprocessing*: (i)Handling Missing Data. (ii)Imputation techniques such as mean, median, or k NN imputation. *Feature Selection:* (i)Identify and retain only the most significant predictors of diabetes. (ii)Normalization and Scaling: Standardize features to ensure consistency across models. *Model Development*: (i)Supervised Learning Models: Logistic regression, decision trees, support vector machines and random forests. *Performance Evaluation*: (i) Compare models based on accuracy, precision, and recall. Visualization and Interpretability: Feature Importance Analysis: Identify the most influential variables in predicting diabetes. *Web Application:* A user-friendly interface for clinicians and researchers.

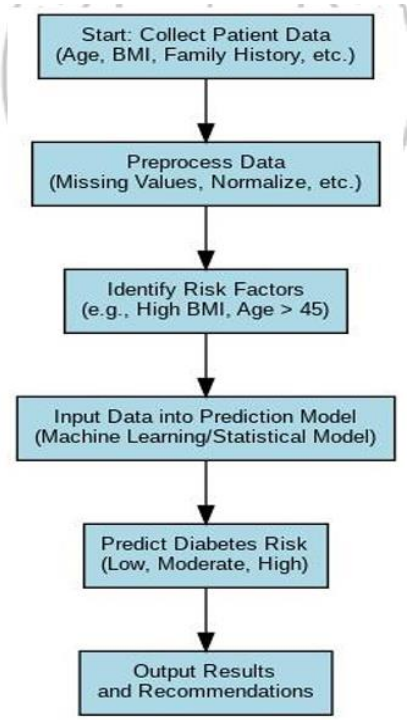

Fig 4: Flowchart

Along with the prediction of diabetes, the work also provides the testing of hypertension which is considered as one of the parameters of diabetes. It also gives the suggestion of the fruits and vegetables to eat to reduce glucose levels and blood pressure. The model shows appreciable accuracy in predicting the outcome. The classification report in predicting diabetes is shown below.

```
Accuracy: 0.8433333333333334

Classification Report:
              precision    recall  f1-score   support

           0       0.87      0.88      0.88       189
           1       0.80      0.77      0.79       111

    accuracy                           0.84       300
   macro avg       0.83      0.83      0.83       300
weighted avg       0.84      0.84      0.84       300
```

Fig 5: Classification Report

5. Future Trends

The future trends of this work is to implement predicting heart issues and glucose levels and blood pressure on a wearable device. Combine EHRs, genetic information, and lifestyle data for comprehensive risk assessment.

5. CONCLUSION

The Diabetes and Hypertension Detection System utilizing machine learning provides a robust framework for predicting the likelihood of individuals developing diabetes or heart failure based on key health features. The process of building this system began with gathering and preprocessing the relevant datasets, followed by applying multiple machine learning algorithms for model training. After thorough evaluation using various performance metrics, the best-performing models were selected for both diabetes and heart failure prediction. The system offers practical utility by enabling early detection, potentially improving patient outcomes through timely interventions. Furthermore, the deployment of the model as an accessible tool for healthcare professionals or patients ensures real-time predictions, making it a valuable asset in healthcare decision-making. As healthcare data evolves and improves, the system can be continuously retrained and updated to maintain its predictive accuracy and adaptability. The project successfully demonstrated how machine learning can be integrated into healthcare to assist in predicting and preventing serious health conditions, ultimately contributing to more proactive and personalized medical care.

REFERENCES

[1] Montagna, S., Pengo, M. F., Ferretti, S., et al. (2023). Machine Learning Hypertension Detection: A Study on World Hypertension Day Data. Journal of Medical Systems, 47(1). https://doi.org/10.1007/s10916-022-01900-5.

[2] Islam, M. M., Alam, M. J., Maniruzzaman, M., et al. (2023). Predicting the Risk of Hypertension Using Machine Learning Algorithms: A Cross-Sectional Study in Ethiopia. PLOS ONE, 18(8). https://doi.org/10.1371/journal.pone.0289613.

[3] Cahn, A., Shoshan, A., Sagiv, T., et al. (2019). Prediction of Progression from Pre-Diabetes to Diabetes: Development and Validation of a Machine Learning Model. Diabetes Metabolism Research and Reviews. https://doi.org/10.1002/dmrr.3252.

[4] Soni, M., & Varma, S. (2020). Diabetes Prediction Using Machine Learning Techniques. International Journal of Engineering Research & Technology, 9(9). http://www.ijert.org.

[5] Mujumdar, A., & Vaidehi, V. (2019). Diabetes Prediction Using Machine Learning Algorithms. Procedia Computer Science, 165, 292–299.

[6] A Novel Diabetes Healthcare Disease Prediction Framework Using Machine Learning Techniques - Hindawi Journal of Healthcare Engineering Volume 2023, Article ID 9872970

[7] Bavkar VC Machine learning algorithms for Diabetes prediction and neural network method for blood glucose measurement. Indian Journal of Science and Technology 14(10): 869-880. https://doi.org/10.17485/IJST/v14i10.2187

[8] Jayroop Ramesh – A remote healthcare monitoring framework for diabetes prediction using machine learning.

[9] N. Krishna and K. Nagamani, "Understanding and visualisation of different feature extraction processes in glaucoma detection," Journal of Physics: Conference Series, vol. 2327, pp. 012023, 2022.

[10] V. Kumar, D. Velpula, D. Sharma, et al., "Glaucoma detection with explainable AI using convolutional neural networks based feature extraction and machine learning classifiers," IET Image Processing,