

Feedback Mechanism in Diabetes Classification

1. Shrey Sandhu

Enrollment No. - 15314803119

(Information Technology scholar at Maharaja Agrasen Institute of Technology, Delhi-India)

2. Shubham Mishra

Enrollment No. - 15614803119

(Information Technology scholar at Maharaja Agrasen Institute of Technology, Delhi-India)

Under supervision of:-

1. Dr. ML Sharma(H.O.D)

(Department of IT , Maharaja Agrasen Institute of Technology, Delhi-India)

2. Dr. KC Tripathi

(Department of IT , Maharaja Agrasen Institute of Technology, Delhi-India)

Abstract.

Diabetes is becoming more common every day, making it important and concerning to predict and diagnose these conditions accurately and efficiently. This research paper focuses on identifying patients who are more likely to have diabetes based on various medical attributes and also examines which attributes are contributing to diabetes. To address this issue, we developed a diabetes prediction system that uses a machine learning algorithm i.e Support Vector Machine (SVM), to predict and classify patients with diabetes. This system helps to relieve a significant amount of pressure on medical professionals by accurately identifying the likelihood of a patient having diabetes. In addition, the diabetes prediction system improves medical care and reduces costs. Overall, this project provides valuable knowledge that can aid in the prediction of diabetes in patients and is implemented in .pynb format.

Introduction

Machine Learning is a method of extracting and manipulating useful information from data, whether previously known or unknown. It is a broad and diverse field that is constantly growing in scope and application. There are several types of machine learning, including supervised, unsupervised, and ensemble learning, which are used to predict and assess the accuracy of a given dataset. Diabetes is a noxious disease in the world. Diabetes caused because of obesity or high blood glucose level, and so forth. It affects the hormone insulin, resulting in abnormal metabolism of crabs and improves the level of sugar in the blood. Diabetes occurs when the body does not make enough insulin. According to (WHO) World Health Organization About 422 million people suffer from diabetes particularly from low or idle income countries. And this could be increased to 490 billion up to the year of 2030. However Prevalence of diabetes is found among various Countries Like Canada, China, and India etc. Population of India is now more than 100 million so the actual number of diabetics in India is 40 million. Diabetes is the major cause of death in the world. Early prediction of diseases like diabetes can be controlled and save human life. To accomplish this, this work explores prediction of diabetes by taking various attributes related to diabetes .

The main focus of the project is Support Vector Machine, a supervised learning technique that uses discrete values. The goal of the project is to determine whether a patient is likely to be diagnosed with diabetes based on their medical attributes, such as glucose, blood pressure ,skin thickness ,and insulin. To do this, we used a dataset from kaggle.com and applied logistic regression to classify those at risk of diabetes.

METHODOLOGY

The system architecture provides an overview of how the system operates. The process begins by collecting data and selecting relevant attributes. The necessary data is then pre-processed into the necessary format and split into two parts: **training and testing data**. The algorithms are applied and the model is trained using the training data. The system's accuracy is evaluated by testing it using the testing data. This system is implemented using the following modules.

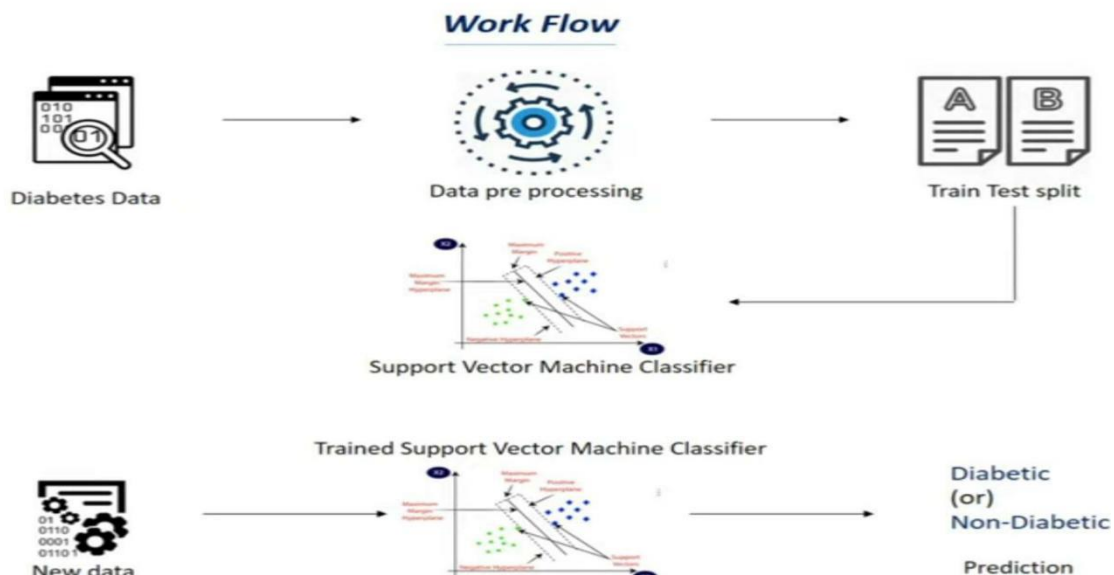
1. Collection of Dataset
 2. Selection of attributes
 3. Data Pre-Processing
 4. Balancing of Data
 5. Disease classification
 6. Feedback
1. Collection of dataset: To build our diabetes prediction system, we first collect a dataset. We then split the dataset into training data and testing data.
 2. Selection of attributes: In order to improve the efficiency of the system, we carefully select relevant attributes, such as glucose, blood pressure, skin thickness and insulin.
 3. Data pre-processing: To ensure the accuracy of our model, we pre-process the data to clean it and transform it into a usable format. This includes importing the dataset, splitting it, and scaling attributes as needed.

4. Classification: We use Support Vector Machine (SVM), which gives the highest accuracy, to classify diabetes.
5. Feedback: Based on the prediction, the model provides feedback on the factors that may be contributing to diabetes.

DATASET

S No.	Attributes
1	Pregnancy
2	Glucose
3	Blood Pressure
4	Skin thickness
5	Insulin
6	BMI(Body Mass Index)
7	Diabetes Pedigree Function
8	Age

WORKFLOW DIAGRAM



CONCLUSION

The goal of this project is to investigate the use of machine learning techniques for classifying diabetes. In the healthcare industry, the role of a classifier is crucial in determining appropriate treatment for patients. In this study, we compare and evaluate existing techniques in order to identify efficient and accurate systems. Our findings suggest that machine learning techniques can significantly improve the accuracy of predicting diabetes, allowing for early identification and prevention of diabetes. Overall, there is significant potential for the use of machine learning algorithms in the prediction and prevention of diabetes.

References and Bibliography

- [1] Z. Sabouri, Y. Maleh, and N. Gherabi, "Benchmarking Classification Algorithms for Measuring the Performance on Maintainable Applications," in *Advances in Information, Communication and Cybersecurity*, Cham, 2022, pp. 173–179. doi: 10.1007/978-3-030-91738-8 17.
- [2] H. EL Massari, S. Mhammedi, Z. Sabouri, and N. Gherabi, "OntologyBased Machine Learning to Predict Diabetes Patients," in *Advances in Information, Communication and Cybersecurity*, Cham, 2022, pp. 437–445. doi: 10.1007/978-3-030-91738-8 40.
- [3] F. Alaa Khaleel and A. M. Al-Bakry, "Diagnosis of diabetes using machine learning algorithms," *Mater. Today Proc.*, Jul. 2021, doi: 10.1016/j.matpr.2021.07.196.
- [4] J. J. Khanam and S. Y. Foo, "A comparison of machine learning algorithms for diabetes prediction," *ICT Express*, vol. 7, no. 4, pp. 432–439, Dec. 2021, doi: 10.1016/j.ict.2021.02.004.
- [5] P. Cihan and H. Coskun, "Performance Comparison of Machine Learning Models for Diabetes Prediction," in *2021 29th Signal Processing and Communications Applications Conference (SIU)*, Jun. 2021, pp. 1–4. doi: 10.1109/SIU53274.2021.9477824.
- [6] M. A. Sarwar, N. Kamal, W. Hamid, and M. A. Shah, "Prediction of Diabetes Using Machine Learning Algorithms in Healthcare," in *2018 24th International Conference on Automation and Computing (ICAC)*, Sep. 2018, pp. 1–6. doi: 10.23919/ICAC.2018.8748992.
- [7] A. Mujumdar and V. Vaidehi, "Diabetes Prediction using Machine Learning Algorithms," *Procedia Comput. Sci.*, vol. 165, pp. 292–299, Jan. 2019, doi: 10.1016/j.procs.2020.01.047.
- [8] M. Rady, K. Moussa, M. Mostafa, A. Elbasry, Z. Ezzat, and W. Medhat, "Diabetes Prediction Using Machine Learning: A Comparative Study,"