K.S.R COLLEGE OF ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

MACHINE LEARNING MODEL FOR DIABETES PREDICTION

GUIDED BY,

Mr.V.RAMESH, M.E.,
(ASSISTANT PROFESSOR/CSE)

TEAM MEMBERS,

1.PRASANTH S (2013076)
2.RAMASUBRAMANIYAN M (2013082)
3.SABARINATHAN M (2013087)

ABSTRACT

- Due to its continuously increasing occurrence, more and more families are influenced by diabetes mellitus. Most diabetics know little about their health quality or the risk factors they face prior to diagnosis. In this study, we have proposed a novel model based on data mining techniques for predicting type 2 diabetes mellitus (T2DM).
- The main problems that we are trying to solve are to improve the accuracy of the prediction model, and to make the model adaptive to more than one dataset. Based on a series of preprocessing procedures, the model is comprised of MLP, NB and RF algorithm.
- Moreover, our model ensures that the dataset quality is sufficient. To further evaluate the performance of our model, we applied it to two other diabetes datasets. Both experiments' results show good performance. As a result, the model is shown to be useful for the realistic health management of diabetes.

INTRODUCTION

- Diabetes is one of the most common diseases in recent years, and its global prevalence is growing rapidly.
- It is a general term for heterogeneous disturbances of metabolism for which the main finding is chronic hyperglycemia. The cause is either impaired insulin secretion or impaired insulin action or both.
- The chronic hyperglycemia of diabetes is associated with long-term damage, dysfunction, and failure of various organs, especially the eyes, kidneys, nerves, heart, and blood vessels.
- The vast majority of diabetes can be divided into two categories, viz. Type 1 and Type 2. The cause of Type 1 diabetes is an absolute deficiency of insulin secretion.
- On the other hand, Type 2 diabetes is much more prevalent, and the cause is a combination of resistance to insulin action and an inadequate compensatory insulin secretary response

OBJECTIVES

- Accurately predict the likelihood of a patient developing diabetes: The primary objective of a machine learning model for diabetes prediction is to accurately predict the likelihood of a patient developing diabetes. The model should be able to take in various patient parameters and historical data and output a prediction of the likelihood of diabetes.
- Early detection of diabetes: Another objective of a machine learning model for diabetes prediction is to enable early detection of diabetes in patients. Early detection can help patients take necessary steps to manage their condition and prevent complications.
- Personalized treatment plans: By accurately predicting the likelihood of diabetes, a machine learning model can help healthcare providers create personalized treatment plans for patients. These plans can be tailored to each patient's individual needs, improving the efficacy of treatment and ultimately leading to better health outcomes.

PROBLEM STATEMENT

- Diabetes is a chronic disease that affects millions of people worldwide. Early diagnosis of diabetes can lead to effective treatment and management of the disease, which can prevent serious complications such as blindness, kidney failure, and heart disease. However, many people with diabetes are not diagnosed until the disease has progressed to an advanced stage, which can make treatment more difficult and less effective.
- Therefore, the problem is to develop a predictive model that can accurately identify individuals who are at risk of developing diabetes. The model should take into account various risk factors such as age, gender, family history, BMI, blood pressure, and glucose levels, among others. The goal is to create a tool that healthcare providers can use to identify at-risk individuals and provide early intervention and preventative care. Such a model can improve the overall health outcomes of patients and reduce the burden of diabetes on healthcare systems.

EXISTING SYSTEM

- In the medical field, it is essential to predict diseases early to prevent them. Diabetes is one of the most dangerous diseases all over the world. In modern lifestyles, sugar and fat are typically present in our dietary habits, which have increased the risk of diabetes. To predict the disease, it is extremely important to understand its symptoms.
- Currently, machine-learning (ML) algorithms are valuable for disease detection. This article presents a model using a fused machine learning approach for diabetes prediction. The conceptual framework consists of two types of models: Support Vector Machine (SVM) and Artificial Neural Network (ANN) models.
- A cloud storage system stores the fused models for future use. Based on the patient's real-time medical record, the fused model predicts whether the patient is diabetic or not.

DRAWBACKS

- Data Quality and Quantity: The effectiveness of machine learning models heavily depends on the quality and quantity of the training data. If the dataset used for training is small, unrepresentative, or contains errors or biases, it can lead to poor model performance and generalization.
- Feature Selection and Engineering: The success of machine learning models often requires careful feature selection and engineering. If important features are not included or irrelevant features are included, the model's predictive power could be compromised.
- Model Complexity: Combining multiple machine learning algorithms (SVM, ANN, and fuzzy logic) can lead to a complex and computationally intensive model. Such complexity might make it difficult to interpret the model's decisions or to debug and troubleshoot issues.

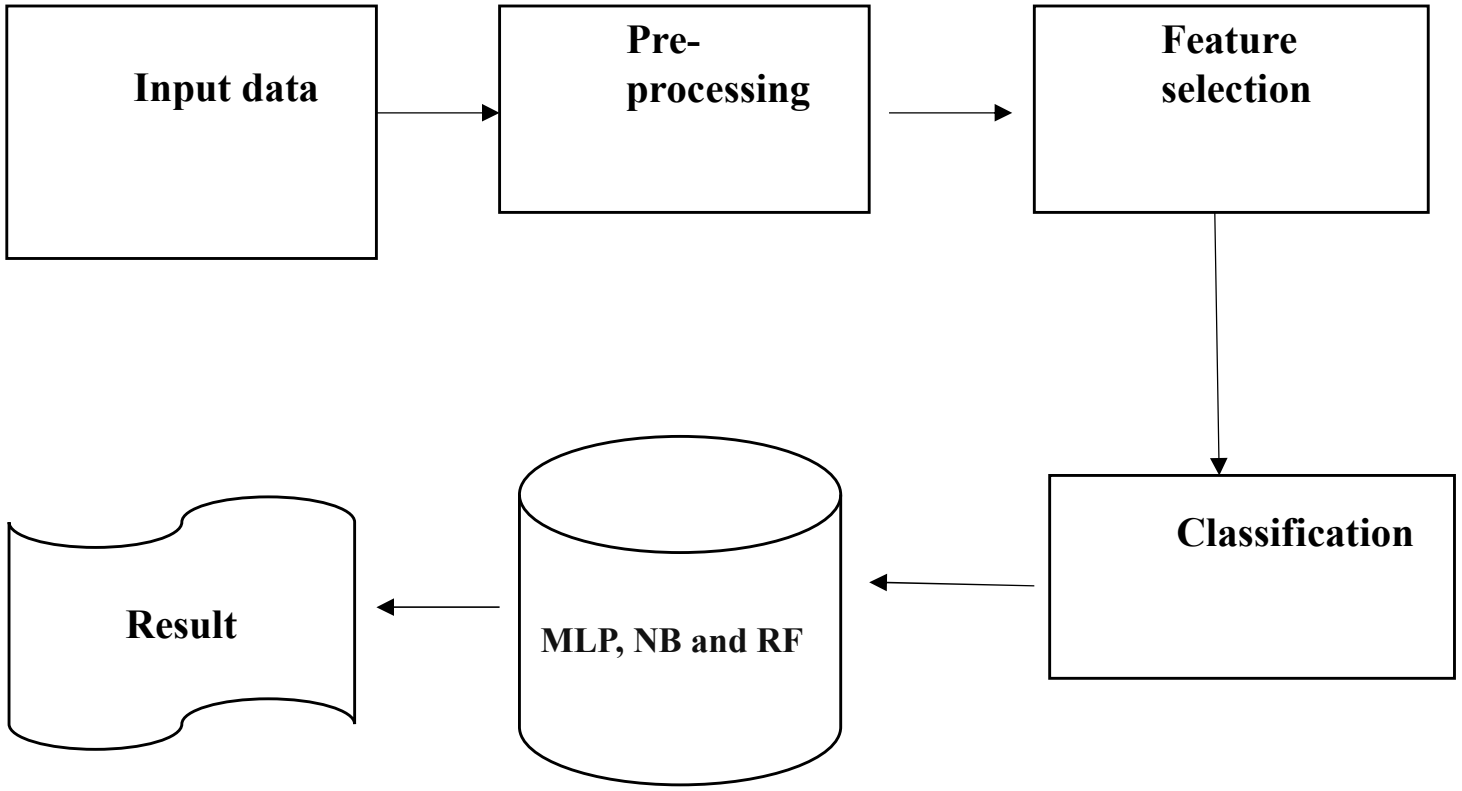
PROPOSED SYSTEM

- In this proposed work, the main objective is to classify the data as diabetic or non-diabetic and improve the classification accuracy. For many classification problem, the higher number of samples chosen but it doesn't leads to higher classification accuracy.
- This survey has analyzed various classification techniques for classification of diabetic and non-diabetic data. Thus, it is observed that techniques like MLP, NB and RF are most suitable for implementing the Diabetes prediction system.
- Machine learning has the great ability to revolutionize the diabetes risk prediction with the help of advanced computational methods Availability of large amount of epidemiological and genetic diabetes risk dataset.
- Detection of diabetes in its early stages is the key for treatment. This work has described a machine learning approach to predicting diabetes levels. The technique may also help researchers to develop an accurate and effective tool that will reach at the table of clinicians to help them make better decision about the disease status.

ADVANTAGES

- In many cases, the performance of algorithm is high in the context of speed but the accuracy of data classification is high.
- The performance of the model can be described by Confusion Matrix as False Negative (FN), False Positive (FP), True Negatives (TN), True Positives (TP).
- These models that were used to fit the training set were compared and estimated their performance in terms of precision, recall , accuracy and ROC-AUC-Score.
- The results obtained in this study have achieved high accuracy rate for predicting diabetes when compared with other existing methods.
- Improved Accuracy.
- Better Performance with Minimal Area Under Curve Values Appeared
- The correctness of the model in predicting the instances is measured in terms of accuracy

DIAGRAM


Input data
Pre-
processing
Feature
selection
Classification
MLP, NB and RF
Result

SYSTEM SPECIFICATION

HARDWARE REQUIREMENTS

- Processor Type : RYZEN 7
- Speed : 3.40 GHZ
- RAM : 8 GB
- Hard disk : 500 GB
- Keyboard : 101/102 Standard Keys
- Mouse : Optical Mouse

SOFTWARE REQUIREMENTS

- Operating System : Windows 10
- Coding Language : Python
- Tools : Weka Tools

REFERENCES

- Sneha, N., Gangil, T. ,”Analysis of diabetes mellitus for early prediction using optimal features selection,” in: Journal of Big Data 6, 13 (2020)
- Q. Wang, W. Cao, J. Guo, J. Ren, Y. Cheng and D. N. Davis, "DMP_MI: An Effective Diabetes Mellitus Classification Algorithm on Imbalanced Data With Missing Values," in IEEE Access, vol. 7, pp. 102232-102238, 2021
- S. Perveen, M. Shahbaz, K. Keshavjee and A. Guergachi, "Metabolic Syndrome and Development of Diabetes Mellitus: Predictive Modeling Based on Machine Learning Techniques," in IEEE Access, vol. 7, pp. 1365-1375, 2020
- M. Goyal, N. D. Reeves, S. Rajbhandari and M. H. Yap, "Robust Methods for Real-Time Diabetic Foot Ulcer Detection and Localization on Mobile Devices," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pp. 1730-1741, July 2021

THANK YOU