IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

Diabetes Prediction using Machine Learning Algorithm

Ms. Ankita A Hulawale, Mr. Abhijit B Deshmukh, Ms.Tejashri B Sahane, Ms. Pooja S Nikam, Prof. Dipali Dube
Department of Computer Engineering.
Vidya Niketan College of Engineering Bota Maharashtra, India

Abstract: *Diabetes is caused when the sugar level increases or the solution which control sugar level i.e. insulin level go down. also it can be caused due to weight gain(i.e. acessive weight) it is long lasting health condition which effect the patient body turn the food to energy/fat .it affect all age people from child to adult everyone and the starting stage symptoms are very negligeous. Later on it may show the symptoms like headache, vision problem and many more and when due to diabetes there is increase in glucose level it affect work of red blood cells (RBC) and also the power of white blood cell is get lower which can make it happens for other disease also diabetes not get to be transmitted from one suffers to another (i.e. non-communication) ,and covid session is going on it more likely get serious for diabetic patient there are many two types of methods for treating diabetes allopathic and medical tasks. We aim to make an affective way for predict diabetes earlier.*

Keywords: machine learning, prediction, support vector machine, random forest, KNN

I. INTRODUCTION

Motivation – the purpose of this project is to develop the system module for prediction of diabetes. The focus of the project is to maintain result accuracy every single time there had been radical and extreme increase in rate of people affected by diabetes since a decade. Our current lifestyle is the main reason for growth in diabetes patient it can not change but we can take treatment ASAP when we test .this system module work on parameter like sugar level, glucose and some more for predicting result and aim to predict result via different methods and machine learning algorithms.

Problem Statement – Diabetes is a persistent disease with prospective to cause a worldwide healthcare crisis. Diabetes is a disease caused due to decrease of insulin level and increase of blood glucose level .thare are different traditional method based on physical and chemical test for the process of determining disease .whenever the early prediction of diabetes is quite difficult, for medical trainee or pharmacist due to various complex interdependence factors as this disease affect organs like kidney, eye, heart etc.

Machine learning methods / algorithms have the potential to illuminate common question the aim of this system is for early prediction of diabetes. This project predict via different approaches, support vector machine, KNN, random forest. We also have aim to purpose fruitful technique for earlier detection of the diabetes disease.

II. LITERATURE SURVEY

In paper first, Diabetes prediction using different machine learning approaches Veena Vijayan V. and Anjali C. they are finding the Diabetes disease how it produced by rising sugar level in the plasma/glucose, various information computerized system were outlined utilizing classifier for diagnosing diabetes and anticipating using the machine learning algorithm – support vector machine (SVM), naïve bayes, decision tree and artificial neural network (ANN).

In paper second MD Kamrul Hasan, MD Ashraful Alam, Dola Das, Eklas Hossain, Mahmudul Hasan.they are published the paper of Diabetes prediction using ensembling of different machine learning classifier. In this paper, we purpose a new pipeline for predicting diabetes from the PIMA Indian diabetic dataset.To ensemble the ML models, soft weighted voting is employed,where the individual model of the weight it was estimated from the respective AUC of the model. they propose a classification of the ML models for boosting the predicting diabetes by ensembling classifier.ML

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

model is chosen as the weight of that model for voting ensembling rather than the accuracy.since AUC is unbiased to the distribution of class

Sivaranjani S,Ananya S,Aravinth J,Karthika R.- Diabetes prediction using machine learning algorithm with feature selection and diamensionality reduction. In recent years plenty of method have been published diabetes prediction. The frame work in machine learning was propose in where the authors implemented the adaboost,random forest,guassian process classification support vector machine linear discriminant analysis quadeatic discriminant analysis ,artificial neural network ,naïve bayes, logistic regression, decision tree with different diamensionality reduction and cross validation.

III. SYSTEM ARCHITECTURE

This section present the listing of proposed/ system architecture for a diabetic patient predicting system. The developed system module measures blood glucose level, sugar and based on age, no of pregnancies etc. the data provided from PIMA Indian diabetic dataset take as input. Then preprocessing through extraction of features from the data will done. We need to train the system over the feature which we need for prediction, then classify input data into parameter like age, sugar, glucose, hoemoglobin and then test the dataset to that according to provide feature and train model machine will be predict the diabetes. The models used are support vector machine, random forest and k-nearest neighbor method. The system architecture help in future for diagnosis diabetes.

It will also predict whether the person is diabetic or not on the basis of using a train dataset.

Steps for prediction

1. Collection of data
2. Data defining
3. Pre-Processing
4. Model-building
5. Result Analysis

Algorithm

- Libraries Importing
- Importing dataset
- Training dataset
- Testing dataset
- Performing algorithms
- Evaluation
- Comparison of results

K-nearest neighbor –

KNN is a simple and supervised machine learning algorithm due to more high accuracy, KNN algorithm is basically for classification and regression purpose. In that data is model which be the reference for future that's why it does not have training period.

Support Vector Machine –

Support vector machine algothm is basically used when there a number of features are high as compared data points in the dataset. There are two types of support vector machine (Svm) like linear, non-linear. Use of support vector machine for clustering i.e. unsupervised learning is now considered in many different ways.

Random Forest –

This algorithm also used for regression and classification. It basically contain decision tree classifier. This algorithm solve the issue of fitting in training set to prefer over decision tree. It majorly used for classification. This classifier collect the majority data for providing final prediction.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

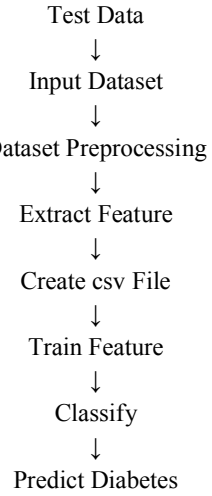

Fig 1- System Architecture

Advantages

- Least cost.
- Easy to use. Saves time.

IV. CONCLUSION

Earlier stage prediction of diabetes is very important as it has negligious System at the starting of disease. The system module is used parameter like sugar level, glucose, age etc and many more to predict the diabetes and here we are only going to predict that the user/patient is diabetic or nondiabetic. this model helps to physician /doctors to start treatment as soon as possible. We used three different machine learning technique for prediction (i.e. support vector machine, random forest, KNN) In which the random forest method has give the more accurate result/closely prediction. Diabetes is rough and and tumble health issue in human society. This paper summarized technique for prediction of this affilication.

REFERENCES

[1]. "Diabetes prediction using machine learning", Bhavya Sanjay Hc[2], Suraj SK[2], Savant Aakash Shivshankar Rao[4], Sanjay M[5],IJARCCE-2020

[2]. Aishwarya muJumdai, Dr – Vaidehi v[b], "Diabetes prediction using machine learning algorithms", on procedia computer science 165(2019) 292-299. (IJARCCE).

[3]. Sivaranjani S, Ananya S, Aaravinth J, Karthika K, "Diabetes prediction using machine learning algorithms with feature selection and diamensionality reduction",IEEE Access.

[4]. Anjali c and Veena Vijayan v, Prediction and Diagnosis Diabetes Mellitus, " A machine learning Approach", 2015 IEEE Recent Advantages in intelligent computational system (RAICS)|10-12 December 2015|Trivandrum.

[5]. Priyanka Sonar and Prof. K Jayamalini prediction several algorithm "Diabetes prediction using Different Machine Learning Approaches", IEEE – 2019

[6]. MD KAMRUL Hasan, MD ASHRAFUL Alam .Dola Das,Eklas Hossain ,MAHMUDUL Hasan, "Diabetes Prediction Using Ensembling of Different Machine Learning Classifier" IEEE Access – 2020/DOI 10119