Type 2 Diabetes Mellitus Prediction and Relevant Anti-Drug Detection

[1]Apoorva M G, [2]Mithilesh Bhagat, [3]Rahul Ranjan, [4]Anmol Patel, [5]Vimuktha E Salis, [5]Associate Professor
[1]Department of Information Science and Engineering,
[1]Global Academy of Technology, Bengaluru, India.

Abstract: Designing of medical decision support system is deceiving and plays an important role in an early prediction of a disease and assists specialists with appropriate medication. Diabetes Mellitus is one of the feebleness diseases which is characterized by increased levels of glucose in blood due to defects in insulin secretion in the body. This results in long term damage and failure of different organs especially heart, kidney, lungs and eyes. Thus, an proper medication and early detection could reduce the risk of these complications. This project aims to make a Data Science Model to Predict Type-2 Diabetes mellitus. India has over 60 million adults with diabetes (7.7% of the population), of which more than 30 million are untreated or undiagnosed, thus increasing the risk of premature morality and developing complications. Objectives of this project are Prevention of complications in diabetic type-2 patients, Reduction of disability and mortality from diabetic complications, Improve in quality of life among diabetic patients and Reduction in expenditures for medical care.

***Index Terms* - Type-2 Diabetes, Machine Learning, Anti-Drug.**

I. INTRODUCTION

Diabetes Mellitus is a chronic illness caused due to reduced production of insulin by pancreas in human body referred to as Diabetes Type-I or the body stops responding to insulin produced Diabetes Type-II which leads to increased sugar level in blood beyond the normal level. In a long run, both these conditions cause severe damage to different organs in the body leading to cardiac arrest, unless controlled by proper medication. Nowadays, people lifestyle is too much busy and most of them do not take care of their health and how to save it. It may cause us to generate many lifestyle diseases; diabetes mellitus, as usual, we can say diabetes is one of the diseases which closely related to our lifestyle. It can be the deadliest disease if it is unidentified. Our body needs energy for working and the main source of energy is blood glucose sourced from the food that you eat. In our body pancreas is an important organ that releases insulin. Insulin is a hormone that plays a vital role in regulating the sugar (glucose) level in the body. Glucose sourced from carbohydrates in the food that human being takes in their diet and responsible for the proper functioning of the body. Insulin maintains the sugar level in the blood such as to prevent the case of too low or too high, both are risky. Diabetics is the multi factorial neurodegenerative disease, which is degrades the vision over time which is responsible for loss vision of 66.8 million people in all over the world. To evaluate the potential risk of the disease is take too much time in diagnose as well as for treatment of the Diabetic disease. The increased pressure of blood is known as intra ocular pressure (IOP) and the progression of the disease increases to dangerous level, and then causes the break optic nerve axon. That's why; in developing Diabetic the risk factor is Increase the pressure in IOP

II. LITERATURE SURVEY

In this survey paper we are discussing various type-2 Diabetes mellitus prediction with various approaches and methodologies.

Gaurav Tripathi et.al [1] proposed a model which aims to develop a model in the health care application using machine learning for predictive analysis of diabetes which uses significant features that are closely related to this disease. Here for the experimental analysis, a Pima Indian Diabetes Database is used. For developing the model four classification algorithms are used in this study namely, Linear Discriminant Analysis (LDA), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF) Algorithm. In this model, the original dataset is prepared into a train-test set to validate the performance. To measure the performance of various classifiers that are used to build model, some important statistical metrics are calculated. The metrics are accuracy, sensitivity, precision, specificity and F-score. The obtained results show that Random Forest (RF) classifier gives a maximum accuracy of 87.66%.This model collects data quickly and accurately.

Shivangi C. Patel et.al [2] proposed a model on image processing technique for the early detection of glaucoma. Here

Glaucoma is detected using retinal fundus image. CDR technique is used on different retinal image for glaucoma detection. Further, analysis of CDR is carried out using different color channel combinations of the fundus images. This model collects data quickly and shows average accuracy.

Sharana Gowda Nawaldgi et.al [3] proposed a model, where a review of automated glaucoma detection techniques are presented. In this model various structural features that are relevant to CFI and OCT images respectively for automated glaucoma detection is used. The model concludes that combining structural features from both CFI and OCT images would result in more accurate glaucoma assessment. This model gives more accuracy than the model which has color Fundus Images.

Priyadarshini et.al [4] proposed a system which presents a simple yet efficient fuzzy bases expert method. The expert system was developed using PIMA Indian Diabetes Database (PIDD). It comprises of 240 rules based on feedback given by specialist in the medical domain. The outcome of the observation in PIDD is either a '0' indicating negative diagnosis and '1' indicating positive diagnosis. Parameters such as glucose level, insulin level, body mass index, diabetes pedigree function and age were recorded as inputs from each patient to calculate probability of occurrence of diabetes. These test results were converted to fuzzy data. Fuzzy logic MATLAB toolbox was used to design the fuzzy inference system. The results obtained showed that ANFIS classification model had better accuracy when compared to that of MLP classification model.

Asif Hassan Syed et.al [5] proposed a mechanism which follows cross-sectional survey in which the researcher measures the exposure in the population, the outcome and study their relationship concurrently. Here the dataset for the preparation of the "T2DM risk predictor system" was built using a cross sectional survey. In this model they have employed the Pearson Chi-squared statistical test to assess the alternative hypothesis. To demonstrate the calculation and analysis of the x2 statistic. They have used stating of Hypothesis. The idea of the Chi-Square test, Measure how different the observed count is different from the expected count steps are implemented. Presently this model can be used by the Physician for assessing the risk of diabetes among Saudi's and expatriates residing in the western province of the Kingdom of Saudi Arabia.

Aishwarya Jaka et.al [6] proposed a model which made a comparative analysis of performance of KNN, decision tree, Naïve Bayes, support vector machine, logistic regression and random forest classifiers using PIMA Indian Diabetes Dataset. This finds the best accuracy level of 77.6% from logistic Regression while KNN classifiers shown 73.43% accuracy.

Subhash Chandra Gupta et.al [7] proposed a model which uses K-Nearest Neighbors Algorithm (KNN) and machine learning methods are used in the prediction model to classify whether the patient is diabetic or non-diabetic. Here the accuracy is 87.01% and error rate is 12.99% respectively. This model can determine type-1 and type-2 Diabetes easily.

Hala Alshamlan et.al [8] proposed a system which predicts diabetic and normal persons by using fisher score feature selection, chi-2 feature selection and Logistic Regression supervised learning algorithm. In this model the obtained accuracy result of logistic regression on two datasets based on fisher score feature selection is higher than ch-2 feature selection. The accuracy results of two data are 90.23% and 61.90% respectively.

Rahul Pradhan et.al [9] proposed a model which has five steps to predict the diabetes. They are data collection, data preprocessing, implementation of various model, evolution of models and prediction of diabetes. This process has used SVM algorithm which provides highest accuracy. Here the mechanism of drugs and therapies are discussed and the general regressive neural network approach is also discussed.

Mamta Arora et.al [10] proposed a system which works in the area on binary classification of diabetic Retinopathy and achieve the accuracy of 88.4% and the work was done with data set of 140 images and considered the result for highest probability for output.

III. COMPARISON

Here we subtly discuss and compare the previously surveyed paper. We also give the methodology and contribution they had.

Table provides the comparison of the survey on Type-2 Diabetes mellitus.

Authors	Methodology	Contribution
Gaurav Tripathi et.al [1]	This model demonstrates the development of convolutional neural network that takes image as input and predicts the diabetic retinopathy. It also describes the development and implementation of ConvNet based algorithm	This model shows the description of the dataset and presents the precision and sensitivity (recall) outcome.
Shivangi C Patel et.al [2]	Image processing technique is used for the early detection of glaucoma. Here Glaucoma is detected using retinal fundus image. CDR technique is used for different retinal image for glaucoma detection. Then, analysis on CDR is	This model collects data quickly and shows average accuracy.

	carried out using different color channel combinations of the fundus images.	
Sharana Gowda Nawaldi et.al [3]	Here various structural features that are relevant to CFI and OCT images for automated glaucoma detection is used. The model concludes that combining structural features from both CFI and OCT images would result in more accurate glaucoma assessment.	This model gives more accuracy.
Priyadarshini et.al [4]	The outcome of this model is categorized into five groups i.e. very low, low, medium, high, very high based on the likelihood and severity od diabetes. The fuzzy model is cost effective and reduce complexity.	Display results systematically.
Asif Hassan Syed et.al [5]	In this model, the dataset for the preparation of the "T2DM risk predictor system" was built using a cross-sectional survey. In this model they have employed the Pearson Chi-squared statistical test to assess the alternate hypothesis.	This model can be used by the Physician for assessing the risk of diabetes among Saudi's and expatriates residing in the western province of the Kingdom of Saudi Arabia.
Aishwarya Jaka et.al [6]	This model makes a comparative analysis of performance analysis of KNN, decision tree, Naïve Bayes, support vector machine, logistic regression and random forest classifiers using PIMA Indian Diabetes Dataset.	It gives the best accuracy level of 77.6% from logistic Regression while KNN classifiers shown 73.43% accuracy.
Subhash Chandra Gupta et.al [7]	This model uses both KNN and SVM algorithm, and the accuracy for KNN is 87.01% and SVM is 88.56%. They have used the dataset which is developed by a model in WEKA	This model can determine type 1 and type 2 Diabetes easily.
Hala Alshamlan et.al [8]	Here the model uses the gene expression with classification data sets to predict the disease. Prediction rate is 90.23%	Worked on two datasets hence accuracy is more
Rahul Pradhan et.al [9]	This method has five steps to predict the diabetic patients. They are Data Collection, Data Preprocessing, Implementation of various model, Evaluation of models, Prediction of diabetes. This process has used SVM algorithm which provides highest accuracy.	Mechanism of drugs and therapies are discussed.
Mamta Arora et.al [10]	This model works in the area of binary classification of diabetic retinopathy and achieve the accuracy of 88.4%.	It is found that the model can be used on the top of pre trained model which can provide a substantial boost to the results.

IV. DATA COLLECTION

We collected our data through an online website known as Kaggle. It is renowned website known to provide datasets on various topics. The dataset included retina pictures of both diabetic and non - diabetic patients.

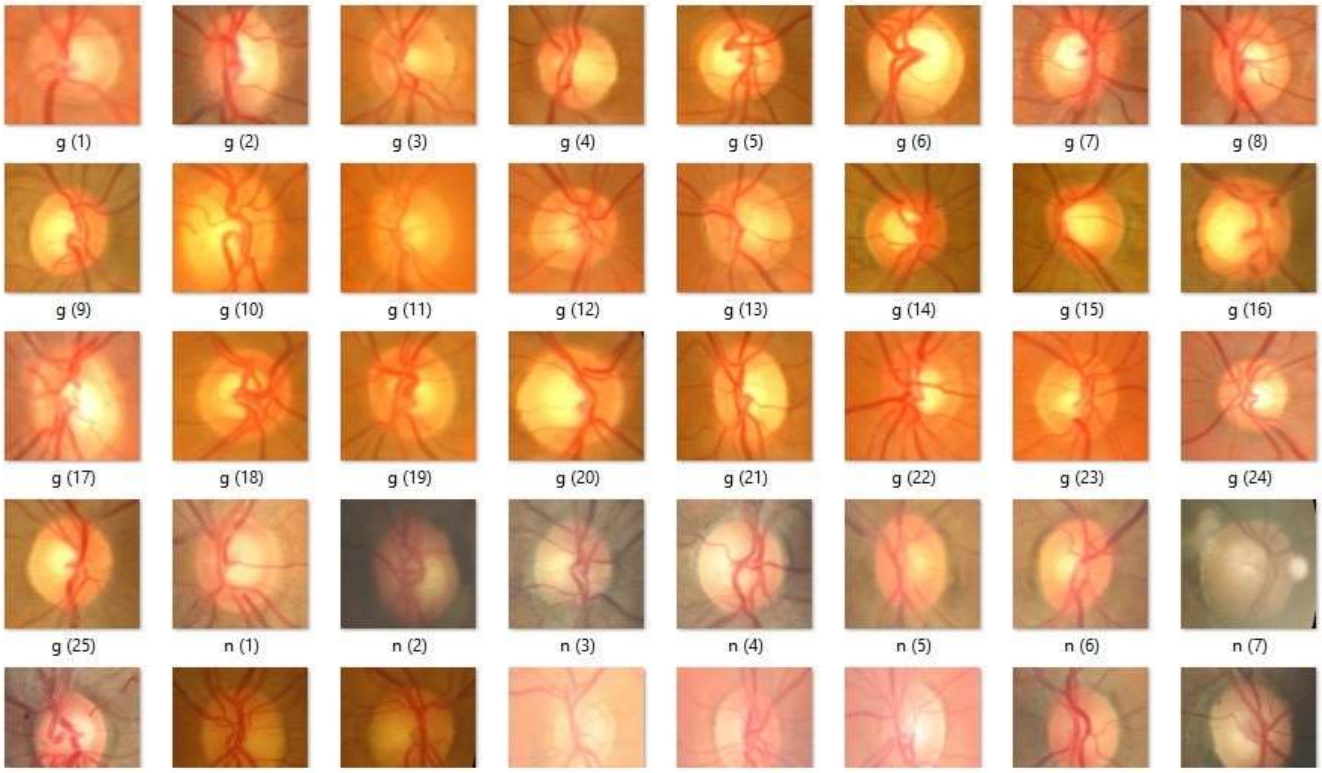

Fig 4.1: Datasets of retina images

V. APPLICATIONS ON TYPE-2 DIABETES MELLITUS PREDICTION

In Real world doctors can use our model to classify the people who are in the risk of getting diabetes. It also tells the anti-Relevant drug selection.

VI. CONCLUSION

In this model, we have offered a novel convolutional neural organization strategy for diabetic retinopathy and discovery utilizing transfer learning approach The method consists of preprocessing including automatic extraction of retinal patches for improvement in results. Furthermore, proposed technique outperformed for various CNN models including VGG16 for DR classification. Experiments show that a inception CNN model shows a better classification performance, also succeeded retinal pathological features are represented. The best results achieved for retinal DR with an accuracy of 96.29% on Messidor database.

In future work we will investigate some of the advanced features for multi - modality of the retinal images.

REFERENCES

[1]Gaurav Tripathi, Rakesh Kumar "Early prediction of diabetes mellitus using machine learning" @ ICRITO volume:5 issue:03 June 2020.

[2]Shivangi C Patel, Manish I. Patel "Analysis of cdr of fungus images for glaucoma detection" @ICoEI vol :4 issue:06 jan 2018.

[3]Sharana Gowda Nawaldg "Review of automated glaucoma detection techniques" @IEEE volume: 3 issue: 05 2018.

[4]Priyadarshini L, Lakshmi Shrinivasan, "Design of an ANFIS based Decision Support System Diabetes Diagnosis" @IEEE volume:05 issue:03 july 2020.

[5]Asif Hassan Syed, Tabrej Khan "Machine Learning Based Application for Predicting Risk T - 2 Diabetes Mellitus (T2DM) in Saudi Arabia: A Retrospective cross - sectional study" @IEEE volume:05 issue: 02 oct 2020.

[6]Aishwarya Jaka and V.R. J, Performance Evaluation of Machine Learning Model for Diabetes Prediction @Int.Journal of Innov. Technol. Exloring Eng, vol: 8, issue:11 2020.

[7]Subhash Chandra Gupta, Noopur G, "Performance enhancement of diabetes prediction by finding optimum K for KNN classifier with feature selection meathod "@ ICSSIT volume: 2 issue: 04 2020.

[8]Hala Alshalman, Hind Bin Taleb, Areej AI Sahow, Prediction Function for Type 2 Diabetes Mellitus using Logistic Regression" @ICICS volume:03 issue:05 2020.

[9]Rahul Pradhan, Mayank Agarwal, Drashti Maheshwari, Ankur Chaturvedi, Dilip Kumar Sharma, "Diabetes Mellitus Prediction and Classifier Comparative Study" @PARC volume:2 issue: 03 2020.

[10]Mamta Arora, Mrinal Pandey, "Deep neural network for diabetic retinopathy detection" @IEEE volume:03 issue:05 feb 2019.