DIABETES PREDICTION USING MACHINE LEARNING

[1]MAYANK GUPTA, [2]PRINCE KARAVADIYA, [3]SAKSHI GORE, [4]SHUBHAM UBALE,
[5]PROF. KANCHAN DHOMSE

B. Tech, Department of IT,
MET's Bhujbal Knowledge City, Adgaon, Nashik, Maharashtra, India.

Abstract: Recently diabetes is a disease caused by a group of metabolic disorders. It is also known as Diabetic mellitus. It affects the human body organs. Diabetes can be controlled by predicting this disease earlier. If diabetics patient is untreated for a long time, it may lead to high level of blood sugar. Now a days, healthcare industries generating large volume of data. Different Machine Learning algorithms and statistics are used to predict the disease with the help of past and current data. Machine learning techniques helps the doctors to predict diabetes at early stage. Patient's medical record and different types of algorithms are added in dataset for experimental analysis. we use logistic regression, random forest, decision tree classifier and gradient boosting to predict whether a patient based on diagnostic measurements has diabetes or not. So, in applied algorithm performance and accuracy is compared and discussed in the project.

Keywords: CNN, FCM, Medical Image, segmentation, SVM

INTRODUCTION

Diabetes is what is going on which makes lack due less measure of insulin in the blood. Cautioning indication of high glucose brings about regular pee, feeling parched, expanded hunger. Extreme troubles lead to cardiovascular sickness foot bruises, and eye fogginess. At the point when there is an ascent inside the sugar level inside the blood, it is alluded to as earlier diabetes.

LITURATURE SURVEY

1.Diabetic Patient Prediction using Machine Learning Algorithm, Malini M etal.,[1]**.**
This paper studied that; Diabetes is a most common disease caused by a group of metabolic disorders. It is also known as Diabetic mellitus. It affects the organs of the human body. It can be controlled by predicting this disease earlier. If diabetics patient is untreated for a long time, it may lead to increase blood sugar. Now a days, Healthcare industries generating large volume of data. Machine Learning algorithms and statistics are used to predict the disease with the help of current and past data. Machine learning techniques helps the doctors to predict early stage for diabetics. Diabetics patient medical record and different types of algorithms are added in dataset for experimental analysis. we use logistic regression, random forest, decision tree classifier and gradient boosting to predict whether a patient has diabetes based on diagnostic measurements. Performance and accuracy of the applied algorithms is discussed and compared.

2.Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms, Lifeng Qiao etal.,[2]
This paper presents the Prognosis of Microaneurysm and early diagnostics system for non-proliferative diabetic retinopathy (PMNPDR) capable of effectively creating DCNNs for the semantic segmentation of fundus images which can improve NPDR detection efficiency and accuracy. An easy yet efficient integrated lesion identification system, coupled with LOG and MF filters accompanied by post-processing procedures, is suggested. Combined sequentially and smartly, these techniques provide a very effective system for the identification of various lesions regardless of their texture, form, scale, etc. Transforming Curvelets is a very good candidate for better dark lesions. The BPF is optimally designed for the improvement of bright lesions. Through MSSIM maximization, the BPF's gain and cutting frequencies are automatically achieved. Data for non-MAs vary widely, the collection of non-microaneurysm training is quite a topic. The huge training set not only takes time and causes class imbalance. In this paper, a sparse Principal Component Analysis based unregulated classification approach for detecting microaneurysm was developed. Once a model that represents MA has been developed, any deviating from the standard MA is detected by statistical monitoring, a scarce Principal Component Analysis is employed to find the latent structure of microaneurysm data.

3. Disease Influence Measure Based Diabetic Prediction with Medical Data Set Using Data Mining, B.V. Baijuetal.,[3] That's what this paper shows, the issue of diabetic forecast has been all around concentrated on in this paper. The infection forecasts have been investigated utilizing different techniques for information mining. The utilization of clinical informational index on the forecast of diabetic mellitus has been investigated. This paper plays out a nitty gritty study on illness expectation utilizing information mining approaches in light of diabetic informational index. The presence of sickness has been distinguished utilizing the presence of different side effects. Nonetheless, the techniques utilize various elements and produces shifting precision. The consequence of expectation varies with the strategies/measures/highlights being utilized. Towards diabetic forecast, an Infection Impact Measure (Faint) based diabetic expectation has been introduced. The technique preprocesses the info informational collection and eliminates the uproarious records. In the subsequent stage, the technique gauges illness impact measure (Faint) in view of the highlights of information data of interest. In light of the Faint worth, the technique performs diabetic expectation. Various methodologies of illness expectation have been thought of and their presentation in sickness forecast has been looked at. The examination result has been introduced exhaustively towards the turn of events.

4.Implementation of Diabetic Retinopathy Prediction System using Data Mining, Siddharekh S. Patil etal.,[4] This paper concentrated on that Diabetic retinopathy (DR) is the most widely recognized reason for recently analyzed visual deficiency consistently. Yearly eye checking for diabetic patients is recommended to find and treat DR promptly, since visual impairment from this condition is preventable with early recognizable proof. DR recognition is exclusively founded on existing patient records. Presently a day's clinical information developing massively and we really want to handle that information for identification. Yet, it is tedious subsequently information mining procedures assists with getting freed from this issue. We utilize brain organization (NN) and innocent bayes for grouping. As per examination results NN gives preferred exactness over innocent bayes and time and memory expected for NN is less when contrasted with guileless bayes

METHODOLOGY

- Further developed J48 Grouping Calculation for the Prediction of Diabetes.
- This work manages productive information digging methodology for anticipating the diabetes from clinical records of patients. Diabetes is an extremely normal illness nowadays in all populaces and in all age gatherings. Diabetes adds to coronary illness, builds the dangers of creating kidney infection, nerve harm, vein harm and visual impairment.
- In this way, mining the diabetes information in proficient way is a basic issue. The Pima Indians Diabetes Informational collection is utilized in this task; which gathers the data of patients with and without having diabetes. The altered J48 classifier is utilized to expand the precision pace of the information mining technique.
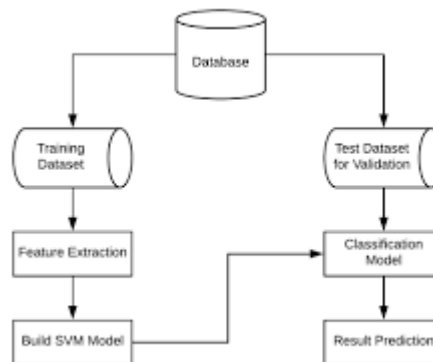

Fig -1: **Flowchart**

OBJECTIVES

- To predict the diabetes at early stage.
- To Avoid the Time-consuming task of feature extraction.
- To increase the accuracy from 70-75% to 80-85 %.
- To provide a cost-effective solution in market.

MOTIVATION

Diabetes is arising as a transcendent sickness in the emerging nations like India. The sickness is turning out to be intense and create numerous different issues in the human body. Many variables are staying as a reason for this sickness in human body. The sickness isn't treatable and must be controlled.
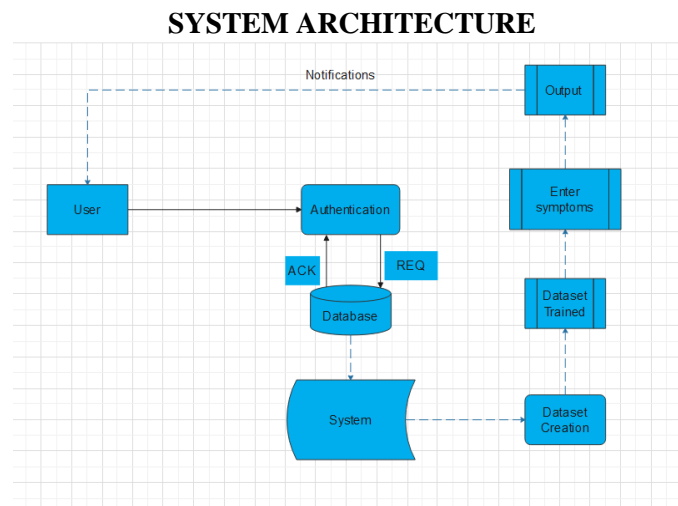
SYSTEM ARCHITECTURE

Fig -2: **System Architecture Diagram**

APPLICATION:

[1] Hospital**:** - In emergency clinics this can be utilized. As we realize that the conventional strategy for diabetes expectation is very tedious. At the point when we go over any side effect, we want to test for it.

[2**]** Personal: - Going to Specialist just to show the report can be a brief period consuming for the patient. Assuming we utilize this framework for the individual use than we can check whether we have diabetes or not. This can be utilized at home additionally by simply entering the information from our reports.

FUNCTIONAL AND NON-FUNCTIONAL REQUIREMENTS

Functional Requirements: It might include estimations, specialized subtleties, information control and handling and other explicit usefulness that characterize what a framework should achieve. Social prerequisites depict every one of the situations where the framework utilizes the practical necessities; these are caught being used cases.

Nonfunctional Requirements: (NFRs) characterize framework credits like security, dependability, execution, practicality, versatility, and ease of use. They act as imperatives or limitations on the plan of the framework across the various overabundances.

SYSTEM REQUIREMENTS

[1] Hardware:

In this project we will require 3GB of RAM or more, likewise Hard Drive of 250 GB or over, extra to this we would utilize the Intel i3 Processor or above.

[2] Technology:

In addition to using Python 3.7 and MySQL 3.2 or greater, this project also calls for the use of Windows Operating System 7.

[3] Tools:

The necessary tools are web browser, notepad++ or VS Code for coding, PyCharm or Jupyter for running the code, Kaggle for the dataset.

CONCLUSION

Thus, we are beating the downside of existing framework, and giving a savvy framework that won't just screen and control our information with security yet in addition supply it too at whatever point vital. We are attempting accomplished over 85% identification exactness.

REFERENCES

1. KM Jyoti Rani, "Diabetes Prediction Using Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, Volume 6 Issue 4, pp. 294-305, July-August 2020. Available at Doi: https://doi.org/10.32628/CSEIT206463 Journal URL: http://ijsrcseit.com/CSEIT206463

2. Dhomse Kanchan B., M.K.M., 2016. Study of Machine Learning Algorithms for Special Disease Prediction using Principal of Component Analysis, in: 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication, IEEE. pp. 5–10

3. Aljumah, A.A., Ahamad, M.G., Siddiqui, M.K., 2013. Application of data mining: Diabetes health care in young and old patients. Journal of King Saud University - Computer and Information Sciences 25, 127–136. Doi: 10.1016/j.jksuci.2012.10.003.

4. Choubey, D.K., Paul, S., Kumar, S., Kumar, S., 2017. Classification of Pima Indian diabetes dataset using naive bayes with genetic algorithm as an attribute selection, in: Communication and Computing Systems: Proceedings of the International Conference on Communication and Computing System (ICCCS 2016), pp. 451– 455

5. Bamnote, M.P., G.R., 2014. Design of Classifier for Detection of Diabetes Mellitus Using Genetic Programming. Advances in Intelligent Systems and Computing 1, 763–770. doi:10.1007/978-3- 319-11933-5.

6. [9]. https://www.kaggle.com/johndasilva/diabetes