IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Predictive Modeling of Diabetic Retinopathy using Classification based Machine Learning Techniques – A Review of Literature

1Sandeep Prakash, 2Pankaj Prajapati

1M Tech Scholar, Industrial Automation and Robotics , 2Head Of Department, Industrial Automation and Robotics

1Ambalika Institute of Management and Technology, Dr A.P.J Abdul Kalam Technical University,Lucknow,

2Ambalika Institute of Management and Technology, Dr A.P.J Abdul Kalam Technical University,Lucknow

Abstract— According to the World Health Organization (WHO), the seventh major cause of human death in 2030 will be diabetes, which of course is a very severe disease and if not treated thoroughly and on time, can lead to critical problems, including death. Millions of people are affected by the disease. The risk of diabetes is increasing day by day and is found mostly in women than men. The diagnosis of diabetes is a tedious process. Diabetic retinopathy is a disease caused by uncontrolled chronic diabetes and it can cause complete blindness if not timely treated. Therefore early medical diagnosis of diabetic retinopathy and it medical cure is essential to prevent the severe side effects of diabetic retinopathy. Manual detection of diabetic retinopathy by ophthalmologist take plenty of time and patients need to suffer a lot at this time. So with improvement in science and technology it is made easy to predict the disease. The purpose is to diagnose whether the person is affected by diabetes or not using classification based machine learning technique. This paper reviews, classifies and compares the algorithms and techniques previously proposed in order to develop better and more effective algorithms.

Index—WHO, Diabetic Retinopathy, Classification based Machine Learning techniques, Algorithms

Introduction:-

One of the major problem diabetic patients suffers from is the Diabetic Retinopathy and blindness. The greatest challenge to current health care is the rapid growth of diabetes. If detected at early stage, laser therapy can be performed to prevent or delay visual loss and may be used to encourage improvement in diabetic control. At the higher stage of diabetes, it is represented by bleeding or the accumulation of fluid in the retina. However, the symptoms can only be noticed by the diabetic patients when they start suffering from some sight disorders. If the blood vessels of the retina are damaged, it results in Diabetic Retinopathy (DR). It occurs when high blood glucose, the characteristic of diabetes, has damaged the small vessels that provide oxygen and nutrients to the retina.

Diabetic Retinopathy (DR) is the leading cause of vision loss in adults aged 20–74 years. From 1990–2010, DR ranked as the fifth most common cause of preventable blindness and fifth most common cause of moderate to severe visual impairment. In 2010, of an estimated 285 million people worldwide with diabetes,

over one-third have signs of DR, and a third of these are afflicted with vision-threatening diabetic retinopathy (VTDR). Every year an estimated 65–70 new cases of blindness per 100 000 occurs. These estimates are expected to rise further due to the increasing prevalence of diabetes, ageing of the population and increasing of life expectancy of those with diabetes. Recent report presented that Diabetic retinopathy is responsible for 4.8% of the 37 million cases of blindness due to eye diseases throughout the world. After 15 years, about 2% of persons with diabetes will become blind, and about 10% will develop severe visual loss. After 20 years, more than 75% of patients will have some form of DR.

Patient perception of complications : A survey says that 84% of diabetic patients knew about the complications of diabetes and 73% knew about the eye complications with 41% knowing that diabetes can lead to vision loss. Indeed, blindness was the complication patients feared the most. Vision loss was the complication feared most by the clients interviewed.

Diabetic Retinopathy is divided into four stages by Ophthalmologists. These stages are:

- First stage is a normal retina that is free from any abnormalities.
- The second stage is Mild (NPDR - non proliferative DR); that is, small red coloured swelling spots shown on the walls of the retina called Microaneurysm.
- Third stage is Moderate (NPDR) where Microaneurysm begins to rupture and appears with small or medium-sized blood on the surface of the retina.
- Fourth stage is Exudates which is called the severe (NPDR) level. exudates are of two types: Hard exudates and Soft exudates Hard exudates have well defined boundaries. The hard exudates

  I. are displayed due to the fragile and weak walls of the blood vessels of the retina. Such weak walls lead to liquid protein leakage from the blood vessels to the surface of the retina.

  II. Soft exudates have unclear boundaries of a whitish colour. These exudates cause blockages in the blood vessels of the retina, and prevent the arrival of food and oxygen for its tissue. They further lead to the emergence of the so-called (Neovascularisation), winding new blood vessels;

Following Figure gives a clear picture of these four stages of DR

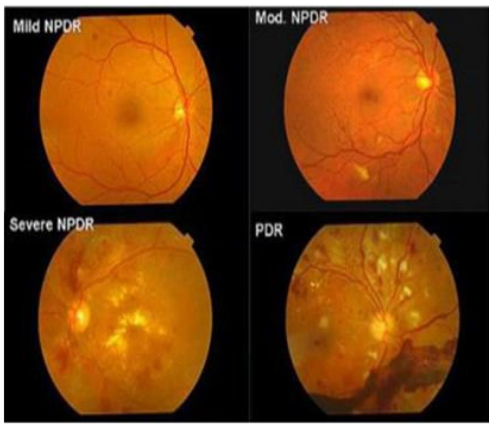

Figure 2: Diabetic Retinopathy Stages

Machine Learning

Machine learning, a branch of artificial intelligence, concerns the construction and study of systems that can learn from data **[1].** Machine learning algorithms use computational methods to "learn" information directly from data without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases. Tom M. Mitchell provided a widely quoted and more formal definition:

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E **[2].**

The core of machine learning deals with representation and generalization. Representing the data instances and functions evaluated on these instances are part of all machine learning systems. Generalization is the ability of a machine learning system to perform accurately on new, unseen data instances after having experienced a learning data instance. The training examples come from some generally unknown probability distribution and the learner has to build a general model about this space that enables it to produce sufficiently accurate predictions in new cases. The performance of generalization is 11 usually evaluated with respect to the ability to reproduce known knowledge from newer examples. There are different types of machine learning, but the two main ones are:

• Supervised Learning

• Unsupervised Learning

Supervised Learning Model

Supervised learning is the machine learning task of inferring a function from supervised training data **[3].** Training data for supervised learning includes a set of examples with paired input subjects and desired output. A supervised learning algorithm analyses the training data and produces an inferred function, which is called classifier or a regression function. The function should predict the correct output value for any valid input object. This requires the learning algorithm to generalize from the training data to unseen situations in a reasonable way.

Related Works and Research

Diabetic retinopathy is the leading cause of blindness in the working-age population of the developed world. Since 1982, the quantification of diabetic retinopathy and detection of features such as exudates and blood vessels on fundus images were studied. A lot of work has been done in this field. Before starting implementation of main task we go through similar paper to know about the whole system such as what are the things we need to consider in order to detect diabetic retinopathy. AkaraS., Matthew N. Dailey has proposed a "Machine learning approach to automatic exudate detection in retinal images from diabetic patients**"[4].** In their paper they presented a series of experiments on feature selection and exudates classification using K- nearest Neighbor (KNN) and support vector machine (SVM) classifiers.

Rajendra Acharya U.,E. Y. K. Ng, Kwan-Hoong Ng and Jasjit S. Suri introduced algorithms for the automated detection of diabetic retinopathy using digital fundus images **[5]** where they improved an

algorithm used for extraction of some features from digital fundus images. Moreover, Varun G. and Lily P. has used deep learning for detection of diabetic retinopathy **[6].**

In "Diagnosis of Diabetic Retinopathy using Machine Learning" research paper S. Gupta and K. AM tried to detect retinal micro-aneurysms and exudates retinal funds from images **[7].** After pre-processing, morphological operations are performed to find the feature and the features are get extracted such as GLCM and splat for classification. They achieved the sensitivity and specificity of 87% and 100% respectively with accuracy of 86%.

Tiago T.G. in his paper "Machine Learning on the Diabetic Retinopathy Debrecen Dataset" has used R language for predicting diabetic retinopathy **[8].** He used a dataset in which the features were extracted from images of the eye of a diabetic patient. In his work he used eight different classification algorithms and also shown some comparisons. He achieved 78% accuracy from his work.

[9] Diabetes is a sickness that takes place whilst glucose content in your blood is too excessive. Insulin is a hormone made through the pancreas, facilitates to separate glucose from meals get into your body- cells for power. On this they used category set of rules techniques of the device mastering on those who are expecting the diabetes. Five machine getting to know algorithms namely SVM, Decision Tree, NaiveBayes , Logistic Regression and KNN are used to hit upon diabetes.

[10] Concluded that the best method of prediction of diabetics is Random Forest. This method gives us an approximate result after the splitting and analysis of the training and testing data. The efficiency of this method is much better compared to that of Naïve Bayes. The analysis done from the PIMA dataset, the aim of splitting the dataset is to find the highest/best accuracy of the algorithms and how they would respond if the data split is varied

[11] In this paper, to predict the persons whether diabetic or not. In this paper classification techniques such as Binary Logistic Regression, Multilayer Perceptron and K-Nearest Neighbor are classified for diabetes data and classification accuracy were compared for classifying data.This work focused the implementation of Binary Logistic Regression, Multilayer Perceptron and k-Nearest Neighbor for the diabetes data. From the analysis, it is examined that the formation of classifications will be different for classification methods. From the histogram, it is seen that the Binary Logistic Regression accuracy is 0.69, Multilayer Perceptron accuracy is 0.71 and KNN gives the accuracy of 0.80.k-Nearest Neighbor is higher than the accuracy of Binary Logistic Regression and Multilayer Perceptron.

[12] This paper explores the early prediction of diabetes using various data mining techniques. The dataset has taken 768 instances from PIMA Indian Dataset to determine the accuracy of the data mining techniques in prediction. The analysis proves that Modified J48 Classifier provide the highest accuracy than other techniques. Modified J48 Classifier gives 99.87% of highest accuracy using WEKA & MATLAB tool.

[13] This paper aims at finding solutions to diagnose the disease by analyzing the patterns found in the data through classification analysis by employing Decision Tree and Naïve Bayes algorithms. The research hopes to propose a quicker and more efficient technique of diagnosing the disease, leading to timely treatment of the patients. It shows how Decision Trees and Naïve Bayes are used to model actual diagnosis of diabetes for local and systematic treatment, along with presenting related work in the field. Experimental results show the effectiveness of the proposed model. The performance of the techniques was investigated for the diabetes diagnosis problem. Experimental results demonstrate the adequacy of the proposed model. Those are some related paper of our topic from where we took knowledge and idea to develop new version. In our work we will use different machine learning classification algorithms to classify diabetic retinopathy.

[14] designed and implemented a system to measure the impact of diabetic retinopathy using data mining techniques In this research, an accurate measurement system of diabetic retinopathy was developed and investigated using data mining technique in which an early analysis of diabetic retinopathy using an accurate and fast technique provides the patient with enough protection treatment time. The color fundus image was used to automatically detect and realize the various lesions of diabetic retinopathy and its normal features,

[15] proposed a system to identify the input image as normal or abnormal. When the input image is found abnormal then analysis for further DR stages is done. To identify abnormal images there are various techniques and methodology used in image mining. Image mining is an extension of data mining technique. Identification of the abnormal images is done using preprocessing, feature extraction and classification algorithms. The obtained result is used to display the image as normal or abnormal and upto what extent.

[16] used Naive Bayes and Support Vercto Machine algorithms to predict early detection of eye disease and DR. Using Rapid Miner tool they have estimated that Naive Bayes gives 83.37% accuracy and SVM gives 64.91% accuracy. Performance of these methods was also measured by specificity as 95% and sensitivity as 96.65%.

[17] performed the diagnosis of DR using morphological process and SVM classifier.Then they have applied erosion operation followed by dilation for exudates feature detection and then segmentation operation is carried out. Severe risk assessed for the degree of abnormality of an image using machine learning classifier. SVM is used to evaluate training data to find a best way to classify images into different cases like moderate or severe. In this paper, the evaluation of the automated diagnosis system of diabetic retinopathy has been performed by using a set of 5 images captured by retinal fundus cameras.

[19] proposed an automated eye screening system based on multi feature extraction which can detect Diabetic retinopathy based on the exudates extraction. Combination of DWT and GLCM features is used for evaluation using sensitivity, specificity parameters. This proposed system has an improved pre-processing stage which will eliminate the noise completely and thus improve overall efficiency of the system.

[20] presented a review paper on Diagnosis of Diabetic Retinopathy using KNN and SVM Algorithms. In this paper authors has mainly focused on automatic detection of Diabetic Retinopathy through detecting exudates in color fundus retinal images and also classify the lesions. they have also discussed on various methods available for detecting the exudates.

Conclusion

Automated DR Disease detection systems remarkably reduce the diagnosis time, cost, and helps as an assistant for ophthalmologists in detecting retinal abnormalities and providing timely treatment. These automated systems play a crucial role in detecting diseases more accurately. Some studies have suggested that while integrating the classification based machine learning techniques, the performance has been increased. Initially, in this paper, a brief description was given on Diabetic Retinopathy risk factors and classifications along with brief theory about classification based machine learning techniques. It is seen that the prediction and diagnosis of Diabetic Retinopathy depends on the presence of haemorrhages and microaneurysms in fundus images. There are many algorithms which have been proposed and developed for the automatic detection of diabetic retinopathy from feature extraction. Here a summarized view of various techniques is presented. This review paper discusses the recent works, and the most useful techniques are put forth, which helps the research community in detecting and classifying DR. It also acts as a resource for the future researchers for the prediction of diabetic retinopathy using data mining techniques. This will also be useful for the researchers to get an outline of this area in order to develop more efficient algorithms.

References

1. BoserB ,Guyon I.G,Vapnik V., "A Training Algorithm for Optimal Margin Classifiers", Proc. Fifth Ann. Workshop Computational Learning Theory,pp. 144-152, 1992.
2. Mitchell, T. (1997). Machine Learning, McGraw Hill. ISBN 0-07-042807-7., McGraw-Hill, Inc. New York, NY, USA. Published on March 1, 1997
3. Alex C, Boston A. (2016).Artificial Intelligence, Deep Learning, and Neural Networks, Explained (16:n37)
4. Varun G., Lily P., Mark C., "Development and validation of a deep learning Algorithm for Detection of Diabetic Retinopathy", December 2016.
5. Tiago T.G. "Machine Learning on the Diabetic Retinopathy Debrecen Dataset", knowledge-Based System60, 20-27. Published on June 25, 2016.
6. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care. 2012;35:556–64
7. Boser B. E, Guyon I. M. and Vapnik V. N. (1992). "A training algorithm for optimal margin classiers".Proceedings of the 5th Annual Workshop on Computational Learning Theory COLT'92, 152 Pittsburgh, PA, USA. ACM Press, July 1992. On Page(s): 144-152

8. Leske MC, Wu SY, Nemesure B, Hennis A Barbados Eye Studies Group. Causes of visual loss and their risk factors: An incidence summary from the Barbados Eye Studies. Rev PanamSaludPublica.2010;27:259–67
9. Naveen Kishore G, V.Rajesh, A.Vamsi Akki Reddy, K.Sumedh, T.Rajesh Sai Reddy , "Prediction Of Diabetes Using Machine Learning Classification Algorithms", International Journal of Scientific and Technological Research, Vol. 9, Issue 01, Jan, 2020
10. G. Geetha, K.Mohana Prasad, "Prediction of Diabetics using Machine Learning", International Journal of Recent Technology and Engineering (IJRTE), Volume-8 Issue-5, January 2020.
11. S.Selvakumar, K.Senthamarai Kannan and S.GothaiNachiyar, "Prediction of Diabetes Diagnosis Using Classification Based Data Mining Techniques", International Journal of Statistics and Systems ISSN 0973-2675 Volume 12, Number 2 (2017), pp. 183-188
12. [20] Mr. R. Sengamuthu1, Mrs. R. Abirami, Mr. D. Karthik, "Various Data Mining Techniques Analysis to Predict Diabetes Mellitus", International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 05 | May-2018
13. Aiswarya Iyer, S. Jeyalatha and Ronak Sumbaly, " Diagnosis of Diabetes using Classification Mining Techniques", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.5, No.1, January 2015.
14. [1] Dr. Karim Hashim Al-Saedi, Dr. Razi Jabur Al-Azawi, Rasha Asaad Kamil, - Design and Implementation System to Measure the Impact of Diabetic Retinopathy Using Data Mining Techniques, International Journal of Innovative Research in Electronics and Communications (IJIREC) Volume 4, Issue 1, 2017, PP 1-6
15. [2] Abhilash Bhaisare, Sagar Lachure, Amol Bhagat, Jaykumar Lachure - Diabetic Retinopathy Diagnosis Using Image Mining, International Research Journal of Engineering and Technology (IRJET), Volume: 03, Issue: 10, Oct -2016
16. [3] K. R. Ananthapadmanaban and G. Parthiban. - Prediction of Chances - Diabetic Retinopathy using Data Mining Classification Techniques. Indian Journal of Science and Technology, Vol 7(10), 1498–1503, October 2014
17. [4] MahendranGandhi and Dr. R. Dhanasekaran - Diagnosis of Diabetic Retinopathy Using Morphological Process and SVM Classifier, International conference on Communication and Signal Processing, April 3-5, 2013, India.
18. [5] Karkhanis Apurva Anant , Tushar Ghorpade, Vimla Jethani - Diabetic Retinopathy Analysis using Image Mining to Detect Type 2 Diabetes, International Journal of Computer & Mathematical Sciences IJCMS ISSN 2347 – 8527 Volume 5, Issue 1 January 2016
19. [9] Shraddha Jalan, A. A. Tayade - Review paper on Diagnosis of Diabetic Retinopathy using KNN and SVM Algorithms, International Journal of Advance Research in Computer Science and Management Studies, Vol.3, Issue 1, January 2015.