Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2016, 3(12) :56-62

Research Article ISSN: 2394 - 658X

Early Glaucoma Detection and Prediction with Perimetry Visual Field using Classification and Regression Trees

Adeleke DS[1], Oladosu JB[1], Iwayemi A[2] and Olawoye OO[3]

[1]*Department of Computer Science, University of Ibadan, Ibadan, Oyo State,Nigeria*
[2]*Department of Computer Engineering, The Federal Polytechnic, Ile - Oluji, Ondo State, Nigeria*
[3]*Consultant Ophthalmologist, University of Ibadan College Hospital (UCH), Ibadan, Oyo State, Nigeria*
iwayemi_ayodeji@yahoo.com

ABSTRACT

Glaucoma is an eye disease characterized by an increase in Intraocular Pressure (IOP) which affects the eye and causes blindness. It is a degenerative disease that damages the nerve fiber layer in the retina of the eye. Its mechanisms are not fully known and there is no fully-effective strategy to prevent visual impairment and blindness. However, if detected early and treatment is carried out at an early stage, it is possible to slow glaucomatous progression and improve the quality of life of possible victims. Despite the great amount of heterogeneous data that has become available for detecting glaucoma, the measures available for early diagnosis are still insufficient, due to the complexity of disease progression and the difficulties in obtaining sufficient measurements. This research aims to develop a framework using visual field data for the detection and prediction of glaucomatous retinal diseases. Classification and regression trees (CART) due its flexibility to continuously check multiple conditions make it a suitable tool to effectively exploit the available data. The processes performed on visual field data include; observation, region extraction and variable preparation, segmentation, classification, etc. Rules set by an ophthalmologist were used as the split criterions which eventually help to detect whether glaucoma was positive or negative.

Keywords: Algorithm, CART, Dataset, Glaucoma, Rules, Tree

INTRODUCTION

A garments industry consists of spinning, dyeing, washing, production, finishing and others sectors. All of these sector deal with the fabrics and there also have some sector that associate chemicals that are hazardous for our health. The textiles sector contains many hazards and risks to workers, ranging from exposure to noise and dangerous substances, to manual handling and working with dangerous machinery. Each processing stage - from the production of materials to the manufacturing, finishing, colouring and packaging -poses risks for workers.

This research work deals with efficient data mining procedure for diagnosing and predicting Glaucoma from perimetric medical records of patients. Glaucoma is the name given to group of eye conditions [1], being the second leading cause of blindness in virtually all parts of the world, Nigeria inclusive. It can be considered as part of a group of eye conditions characterized by optic nerve damage, which may cause loss of vision that can rarely be regained or cured. With prompt and appropriate treatment, the decline can be slowed or halted as it gives no noticeable sign till it gets to an advanced stage with treatment options relatively limited. Current diagnosis of retinal disease relies upon examining retinal fundus image using image processing. The key image processing techniques to detect eye diseases include image registration, image fusion, image segmentation, feature extraction, image enhancement, morphology, pattern matching, image classification, analysis and statistical measurements [2].

For normal eye, cup-to-disc ratio is 0.3 to 0.5 but for glaucomatous eye the ratio becomes 0.8 [3]. Glaucoma can be classified by types, with the most prevalent as the primary open angle glaucoma that is associated with an increase in the fluid pressure inside the eye, called intraocular pressure (IOP), such that reducing IOP can decrease the risk of glaucoma developing and progressing [4]. Visual fields are where the retina is divided into a set of points, and the patient is tested to see how sensitive their eyesight rated at each point which is a number between 0 and 60 dB/log unite, which is a measure of retinal sensitivity, 0 indicates no sensitivity while a value above 50 indicates high sensitivity [5]. Suffice to say that computerizing the perimetry data provides visual field measurements and analyses which help guides clinical decisions, which is the bane of our research in measuring the variability of sensibility using perimetric data from the university of Ibadan teaching hospital UCH Ophthalmology Department. Data mining has proven to be very beneficial in medical analysis as it increases diagnostic accuracy, to reduce costs of patient treatment and to save human resources [6].

The purpose of this work is to develop a framework using the available perimetric visual field data for the detection and prediction of glaucomatous retinal diseases. The objectives are to:

1. develop a technique and analytical method for detecting and predicting glaucoma;
2. use a CART based tree classification model in (MATLAB) and for evaluating the Specificity and sensitivity of visual field on glaucoma progression; and
3. predict glaucoma progression using perimetry field datasets.

LITERATURE REVIEW

Glaucoma is a neurodegenerative disorder of the optic nerve, which causes partial loss of vision with retinal fundus image examination using image processing which are image registration, image fusion, image segmentation, feature extraction, image enhancement, morphology, pattern matching, image classification, analysis and statistical measurements. [2], [7-8]. One of the indicators of glaucomatous eye is change in appearance of optic disk such that Fundus camera imaging of the retina is widely used to diagnose and manage ophthalmologic disorders including diabetic retinopathy, glaucoma, and age-related macular degeneration. [9]. With great improvement in field of medical imaging, Image processing technique helps in early diagnosis of glaucoma and other eye disease. [2]. The early detection of glaucoma can help prevent the loss of vision in the patient, whence he went further to propose a framework for early detection and diagnosis of glaucoma using Perimetry data and Optical Coherence Tomography (OCT) images. [10]. Applying the Data Mining Support Vector Machine (SVM) on the pre-processed data in their result, they proposed that the development of such framework will help in diagnosing the disease which will provide detailed analysis of the progression of glaucoma with the help of graphical representation of the Perimetric as well as OCT data regarding the fact that Perimetry and OCT are the widely performed tests due to their high sensitivity and specificity. Multivariate Time Series (MTS) modeling is considered in glaucoma prediction and based argument on the application to biomedical data where there is small observation but large dataset of variables to consider. [11]. It is believed that application using genetic algorithm and computational method can bypass the size restriction on traditional statistical MTS method which is a novel study. Several studies are reported in literature for detection and prediction of glaucoma. The methods used to detect main features of retinal fundus images such as optic disk, fovea, and exudates and blood vessels using different techniques are specified in [12]. Validation of Retinal Image Registration Algorithms by a Projective Imaging Distortion Model is necessary. [13]. A variety of methods for retinal image registration have been proposed. The authors also presented the validation tool for any retinal image registration method by tracing back the distortion path and accessing the geometric misalignment from the coordinate system of reference standard. The retinal fundus image is widely used in the diagnosis and treatment of various eye diseases such as diabetic retinopathy and glaucoma. [14]. The proposed methodology consists of two steps: in the first step, region of interest (ROI) is found by image by means of morphological processing, and in the second step, optic disk is detected using the Hough transform. [15] reported the research work: 'ORIGA-light: An Online Retinal Fundus Image Database for Glaucoma Analysis and Research'. An online dataset, ORIGA-light, which aims to share clinical retinal images with the public was presented. [16] presented the work,' Retinal Blood Vessel Segmentation Using Gabor Filter and Tophat Transform'. In the paper, a method for retinal blood vessels segmentation by applying firstly Gabor filter to enhance blood vessels was given and then applying top-hat transform. Then, the output was converted to binary image with p-tile thresholding. [3] reported the work,' Optical Cup to Disc Ratio (CDR) Measurement for Glaucoma Diagnosis Using Harris Corner'. In this paper, CDR was determined using Harris Corner. Harris comer detector measures the local changes of the signal with patches shifted in different directions by a small amount.

METHODOLOGY

Design Overview

The objective is to help clinicians better determine the probable progression course of any glaucomatous eye and its diagnosis through a tree-based classification and regression model, using the collated visual field data as the prognostic factor for predicting glaucomatous visual field progression. Tree based algorithm is efficient for the easy understanding of tree structure results generated from CART, nonparametric variables that deals extensively and efficiently with large dataset as the MTS glaucoma datasets used in this study that have the ability to handle mixtures of categorical and continuous explanatory variables. This good omen prompts our decision in the use of CART algorithm applying it to visual fields data for mining and prediction of glaucomatous eye datasets for diagnosis. This we believe that the CART analysis will present results which will be an excellent tool in assisting medical prognosis and diagnosis.

In order to build a tree, we will follow the CART methodology to develop tree models which contain the following three major steps:

1. Grow a large initial tree T_0
2. Iteratively truncate branches of T_0 to obtain a sequence of optimally pruned (nested) sub-tree
3. Select the best tree size based on validation provided by either test sample data or cross validation data.

Tree based CART Algorithm

The tree based CART algorithm is split into two classes, the classification tree and the regression tree. Both of them will be used in this paper for analysis and implementation the model. The two tree algorithm are discussed below

Classification Algorithm

The tree construction starts by splitting a sample or root node into two child nodes according to a binary question. If the observations satisfy the rule, they are sent to the left child node, otherwise they are sent to the right child node. Suppose Y is a binary response variable, and X is a set of covariates where X = (X_1,X_2, ...…… Xp;).

For continuous covariates, the splitting rules **{ Xi>= C } or {Xi > C }**where **C** is a constant observed value. For categorical predictors, the splitting rules **{Xi ξ A} or {Xi ξ A },** where **A** can be any subset of its categories. For simplicity and interpretability, only univariate split or splits on a single covariate is considered.

Regression Tree Algorithm

The method for building a regression tree is similar to that for a classification tree. However, with a regression tree where the dependent variable is continuous, the classification splitting rules cannot be applied. The split criterion employed in CART is the within-node variation of the response, which implies that the node impurity is defined as in equation (1)

$$i(t) = \sum (yi - \bar{y}t)^2 \qquad i \in t \tag{1}$$

Where y_i is individual values of the dependent variable at node **t** and $\bar{y}$ is the average of y_i's at node t. To evaluate a split s at node t into left node $t_L$and right node t_R, we maximize the goodness of split criterion as in equation 2

$$\Delta i(s,t) = i(t) - [i(t_L) + i(t_R)] \tag{2}$$

The Operational Model

The framework creates a view of the entire model where the totality of the design is to predict the progression outcome of the expected sensibility of an individual having his visual field data from the standard automated Perimetry machine. Diagnose a patient to verify if he/she is in danger of the glaucoma impairment for early treatment and cure. As modelled in Fig. 1, the operation model depicts the operational functionality of the entire system. The operational process starts with the Perimetry test which is diagnosed and if the result is positive it refers the doctor decision else the progressive engine considers the expected probabilistic danger the patient stand through the predictive algorithm.

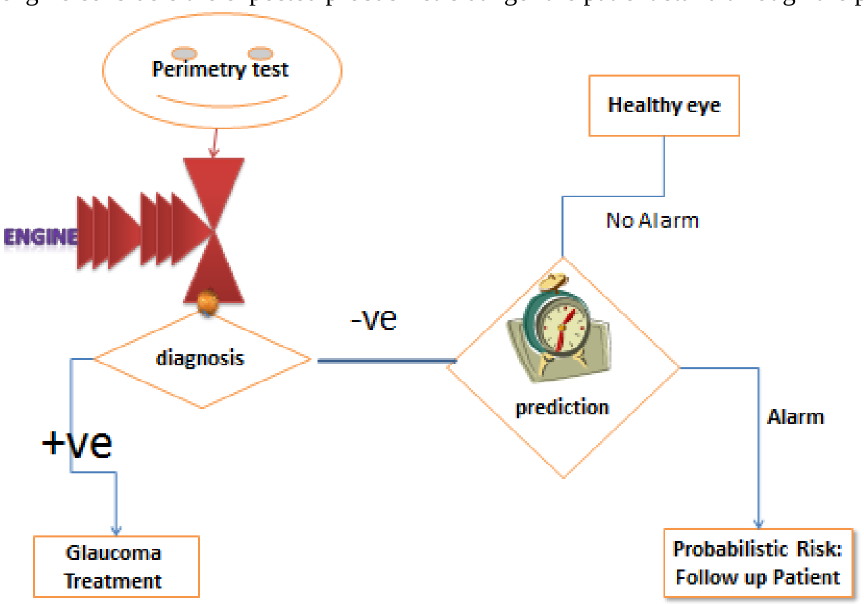

Fig. 1 Operational Model

Data Preparation and Segmentation

Diagnosing glaucoma progression is the prime cause of this research and it has great concern to medical classification task, ideally for predefined training data set though arduous to collect. Data was collected retrospectively from patient charts of glaucoma patients of the University of Ibadan College Hospital (UCH) Ophthalmology Department, with visual fields measurements of accumulated years through their regular ophthalmologic examination. Over the data used, the evaluation of patient reliability test is done for the quality of the visual field by grouping the specimen into progressing, stable and unknown. Finally, we will adjust the visual field threshold values for age measurement by 1dB per decade and half as against per decade in the previous works. The data set collected is pre-processed through an attribute selection by region extraction, using the data filtering technique to filter the data to the reliability set before it is used in the study. After filtering, the data set is segmented into groups of splits and the observation is treated with the CART algorithm for measurement and analysis as depicted in the Fig. 2.

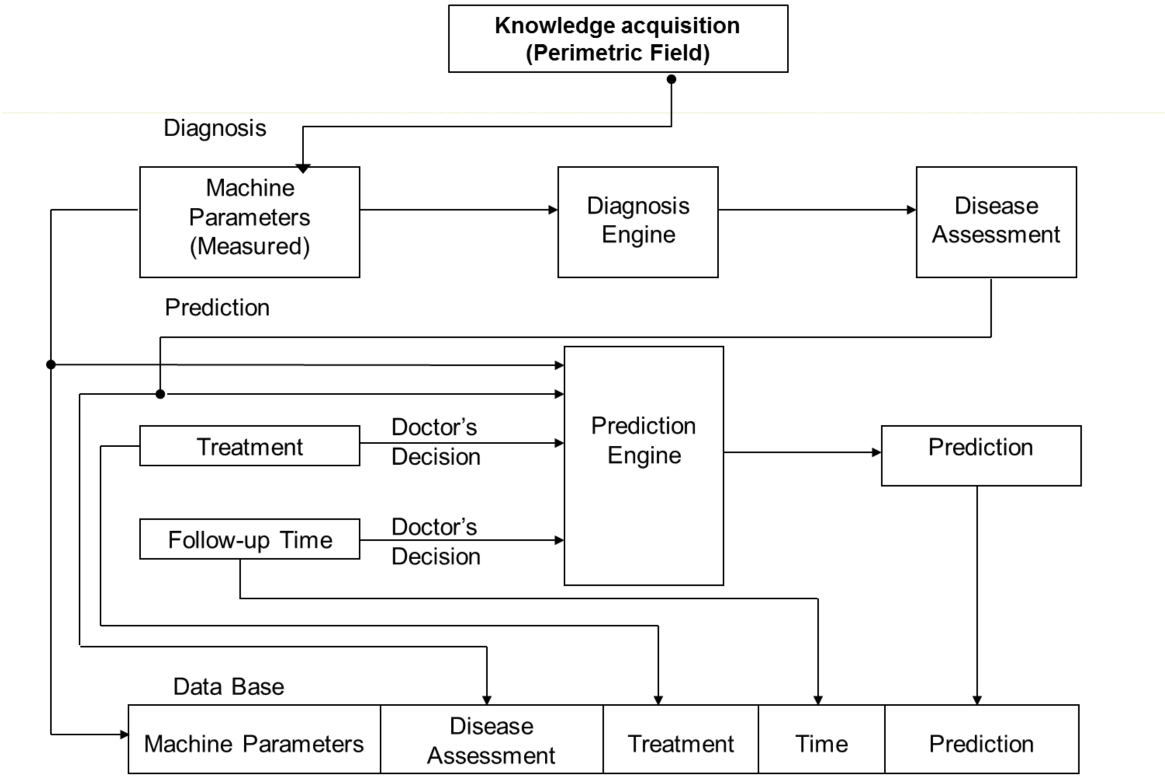

Fig. 2 System framework

DISCUSSION OF RESULT/IMPLEMENTATION

The Simulation Interface

Glaucoma, as aforementioned, is said to be a dangerous disease of the eye that causes blindness if not well taken care of at the early stage. Having inspired this paper, the CART classification algorithms were applied for accurate diagnosis and prediction of glaucoma using the Perimetry visual field data from the university of Ibadan ophthalmology department. The result so far reaching has a great impact in diagnosing and predicting glaucoma in patients as discussed in the succeeding sections. Classification of glaucoma progression is a prime example of a medical classification task that is ideally suited for data mining approaches because pre-classifier training data is available. In this result we present an empirical result of using soft computing in diagnosing and predicting glaucoma using the visual field Perimetry data. Data in health and medical sciences are often longitudinal, where we obtain data on each individual at several points in time. In longitudinal studies, repeated measurements of the same individual results in a source of variation that violates the independence assumption of traditional fixed-effects models. CART has led to considerable related advances in data mining. In this thesis we introduced a soft computing technique and we also describe the algorithm for the bagging and boosting of the proposed tree structure to improve predictive performance. The implementation of the system was carried out using MATLAB simulation software with the classification API in developing the prediction machine used for the diagnosis of glaucoma disease and progression. It is well suited for such implementation and useful in getting accurate results and good tools for analyses and pattern generation. The MATALAB simulation was used in the implementation and development of the system. The MATALAB library toolbox Classification Tree Algorithm was utilized in testing the performance of this model.

Analysis of Results

The Glaucoma Diagnosis and the prediction system also named Glaucoma diapredictor (GDP) is an interactive module designed towards the processing functional framework using the CART algorithm as designed in the methodology, which has enhanced the usability of the system for the prediction and diagnosis. Fig. 3 below shows the Interface.

Dataset Simulation Analysis

The data used was divided into three sets: the first set of data was used for training of the model; the second used for validation and the third set was used for testing of the model generated during validation and training. The data values are generated matrix of the visual field images as shown in Fig. 4 which were used for testing. The approach used was extracting the image spot and converting it to numerical values in the matrix for the training.In the Fig. 4, the dataset is loaded alongside the visual image cup which are the values used in the diagnosis and the prediction of the glaucoma dataset using the trained model. This data is the case data under consideration. When loaded using the prediction engine, the classification model uses the dataset information to make prediction and accurate diagnosis.

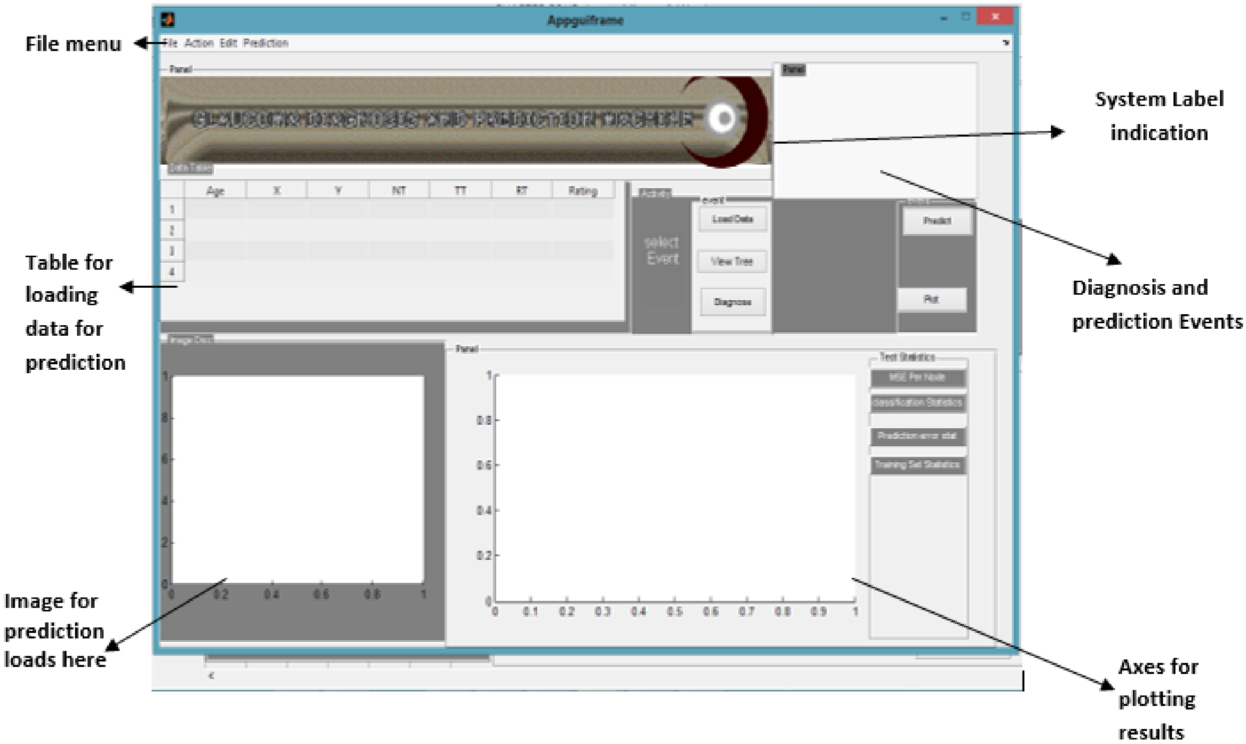

Fig. 3 Glaucoma Diapredictor(GDP) Interactive Interface

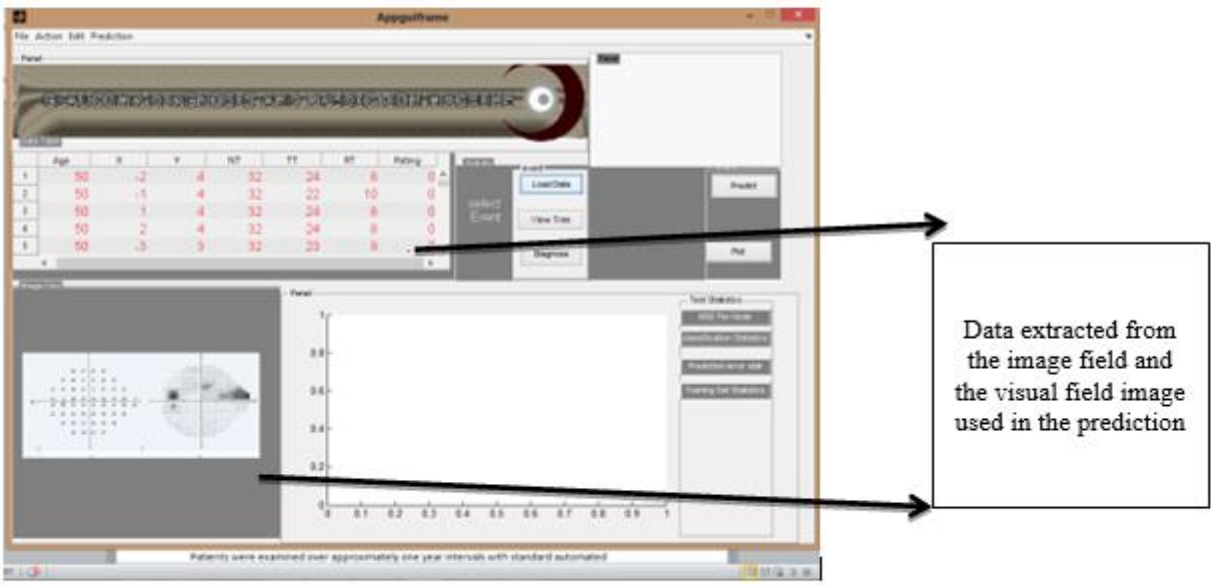

Fig. 4 The visual data and the image for diagnosis and prediction

Fig. 5 above depicts the visual spots of the visual data set under consideration; the coordinate spots are as shown in the axes. According to the visual field glaucoma determination, the deviations of the values in the spot are some of the statistical parameters used for analysis in considering the glaucoma cases. This visual spots are extracted and fielded in a file and later converted into matrix array used in the classification, training, final detection and prediction. The values of the dataset are numeric values measured in decibel, ranging from 0 to 32.

GDP Regression Tree

The regression tree, as discussed, helps to determine the terminal node and the classification deepness using the if-then approach. The principal, splitting could continue until all cases are perfectly classified. From the statistics the data is well classified. In the tree, the classification pattern was generated with the threshold showing the tree traversal of the classification the variables. From the tree traversal, it was seen that the algorithm classified the dataset accurately. The statistics chart indicates that the major classification variables are the visual spots in the matrix data. Some of the variable was reduced and yet a good prediction with minimal classification error was got

Glaucoma Detection Result

In Fig. 6, using the test data in the prediction and diagnosis, we found out from the test that there was 90% accuracy in prediction of the data and the detection was also perfect considering the test data supplied. The case was predicted correctly and the diagnosis was almost perfect considering the appreciable timeline. When GDP was used to test a data, the visual field generated matric dataset and an accurate prediction was observed (Fig. 7). The prediction indicates the rate at which glaucoma is progressing whether it is normal or baseline depending on the state of the visual field of the patient diagnosed.

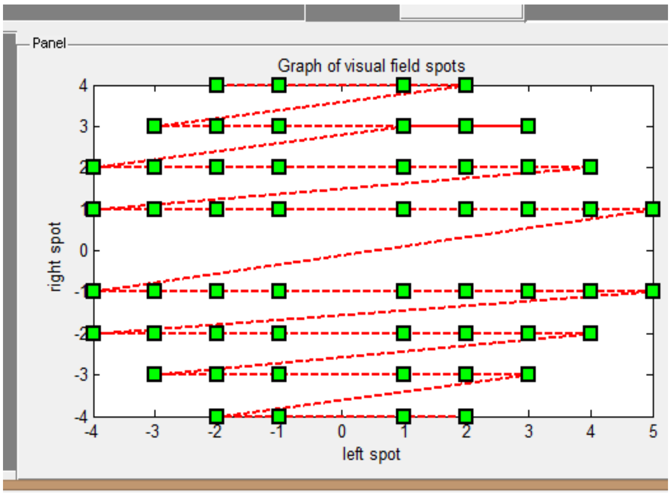

Fig. 5 The graph of visual field spots

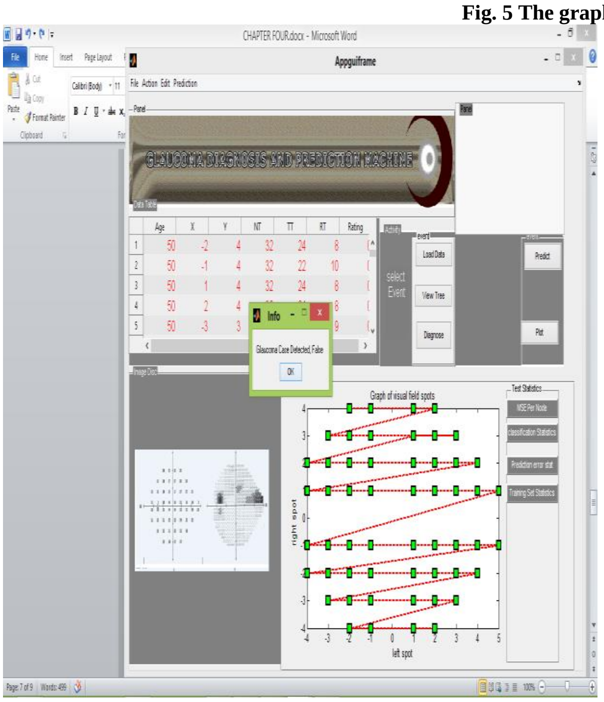

Fig. 6 Glaucoma Detection Result

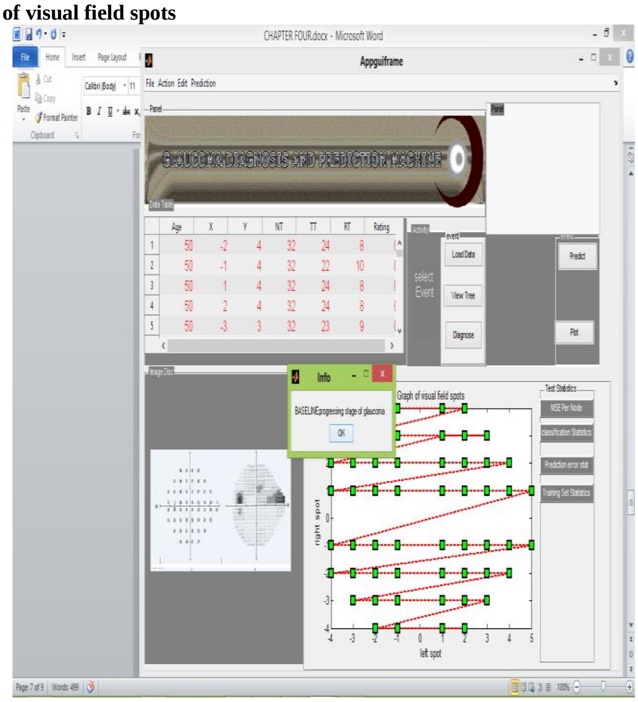

Fig. 7 GDP Prediction Result

CONCLUSIONS

Determining whether a glaucoma patient has deteriorating vision field information is challenging, though an important task. Each visual fields assessment provides a large number of attributes, and the attributes which lends itself to data mining approach as seen using the CART. Our findings from the paper indicate that soft computing approach to prediction of glaucoma better that the point-wise univariate regression with a better result in terms of precision of dataset. In nutshell we can conclude that early detection of Glaucoma can prevent early vision loss in the patient, having developed a system which can help the ophthalmologist in diagnosing the disease which provides a detailed analysis of the progression of. It is also certified in this paper that CART algorithm performs well and fast in the prediction of large multivariate data.

REFERENCES

[1] RN Weinreb and PT Khaw, Primary Open-Angle Glaucoma, *Lancet Journal*, **2004**, 363(9422), 1711-1720.
[2] JP Preeti, Review of Image Processing Technique for Glaucoma Detection, *International Journal of Computer Science and Mobile Computing*, **2013**, 2 (11), 99-105.
[3] D Nilan, Optical Cup to Disc Ratio Measurement for Glaucoma Diagnosis Using Harris Corner, *International Conference on Computing Communication and Networking Technologies*, **2012**.
[4] M Kass, D Heuer, E Higginbotham, C Johnson, J Keltner, J Miller, R Parrish, M Wilson and G Gordon, The Ocular Hypertension Treatment Study: A Randomized Trial Determines that Topical Ocular Hypotensive Medication Delays or Prevents the Onset of Primary Open-Angle Glaucoma, *Archive of Ophthalmology*, **2002,** 120(6), 701-713.
[5] MJ Haley, *The Field Analyser Primer*, San Leandro: Allergan Humphrey, **1987**.
[6] Y Guo, B Guohua and H Yan, Using Bayes Network for Prediction of Type-2 Diabetes, *In Internet Technology and Secured Transactions*, **2012.**
[7] R ManjulaSri and KMM Rao, Novel Image Processing Techniques to Detect Lesions using Lab View, *Annual IEEE India Conference INDICON,* **2011**, 1-4.
[8] B Viralkumar, B Surabhi, S Poonam and KMM Rao, Low Cost Medical Image Processing System for Rural/Semi Urban Healthcare, *IEEE Conference Publications on Recent Advances in Intelligent Computational Systems*, **2011**, 724 – 728.
[9] VS Charles, T Chia-Ling T and R Badrinath, The Dual-Bootstrap Iterative Closest Point Algorithm with Application to Retinal Image Registration, *IEEE Transactions on Medical Imaging*, **2003,** 22(11), 1379- 1394.
[10] C Kinjan and G Ravi, A Proposed Framework for Diagnosis of Glaucoma - A Data Mining Approach, *International Journal of Engineering Research and Development*, **2012**, 3(5).
[11] S Swift, J Kok and X Liu, *Learning Short Multivariate Time Series Models Through Evolutionary and Sparse Matrix Computation*, Brunel University, Uxbridge, Middlesex, United Kingdom, **2005**.
[12] N Kevin, N Jagadish and SN Bhat, Enhancement of Retinal Fundus Image to Highlight the Features for Detection of Abnormal Eyes, *TENCON 2006 IEEE Region 10 Conference***, 2006**, 1- 4.
[13] L Sangyeol, D Michael, A Amoff and MR Joseph, Validation of Retinal Image Registration Algorithms by a Projective Imaging Distortion Model, *29th Annual International Conference of the IEEE EMBS*, France, **2007**.
[14] S Sekhar, *Automated Localisation of Retinal Optic Disk using though Transform*, Department of Electrical Engineering and Electronics, University of Liverpool, UK, **2008**.
[15] Z Zhuo, ORIGA-light: An Online Retinal Fundus Image Database for Glaucoma Analysis and Research, *32nd Annual International Conference of the IEEE EMBS*, Argentina, **2010.**
[16] Y Zafer, Retinal Blood Vessel Segmentation Using Gabor Filter and Tophat Transform, *IEEE 19th Signal Processing and Communications Applications Conference*, SIU, **2011**.