EasyChair Preprint

№ 12986

Development of a Model for Diabetic Retinopathy Image Classification

Olusogo Adetunji, Oladapo Ibitoye, Ayobami Olusesi and Oluwafunso Osaloni

April 10, 2024

Development of a Model for Diabetic Retinopathy Image Classification

Olusogo Julius Adetunji[1, a)], Oladapo Tolulope Ibitoye[2, b)], Ayobami Taiwo Olusesi[3, c)], and Oluwafunso Oluwole Osaloni[4, d)]

[1]*Department of Computer Engineering, Federal University Oye Ekiti, Ekiti State, Nigeria*
[2]*Department of Electrical, Electronics and Computer Engineering, Afe Babalola University Ado Ekiti, Nigeria*
[3]*Department of Computer Engineering, Bells University of Technology, Ota, Nigeria*
[4]*Department of Electrical Engineering, Afe Babalola University, Ado Ekiti, Nigeria*

[a)] Corresponding author: ojadetunji@bellsuniversity.edu.ng
[b)]ibitoyeo@abuad.edu.ng
[c)]atolusesi@bellsuniversity.edu.ng
[d)]wole_osaloni@yahoo.com

Abstract. Diabetic retinopathy is a disease capable of causing vision loss in humans if not treated promptly. This study aims to predict the occurrence of diabetic retinopathy by applying a stacking ensemble classification technique to features extracted from various retinal images. The proposed system was implemented on an embedded system using a Python-based integrated development environment and executed on an Apple M1 computer with a 3.20 GHz CPU and 8 GB RAM running Mac OS Mojave. Diabetic retinopathy prediction involved comparing the performances of individual classifiers, including decision trees, support vector machines, stochastic gradient descent, and extreme gradient boosting, with their ensemble using the Stacking classification technique. The novel ensemble model outperforms individual models, achieving an accuracy of 85%, sensitivity of 81%, and specificity of 86%.

Keywords. Decision tree, Diabetic retinopathy, ensemble model, extreme gradient boosting, stochastic gradient descent

INTRODUCTION

Diabetes, or diabetes mellitus (DM), is a condition that affects the pancreas, often due to a lack of insulin, which regulates blood sugar levels [1]. Insufficient insulin production leads to increased blood glucose levels, damaging various organs, including retinal vessels [2]. In 2020, the World Health Organization reported that 422 million people worldwide had diabetes. A 2021 study by the International Diabetes Federation found that 52 million Africans had diabetes, and as of September 2017, Nigeria's population was estimated at 193.3 million by the United Nations. Meta-analysis indicates that approximately 1 in 17 Nigerian adults, totaling 11.2 million people, had diabetes, with the highest incidence in the south-south zone and the lowest in the north-western zone. Regional variations in diabetes prevalence mirror those of obesity, a major risk factor for type 2 diabetes. A study [3] encompassing 91 countries predicts a significant increase in diabetes cases in the coming years, with a 69 percent rise in developing nations and a 20 percent increase in industrialized nations.

Diabetes is a silent illness characterized by a range of primary and secondary symptoms [4]. If left undetected or untreated, it can progress, leading to the destruction of retinal blood vessels and, in rare cases, complete blindness. Studies indicate that individuals with diabetes for over a decade face an 80 percent higher risk of developing Diabetic Retinopathy (DR) [5]. Research spanning from 1990 to 2020 predicts a 64% increase in DR-related blindness and a 27% increase in visual impairments [6]. Currently, more than 100 million people worldwide are affected by these health issues [7]. DR can remain asymptomatic in its early stages, only revealing itself after significant retinal damage

has occurred. It is considered a severe ocular disease, marked by symptoms such as blurred vision, nighttime vision difficulties, swollen blood vessels within the eye, fluid leakage, and the development of abnormal blood vessels on the retina's surface. As the condition progresses, these abnormal vessels can lead to blurry vision. It's important to note that both eyes may be affected by this retinal degeneration, causing gradual changes in vision for individuals with diabetes [7]. The long-term complications of diabetes can result in diabetic retinopathy, leading to damage to the retinal blood vessels. There are two fundamental pathophysiological mechanisms involved: increased capillary permeability and capillary closure, resulting in clinical signs of retinopathy. If left untreated, it often affects both eyes and can impair vision. Diabetic retinopathy is categorized into two types: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR) [8].

Small retinal blood vessels in NPDR patients have some damage, which leads to fluid buildup, leakage, or rupture. The initial step is NPDR, followed by PDR, which is a later stage in the development of new aberrant capillaries. It is well recognized that retinal vein expansion indicates the presence of diabetic retinopathy. Additionally, pre-occlusive changes may occur in microscopic capillaries. Micro aneurysms (MA), which are minute bulges in the vascular walls, are the result of this. A person may be at risk of acquiring mild non proliferative diabetic retinopathy in addition to vision problems [9]. In the retinal vasculature, micro aneurysms are spherical deformations brought on by blood or lipid blockage. Micro aneurysms, also known as red lesions, are an early stage of the illness that must be identified. Only two or three red dots are observed at the beginning of NPDR. Exudates (EX) and Cotton Wool Spots (CWS), two bright lesions, also increase in number as the illness worsens [8]. Currently, diabetic retinopathy is clinically detected by fundus imaging, which mostly consists of digital fundus imaging (DFI). Given that it is the cheapest, the DFI imaging technique is still one of the ones that is advised for DR detection. A computer-aided diagnosis (CAD) system can be beneficial in the case of a large-scale screening programme where ophthalmologists are only involved in a tiny fraction of the total programme [10].

Given the rise in the prevalence of diabetic retinopathies, it may assist reduce waiting times for patients. Automated diagnosis of these conditions using retinal images is essential.

for determining the aetiology, diagnosing, and treating retinal problems, as well as for organizing surgery [10]. Exploring a hypothesis space in search of appropriate hypotheses that will produce excellent predictions for a specific instance is the aim of supervised learning algorithms [11]. Finding a good hypothesis can be difficult, even if the space for hypothesis contains ones that are appropriate for a particular circumstance [12]. To create a more effective hypothesis, ensembles combine many theories which entail building a number of basic classifiers from which a superior classifier is built. The dataset employed for this work was obtained from kaggle machine learning repository [13]. Apart from the introductory part, this paper is further section into related works (section II), section III represents design methodology, section IV describes the experimental results and discussions and the last section is the conclusion part.

RELATED WORKS

Varun C. et al [14] employed deep learning algorithm to accurately detect and quantifying diabetic retinopathy disease through utilizing digital angiography and discern, trace the vessel boundary boundaries and concentrate vital metrics of clinical value. The author's computation relied on a coordinated filter approach in conjunction with existing knowledge about the retinal blood vessel features. By displaying the vessel profile using Gaussian capabilities, improved assessments of vessel widths above prior computations are obtained by creating unique algorithms, however not much variance could be detected. Raju M. et al [15] Adopted deep learning algorithm for automatic prediction of diabetic retinopathy. The result indicated that accuracy; sensitivity and specificity of 93.28%, 80.28% and 92.29% respectively were observed. Mann K.S. et al [16] Employed artificial neural network concepts in order to observe the illness at preliminary stage. Artificial neural network and the algorithm were utilized to identify the diabetic retinopathy in humans at an early stage, but they were only tested on a tiny sample of photographs from the common database. The authors tried supervised techniques for detection and finally concluded that they could get excellent results by utilizing artificial neural network. The failure to account for noise was a significant flaw in their approach.

In the work of Carrera E.V. et al [17], performance of support vector machine (SVM) was compared with decision tree algorithm when tested on 400 retinal image datasets. When identified with 4-grade scale of non-proliferative diabetic retinopathy. Support vector machine algorithm performed better than decision tree algorithm with maximum sensitivity of 95% and average accuracy of 85%. Also, Tsao H. et al [18] applied and compared support vector machine algorithm with decision tree and logistic regression algorithms. The observed results indicated that support vector machine performed better than both logistic regressions and decision tree algorithms with accuracy and AUC (area

under the ROC curve; ROC (receiver operating characteristics curve)) of 79.5% and 0.839 respectively. Logistic regressions and decision trees algorithms used both insulin and duration of diabetes as distinct features of the occurrence of diabetic retinopathy. Oladele T.O. et al [19] Compared the performance of four algorithms with and without feature selection of diabetic retinopathy dataset. The algorithms compared are support vector machine, decision tree, K- nearest neighbor algorithm and multilayer perception. The experimental results proved that the performance is better with feature selection than without feature selection.

Wrapper feature selection was employed and the results showed that support vector machine was more pronounced as its sensitivity increased from 26.3% to 74.8% and also, its accuracy increased from 58.21% to 70.37%. Also, it is observed that adoption of feature selection before prediction increased its time taken for the prediction of diabetic retinopathy. Rathi P. et al [20] applied different machine learning algorithms; such algorithms include support vector machine, bagged trees, logistic regressions and k neighbor algorithms. The 300 patients' record of dataset was classified with these algorithms. The experimental results proved that bagged trees perform better than all other classifiers with the accuracy of 93% Shen Z. et al [21]. Romero-Aroca P. et al [22] Built version of clinical decision support system (CDSS) for the prediction of diabetic retinopathy in type 1 diabetes patients. The model achieved an accuracy, specificity and sensitivity of 79.5%, 83% and 65.7% respectively for any kind of diabetic retinopathy and an accuracy, specificity and sensitivity of 91.8%, 87.1% and 87.8 respectively for sight threatening diabetic retinopathy. Das D. et al [26] developed deep learning model for the prediction of diabetic retinopathy using fundus images. Experimental results indicated that NetB4 algorithm has the highest validation accuracy of 79.11%.

METHODS

The method and approach applied for this study are covered in this section. The recommended model, dataset collection, dataset description, data visualization, and classification algorithms are among the strategies that are utilized to carry out the analysis of this work. Data gathering is the study's first stage. An online dataset, a combination of the Messidor_2 and Eye Pac dataset used for this work was obtained from the Kaggle Machine Learning repository. To determine if a picture has symptoms of diabetic retinopathy or not, the dataset comprises characteristics that were taken from the image collection. The dataset's characteristics and labels are then determined. The dataset is then split into two sets: one for training, where the majority of the data is used, and one for testing. Four different classification methods have been fitted in the training set for the model's analysis performance. The decision tree, support vector machine, stochastic gradient descent, and XG Boost were the techniques employed in this study. Newer data is delivered without outputs once the algorithm has finished learning from training datasets. The final model produces the output based on the information it learned from the training set of data. The study determines how stacking classification ensemble of the algorithms will perform and vary based on the data. The block diagram of the developed model is presented in Figure 1.

Data Collection

A dataset from the Kaggle Machine Learning Repository is utilized in this study [13] To determine if a picture has symptoms of diabetic retinopathy or not, this dataset comprises characteristics that were taken from the Messidor and Eye Pac image collection. Each feature is a descriptor of an image-level descriptor, a descriptive feature of an anatomical component, or a detected lesion. The Messidor database was created to aid research on computer- assisted diabetic retinopathy diagnosis. Kaggle, GitHub, and other websites provide many datasets that have been utilized for various projects focused on diabetic retinopathy. This dataset will be suitable for the job as the study sought to deal with the identification of diabetic retinopathy and it contains a variety of characteristics. The dataset consists of five categories namely no diabetic retinopathy (0), mild non-proliferative diabetic retinopathy (1), moderate non-proliferative diabetic retinopathy (2), severe non- proliferative diabetic retinopathy (3), proliferative diabetic retinopathy (4).

Image Pre-processing

In machine learning, image pre-processing is the process of further modification to the acquired images which constitutes the dataset. The fundus images are high in resolution and requires "down sampling" in accordance to the standard compatible with different machine learning backbones. In this study the images are down-sampled to 224 by 224 by 3. This assists the model to effectively understand the images [26, 27]. In this study, parameters such as

variance and standard deviation are included in the algorithms that performed feature extraction of the Diabetic Retinopathy images.

Classifiers/Algorithms

Different algorithms were used for the prediction of diabetic retinopathy. This study employed stacking ensemble classification technique of four algorithms namely decision tree, support vector machine, stochastic gradient descent, and XG Boost.

1. XG Boost Classifier: A networked, scalable gradient boosted decision tree (GBDT) machine learning system is called Extreme Gradient Boosting (XGBoost) [23]. It provides parallel tree boosting and is the best machine learning package for regression, classification, and ranking problems. Understanding supervised machine learning, decision trees, ensemble learning, and gradient boosting, among other machine learning theories and methods, is crucial to comprehending XGBoost.
2. Support Vector Machine: The Support Vector Machine (SVM), a state-of-the-art classification method, was created in 1992 [24]. According to a more detailed explanation, a support vector machine produces a hyper plane or collection of hyper planes in a high or infinite-dimensional space that may be used for classification, regression, or other tasks. Assuming that the bigger the margin, the less the classifier's generalization error is, the hyper plane with the largest distance from the nearest training data point for each class provides a suitable separation. Support vector machine falls under the broad heading of kernel approaches. Kernel methods utilize techniques created for linear classifiers, it is possible to create non-linear decision boundaries and also, apply classifier to data without a clear fixed-dimensional vector space representation is possible with kernel functions.
3. Stochastic Gradient Descent (SGD) Classifier: Stochastic Gradient Descent (SGD) Classifier is an optimization technique frequently used in machine learning applications to identify the model parameters that correlate to the greatest match between anticipated and actual outputs [25]. It is a rough yet effective method. Machine learning applications frequently employ SGD. Although SGD occasionally becomes trapped at a local minimum or a saddle point rather than locating the global minimum, it is often used in practice. For instance, neural networks employ gradient descent to discover weights and biases.
4. Decision Tree Classifier: Applications for decision tree classifiers are numerous and productive. Their main strength is their ability to extract descriptive decision- making information from the given data[29]. Decision trees may be made using training sets.
5. Stacking Ensemble Classification Technique: Stacking ensemble model involves combination of base estimators to reduce bias and variance in order to improve predictive performance of final estimator. The final estimator is trained through cross validation. The research would be carried out using stacking ensemble model. Stacking ensemble techniques has the advantage of reducing variance and bias, hence eliminating the over fitting effect of models in the procedure.

Algorithm Listing 1: Stochastic gradient descent (Source:[25])
Let $(x^{(i)}, y^{(i)}$ be the training example
$Cos\ t\ (\theta), (x^{(i)}, y^{(i)})) = (1/2) \sum (h\ \theta\ (x^{(i)} - y^{(i)}\ y^{(i)})^2$
$J_{train}(\theta) = (1/m) \sum cost(\theta, (x^{(i)}, y^{(i)}))$
Repeat
{
For i = 1 to m
{
$\theta_j = \theta_j$ - (learning rate) $\times \sum h_\theta(x^{(i)} - y^{(i)}x^{(i)})$
For every $j = 0$……
…….. n
}

Algorithm Listing 2: Extreme gradient boosting (Source:[23])

$\widehat{\boldsymbol{y}}_i = \sum_{k=1}^{K} \boldsymbol{f_k}\ (\boldsymbol{x_i}), \boldsymbol{f_k}$ (1)

where k = number of trees, f = functional space of f, f is a set of possible

Objective function for the above model is given by

$\text{obj}(\theta) = \sum_{i}^{n} l(y_i, \hat{y}i) + \sum_{k=1}^{k} \Omega(f_k)$ (2)

$\widehat{\boldsymbol{y}}_{\boldsymbol{i}}^{(\mathbf{0})} = \mathbf{0}$

$\hat{y}_i^{(1)} = f_1(x_i) = \hat{y}_i^{(0)} = \hat{y}_i^{(0)} + f_1(x_i)$

$\hat{y}_i^{(2)} = f_1(x_i) + f_2(x_i) = f_2(x_i) + \hat{y}_i^{(1)}$

……

$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i)$

$obj^{(t)} = \sum_{i=1}^{n} i(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^{t} \Omega(f_i)$

$= \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) + constant$

$obj^{(t)} = \sum_{i=1}^{n} l(((y_i, y_i^{(t-1)} + f_t^{(t-1)})))^2 + \sum_{i=1}^{t} \Omega(f_i)$

$= \sum_{i=1}^{n} [2(\hat{y}_i^{(t-1)} - y_i) f_t(x_i)^2 + f_t(x_i)^2] + \Omega(f_t) + constant$ (3)

Equation 1 represents the sum of prediction scores of each individual decision tree; equation 2 represents objective function of the model in equation 1; equation 3 represents the redefined expression for objective function.

Algorithm Listing 3: Support vector machine (Source: [28])

Let $(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{i})$ be training data points

Step 1: Compute matrix $\boldsymbol{H} = [\boldsymbol{Hi},]$ where $\boldsymbol{Hi},= \boldsymbol{y}^{(i)}\boldsymbol{y}^{(j)}(\boldsymbol{x}^{(i)}. \boldsymbol{x}^{(j)})$

Step 2: Select value Q that controls misclassification.

Step 3: Obtain $a = (a\mathbf{1}, a\mathbf{2}, \ldots, a\boldsymbol{n})$

Maximize $(\sum a + \mathbf{1} a^T \boldsymbol{H} a)$ subject to the constraints

$\sum \boldsymbol{i}\ a\boldsymbol{i}\boldsymbol{y}^{i} = \mathbf{0}, \mathbf{0} \leq a\boldsymbol{i} \leq Q$

Step 4: Calculate $a = \sum \boldsymbol{i}\ a\boldsymbol{i}\boldsymbol{y}^{(i)}\boldsymbol{x}^{(i)}$

Step 5: Identify the supporting vectors. These are all the points for which

$\mathbf{0} < a\boldsymbol{i} \leq Q$

Step 6: Compute $\boldsymbol{b} = \mathbf{1} \sum (\boldsymbol{y}^{s} - \sum_{\boldsymbol{n_s}} \boldsymbol{a}\ \boldsymbol{s}^{(i)(i)}. \boldsymbol{x}\ \boldsymbol{s}^{(s)})\ \boldsymbol{i}$

Step 7: Compute $\boldsymbol{sign}(a^T \boldsymbol{x}' + \boldsymbol{b})$ for the classification of the given point $\boldsymbol{x}'$

Algorithm Listing 4: Decision Tree Algorithm (Source: [29])

Algorithm:

Input: T// Decision Tree

D// Input Database

Output: M// Model Prediction

DT Proc algorithm:

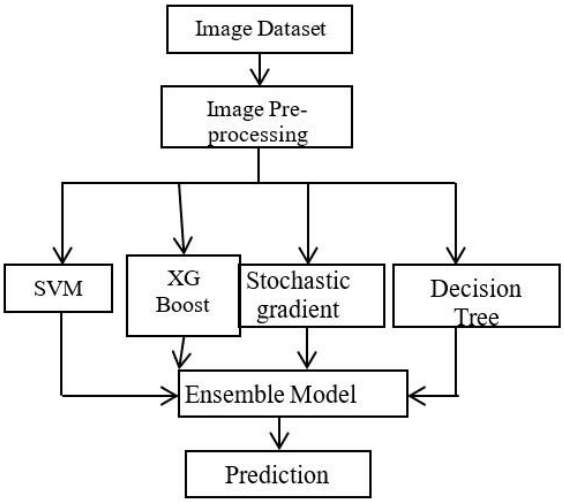
// simplest algorithm to illustrate prediction technique using DT.

For each t Ɥ D do n = root node of T;

While n not leaf node do Obtain answer to question on n applied to t;

Indentify arc from t, which contains correct answer;

n = node at end of this arc;
Make prediction for t based on label of n;..



Experimental Setup

The developed system has been implemented in Python 3.5 and executed in an Apple M1 chip computer with 3 GHz CPU and 8 GB RAM under Mac OS Big Sur (64 bit). The study also covers the evaluation and comparison of the proposed technique with other classification methods in order to interpret the performances of classifiers in predicting diabetic retinopathy. The image dataset was obtained from the Kaggle machine learning repository website. The dataset consists of five different labels of no diabetic retinopathy (0), mild non-proliferative diabetic retinopathy (1), moderate non-proliferative diabetic retinopathy (2), severe non-proliferative diabetic retinopathy (3), proliferative diabetic retinopathy (4). Percentage split method was used which contains 75% training set of 1350 images and 25% testing set of 450 images. Performance evaluation of the developed system was done using accuracy, specificity and sensitivity metrics.

RESULTS

The algorithms employed in this study are decision tree, support vector machine, stochastic gradient descent, and XG Boost. Stacking ensemble classification technique was then used to combine these single classifiers. The performance evaluation is done by comparing the performances of singled classifiers with stacking ensemble model.

TABLE 1. Results of Singled Classifiers and Ensemble Model

Classifier	TN	TP	FN	FP	Accuracy	Specificity	Sensitivity (%)
XG	121	194	85	50	70	70.8	69.5
SGD	168	151	59	72	71	70	72
DT	199	85	44	122	63	62	66
SVM	281	92	23	54	83	84	80
EM	326	56	13	55	85	86	81

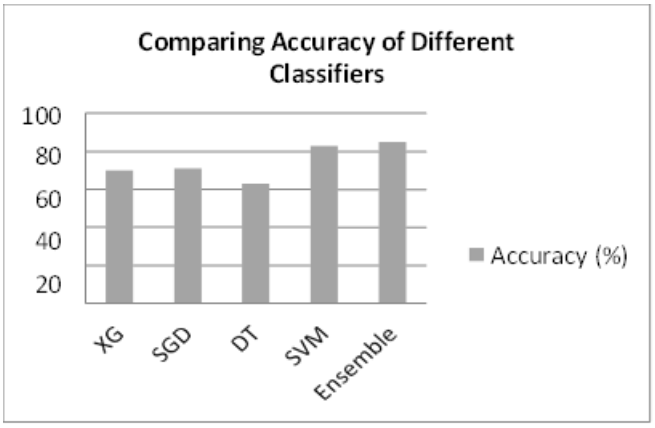


FIGURE 1. Flow chart of proposed model

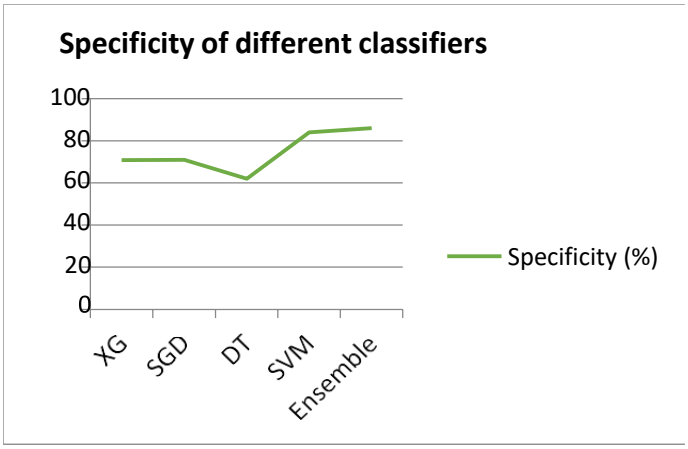


FIGURE 2. Graphical representation of accuracy of singled

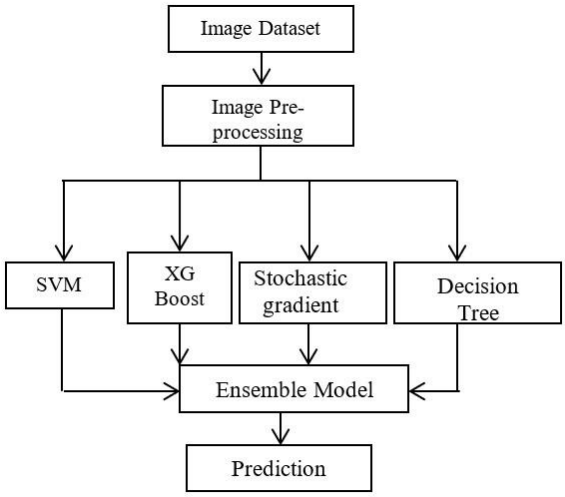


FIGURE 3. Graphical representation of specificity of singled and Ensemble model

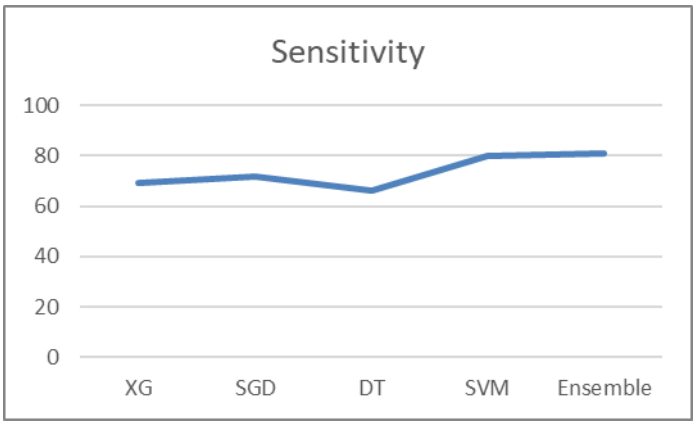


FIGURE 4. Graphical representation of sensitivity of singled and Ensemble model

Performance Comparison of the Singled Classifiers with an Ensemble Model

Table 1 shows the results of the application of singled and ensemble-based models. Performance evaluation metrics which include accuracy, sensitivity and specificity of singled classifiers were compared with an ensemble model in the figures 2, 3 and 4. Support vector machine (SVM) performed better than other singled classifiers with 83%, 84% and 80% accuracy, specificity and sensitivity respectively. Furthermore, developed ensemble model performed better than support vector machine with accuracy, specificity and sensitivity of 85%, 86% and 81% respectively.

CONCLUSION

It can be deduced from the results obtained using selected performance evaluation metrics that stacking ensemble classification technique outperformed the singled classifiers. The results obtained can be compared using other algorithms or on other diseases. Future research can also take advantage of the optimization prowess of stacking ensemble to improve the most appropriate blend of models, in order to get the best combination of multiplicity and predictive performance from the models.

REFERENCES

1. P. L. Gunawardhana, R. Jayathilake, Y. Withanage, and G. U. Ganegoda, “Automatic Diagnosis of Diabetic Retinopathy using Machine Learning: A Review,” Proc. ICITR 2020 - 5th Int. Conf. Inf. Technol. Res. Towar. New Digit. Enlight., pp. 0–5, 2020, doi: 10.1109/ICITR51448.2020.9310818.
2. E. HENSLEY DUKU, M.; GU, S.;BENHAGAN, “A Novel Techniques for Detection of Diabetic Mellitus in Human Eyes,” Digit. Image Process. [S.l.], v. 11, n. 9, p. 157-168, Jan. 2020. ISSN 0974 – 9586. Available<http//www.ciitresearch.org/dl/index.php/dip/article/vie w/DIP092019002.>. Date accessed 20 Jul. 2022., 2020.
3. J. E. Shaw, R. A. Sicree, and P. Z. Zimmet, “Global estimates of the prevalence of diabetes for 2010 and 2030,” Diabetes Res. Clin. Pract., vol. 87, no. 1, pp. 4– 14, 2010, doi: 10.1016/j.diabres.2009.10.007.
4. Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, “Predicting Diabetes Mellitus With Machine Learning Techniques,” Front. Genet., vol. 9, no. November, pp. 1–10, 2018, doi: 10.3389/fgene.2018.00515.
5. R. O. Ewuga, O. O. Adenuga, P. D. Wade, and J. O.Edah, “Prevalence and risk factors for diabetic retinopathy in north-central Nigeria,” Ghana Med. J., vol. 52, no. 4, p. 215, 2018, doi: 10.4314/gmj.v52i4.8.
6. R. R. A. Bourne et al., “Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: The Right to Sight: An analysis for the Global Burden of Disease Study,” Lancet Glob. Heal., vol. 9, no. 2, pp. e144–e160, 2021, doi: 10.1016/S2214-109X(20)30489-7.
7. J. L. Leasher, “Global Estimates on the Number of People Blind or Visually Impaired by Diabetic Retinopathy,” Diabetes Care, 2016.39(9) p. 1643-1649., 2016.

8. V. M. Mane, D. Jadhav, and R. B. Kawadiwale, “Preprocessing in Early Stage Detection of DiabeticRetinopathy Using Fundus Images, in Advancements of Medical Electronics.2015,” Springer. p. 27-38., 2015.
9. N. Singh, R., Kumar, P., Booth, “Diabetic retinopathy: an update,” Indian J. Ophthalmol. 2015.56(3) p.179., 2015.
10. W. S. Ng, W. M. H. W. Mahmud, A. K. C. Huong, W.N. H. W. Kairuddin, H. S. Gan, and R. M. A. R. Izaham, “Computer Aided Diagnosis of Eye Disease for Diabetic Retinopathy,” J. Phys. Conf. Ser., vol. 1372, no. 1, 2019, doi: 10.1088/1742- 6596/1372/1/012030.
11. A. O. Julius et al., “Early Diabetic Risk Prediction using Machine Learning Classification Techniques,”International Journal of Innovative science and research Technology vol. 6, no. 9, pp. 502–507, 2021.
12. T. Ching et al., Opportunities and obstacles for deep learning in biology and medicine, vol. 15, no. 141. 2018. doi: 10.1098/rsif.2017.0387.
13. A. Mohammed,“DiabeticRetinopathy_Messidor_EyePac_PreProcesse d Diabetic Retinopathy,”Available online Https//www.Kaggle.Com/Datasets/Mohammadasimblu emoon/Diabeticretinopathy-Messidor-Eyepac-Preprocessed, Accessed June 09, 2022., 2020.
14. C. Varun, G.; Lily, P.; Mark, “Development and validation of a deep learning Algorithm for Detection of Diabetic Retinopathy,” Int. J. Image, Graph. Signal Process., 2016.
15. M. Raju, V. Pagidimarri, R. Barreto, A. Kadam, V. Kasivajjala, and A. Aswath, “Development of a deep learning algorithm for automatic diagnosis of diabetic retinopathy,” Stud. Health Technol. Inform., vol. 245, no. January, pp. 559–563, 2017, doi: 10.3233/978-1- 61499-830-3-559.
16. K. S. Mann and S. Kaur, “Segmentation of retinal blood vessels using artificial neural networks for early detection of diabetic retinopathy,” AIP Conf. Proc., vol. 1836, no. June 2017, 2017, doi: 10.1063/1.4981966.
17. E. V. Carrera, A. Gonzalez, and R. Carrera, “Automated detection of diabetic retinopathy using SVM,” Proc. 2017 IEEE 24th Int. Congr. Electron. Electr. Eng. Comput. INTERCON 2017, no. August, 2017, doi: 10.1109/INTERCON.2017.8079692.
18. H. Tsao, P. Chan, and E. C. Su, “Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms,” 13th Int. Symp. Bioinforma. Res. Appl. (ISBRA 2017) Honolulu, Hawaii, USA, vol. 19, no. Suppl 9, 2018.
19. T. O. Oladele and R. O. Ogundokun, “Application of Data Mining Algorithms for Feature Selection and Prediction of Diabetic Retinopathy,” Springer Nat. Switz. AG 2019, no. June, 2019, doi: 10.1007/978-3- 030-24308-1.
20. P. Rathi and P. Shrivastava, “Prediction of Diabetic Retinopathy Using Classification Techniques,” Solid State Technol. Vol. 63 Issue 2s, no. December, 2020.
21. Z. Shen, Q. Wu, Z. Wang, G. Chen, and B. Lin, “Diabetic Retinopathy Prediction by Ensemble Learning Based on Biochemical and Physical Data,” Sensors, pp. 1–19, 2021.
22. P. Romero-Aroca et al., “Validation of an Algorithm for the Prediction of Diabetic Retinopathy in Type 1 Diabetic Patients in a Spanish Population,” Clin. Ophthalmol., vol. 16, no. March, pp. 715–722, 2022, doi: 10.2147/OPTH.S351790.
23. C. Chen, T. Guestrin, “Xgboost: A Scalable tree
24. boosting system,”Proc.SIGKDD, pp 785-794, 2016.
25. D. K. Srivastava and L. Bhambhu, “Data classification using support vector machine,” J. Theor. Appl. Inf. Technol., vol. 12, no. 1, pp. 1–7, 2010.
26. N. Ketkar, “Deep Learning with Python,” Deep Learn. with Python, pp. 113–132, 2017, doi: 10.1007/978-1-4842-2766
27. D. Das, S. K. Biswas, and S. Bandyopadhyay, “Detection of Diabetic Retinopathy using Convolutional Neural Networks for Feature Extraction and Classification (DRFEC),” Multimed. Tools Appl., 2022.
28. Y. Jung and D. Kim, “Feature Extraction of Non-proliferative Diabetic Retinopathy Using Faster R-CNN and Automatic Severity Classification System Using Random Forest Method 2 . Diagnosis of Diabetic Retinopathy using Faster R-CNN and This study designed a system that can quic,” | J Inf Process Syst, Vol.18, No.5, pp.599~613, vol. 18, no. 5, pp. 599–613, 2022.
29. Vapnik, V. The Nature of Statistical Learning Theory. Springer-Verlag, 2000
30. Seema, M. Rathi and Mamta. Decision Tree: Data Mining Techniques. International Journal of Latest Trends in Engineering and Technology (IJLTET), 1(3), 150–155, 2012