Science World Journal Vol. 20(No 4) 2025 https://dx.doi.org/10.4314/swj.v20i4.14
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Full Length Research Article

MACHINE LEARNING MODELS FOR PREDICTION OF GLAUCOMA STATUS USING SIGNIFICANT DEMOGRAPHIC, CLINICAL, AND LIFESTYLE RISK FACTORS

[1]Gerald I.O., [1]Babayemi A.W., [2]Sadiq Mohammed Abdullahi, *[1]Buhari S.A.

[1]Faculty of Physical Science, Department of Mathematics, Abdullahi Fodio University of Science and Technology, Aliero, Nigeria
[2]National Eye Centre, Kaduna, Nigeria

*Corresponding Author Email Address:buharishehukaura@gmail.com

ABSTRACT
Glaucoma is an eye condition that damages the optic nerve, leading to irreversible vision loss. Despite advances in diagnostic techniques, it remains a significant public health concern, particularly in resource-constrained settings. Previous studies on Machine Learning (ML) used Demographic Risk Factors (DRFs), Clinical Risk Factors (CRFs), and Fundus Images (FI) to predict Glaucoma, but did not emphasize the use of lifestyle risk factors (LRFs) alongside DRFs and CRFs. A dataset of sample size 200 patients was collected from the National Eye Centre (NEC), Kaduna, from 25th November 2024 to 19th December 2024 using personal interview. The results of the study revealed that all fourteen risk factors of Glaucoma were significant. SVM, DT, K-NN, NB, and MLP ML models predicted 97.1%, 99.0%, 96.1%, 96.1%, 98.0% and 92.6%, 100%, 92.6%, 96.3%, 96.3% Glaucoma patients in the training and test sets. 94.8%, 98.3%, 94.8%, 93.1%, 96.6% and 84.6%, 92.3%, 76.9%, 84.6%, 92.3% Non-Glaucoma patients in the training and test sets. The models have perfect performance and better 5-folds and 8-folds cross-validation. The study concludes that the use of significant LRFs alongside with DRFs and CRFs could help to predict Glaucoma and Non-Glaucoma patients effectively.

Keywords: Glaucoma, Significant risk factors, Machine learning, Blindness, Training and Test sets

INTRODUCTION

Glaucoma is the second leading cause of irreversible loss and accounting for 8% of all blindness worldwide. It is characterized by progressive optic nerve damage and visual field loss. Approximately 17% of blindness in Nigeria was due to Glaucoma. It is the second leading cause of blindness in Nigeria after Cataract. In its early stages, Glaucoma patients might not notice any symptoms or even experience poor vision because it affects the peripheral vision first before the central vision (NIH, 2023). Early prediction is crucial for preventing vision loss. Despite advances in diagnostic techniques, Glaucoma remains a significant public health concern, particularly in resource-constrained settings and Machine Learning (ML) models can offer a promising solution for enhancing predicted cases of Glaucoma, but its effectiveness depends on applying Significant Risk Factors (SRFs) (Pascolini & Mariotti, 2010).

Cases of Glaucoma are being reported frequently, and it has reached the extent that even children less than 10 years of age are affected, and some become blind due to late diagnosis (Santos *et al.,* 2023). The individual affected by blindness finds it difficult to gain employment, and this leads to dependence on family members and society. This eye disease was a threat to the well-being of people, and it needs urgent attention. Treatment can be expensive, and using the significant risk factors to predict the disease is essential. Some Demographic Risk Factors (DRFs) and Clinical Risk Factors (CRFs) such as age, family history of eye diseases, presence of diabetes, high blood pressure, obesity were used to predict Glaucoma but Lifestyle Risk Factors (LRFs) could also cause the risk of having the disease, therefore the inclusion of these risk factors could assist in predicting Glaucoma and Non-Glaucoma patients effectively. Because many eye diseases have LRF contributions, for instance, smoking accelerates Cataract formation and Age-related macular degeneration; near work is linked to Myopia progression; poor diet and light exposure influence Glaucoma and Diabetic retinopathy. Similarly, studies revealed that when LRFs are incorporated, it would lead to a consistent area under the curve and improve model performance. Therefore, the incorporation of these risk factors would help immensely in identifying Glaucoma and Non-Glaucoma patients more effectively (NIH, 2024).

Existing studies, such as Anshul *et al.* (2020), used an SVM model for the prediction of Glaucoma development before and after disease onset using the significant CRFs. Wang *et al.* (2020) used ML models for predicting the 10-year risk of Cataract surgery using significant DRFs and CRFs. An *et al.* (2020) employed ML models for the diagnosis of patients with Glaucoma using a fundus images dataset. They discovered that inception-V_3, Visual Geometry Group -19 (VGG-19) and Resnet-50 achieved prediction accuracy of 94.0%, 93.5% and 92.6% in diagnosis of Glaucoma with AUROC curve of 91.2%, 90.5% and 88.6% respectively and Malik *et al.* (2020) trained ML models namely RF, DT, Convolutional Neural Network (CNN) and NB for prediction of Glaucoma using demographic and CRFs. The models achieved prediction accuracy of 93.2%, 92.8%, 89.5% and 86.7% respectively. Shuldiner *et al.* (2021) used CRFs to assess whether an SVM model could predict eyes that would undergo rapid Glaucoma progression based on the Initial Visual Field (IVF) test. Mahyar *et al.* (2021) employed an ML model (K-NN) for the prediction of Glaucoma using DRFs and CRFs. Fei *et al.* (2022) used an ML model to predict Glaucoma incidence. Their study demonstrated the feasibility of ML in the early detection and prediction of Glaucoma progression. Marouf *et al.* (2023) formulated an SVM model for the prediction of eye diseases using DRFs and CRFs. They discovered that the model achieved a prediction accuracy of 94.8%. The work of Ravindranath *et al.* (2025) applied ML and Deep learning models (XGBoost and Logistic regression) to predict patients with Glaucoma using demographic patient data only collected from a

health survey. The results of their work revealed that XGBoost achieved 90.4% accuracy, 77.1% balanced accuracy, 57.2% precision, 58.6% recall, and an AUROC curve of 89.0%. Logistic regression had 79% accuracy, 28.1% precision, and 37.2% F1 score, and an AUROC curve of 77.2%. But none of the studies focused on using significant LRFs alongside DRFs, CRFs for the prediction of Glaucoma. This is the gap identified and addressed in this study.

MATERIALS AND METHODS

Instrument for Data Collection

This study used the personal interview method as an instrument to collect data on DRFs, CRFs, and LRFs of Glaucoma at the National Eye Centre (NEC), Kaduna, after being given approval by the ethical committee. The dataset used in this study was from ongoing research, which started in 2024 by the corresponding author on eye diseases. It was collected by the corresponding author from 25th Nov. to 19th Dec. 2024 (Daily observations)

The dataset holds fourteen (14) risk factors, twelve (12) nominal and two (2) numerical. Glaucoma dataset had two (2) DRFs sex and age; six (6) CRFs namely Intraocular Pressure (IOP), Family History of Glaucoma (FHG), Previous Eye Injury (PEI), Decrease Corneal Thickness (DCT), Presence of Diabetes (POD) and High Body Mass Index (HBMI); six (6) LRFs were Taking Low Diet (TLD), Smoking Cigarette (SMC), Too Much Coffee Consumption (TMCC), Inadequate Exercise (IE), Medication with Steroid (MWS) and Too Much Alcohol (TMA). The statistical properties of the dataset and the Glaucoma risk factors description/characteristics are presented in Tables 1 and 2, respectively.

Table 1: Statistical Properties of the Dataset

Properties	**Description**
Number of patients	200
Age group	10-80 years
Gender of patients	Male or Female
Data Collection Process	Interview with patients
Type of Interview	Personal interview
Types of pre-defined questions	Binary closed question (Yes/No)
Total number of risk factors	14
Types of eye disease	Glaucoma

Table 1: Glaucoma Risk Factors Description/Characteristics

		Description/Characteristics of the Risk Factors	
SN	**Glaucoma Risk Factors**	**Name of Risk Factors**	**Type of Risk Factors**
1.	Age	Demographic	Numerical
2.	Sex	Demographic	Nominal
3.	Intraocular pressure level	Clinical	Numerical
4.	Family history of Glaucoma	Clinical	Nominal
5.	Previous eye injury	Clinical	Nominal
6.	Decrease in corneal thickness	Clinical	Nominal
7.	Presence of diabetes	Clinical	Nominal
8.	High body mass index	Clinical	Nominal
9.	Low diet	Lifestyle	Nominal
10.	Smoking cigarette	Lifestyle	Nominal
11.	Too much coffee	Lifestyle	Nominal
12.	Medication with steroid	Lifestyle	Nominal
13.	Inadequate exercise	Lifestyle	Nominal
14.	Too much alcohol	Lifestyle	Nominal

To address potential biases in data collection, the study used bias-mitigation techniques during model development by re-weighting the sample so that under-represented groups have a larger influence on the loss function and use of fairness-aware algorithms to penalize disparate performance across groups, as well as implementation of a detailed data collection protocol, such as preprocessing and handling of outliers.

Population and Sample Size of the Study

The population of the study consists of patients who visited the Glaucoma Department at NEC Kaduna with eye problems for medication from 25th Nov, 2024, to 19th Dec, 2024, and covered patients within the ages of 10-80 years. The study used a non-probability sampling method called volunteer sampling to select a sample size of 200 patients because only those willing to participate would be part of the sample, and limited to 200 patients because it was computed based on the average prevalence of eye diseases. To ensure representativeness of the sample, the study interviewed 35 patients from Kano and Kaduna states; 26 patients each from Katsina, Jigawa, Sokoto, Zamfara, and Kebbi states, which makes up total sample size of 200 patients.

Validity and Reliability of the Instrument

The instrument used was checked for validity using the construct validity method. This method of checking validity employed Pearson`s correlation coefficient (r) given by

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}} \quad (1)$$

where r is the Pearson`s correlation coefficient, n is the number of valid responses, x represents the score of an item, and y represents the total score of each patient with valid responses. A Western & Rosenthal (2003) criterion was applied to interpret Pearson correlation coefficient values for the instruments used to collect the Glaucoma dataset as follows: above 0.35 very beneficial or strongly valid, 0.21-0.35 likely to be useful, 0.11 – 0.20 depends on circumstances, and below 0.11 unlikely to be useful.

To check the reliability of the instrument, the study used Cronbach`s alpha method, which measures the internal

consistency of the items in the instrument. To estimate the internal consistency reliability, Cronbach`s alpha coefficient was used, and is given by

$$\alpha = \left[\frac{n}{n-1}\right]\left[1 - \frac{\sum_{i=1}^{n}\sigma_i^2}{\sigma_X^2}\right] \quad (2)$$

where α is a lower bound estimate of the true reliability, n is the number of items in test X, σ_X^2 is the observed score variance of test X and $\sum_{i=1}^{n}\sigma_i^2$ is the variance of item i. The interpretation of Cronbach`s alpha coefficient was carried out using Cook & Beckman's (2006) criteria as follows: greater than 0.90 excellent reliability, 0.80 - 0.90 good reliability, 0.70 – 0.79 adequate reliability, and below 0.70 less applicable.

Data Preprocessing and Preparation

The main steps of data preprocessing and preparation performed in this study are divided into two main categories: data cleaning and balanced sampling. Data cleaning steps are outlier detection and removal, and missing value handling. For outlier detection, numerical variables are analysed using the interquartile range, and according to this method, outlier was detected in numerical variables. No missing value because it was a face-to-face interview between the patients and two ophthalmologists, a researcher, and a staff member from the research and management information system at NEC Kaduna, and all the patients interviewed gave complete and detailed information.

To handle outliers, the study used the One-sided Winsorization technique by replacing only the upper or lower extreme values with a specified percentile (95th percentile).

The glaucoma dataset was imbalanced because the distribution of the dataset between the classes was not equal. That is, 149(74.5%) of the considered patients belong to Glaucoma (majority class) while 71(35.5%) belong to non-Glaucoma (minority class). A previous study by Krawczyk (2016) showed that the classifiers trained on an imbalanced dataset have higher accuracy for predicting the majority class, and the minority class could not be trained with higher accuracy. To address this problem of imbalanced datasets, the study used the Synthetic Minority Over-sampling Technique (SMOTE).

Data Normalization

Data normalization was performed because the Glaucoma dataset has risk factors that differ in range and unit; this would reduce the model's performance and accuracy, as well as prevent risk factors with larger scales from dominating the learning process. There are different types of data normalization, but this study used min-max to transform risk factors of the datasets to a specified range, usually between zero (0) and one (1), in order to maintain the interpretability of the original values within the specified range. The min-max scaling formula is given by

$$X_{normalized} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (3)$$

where X and $X_{\min}$ indicate the random risk factor value to be normalized and the minimum risk factor value in the dataset, respectively. $X_{\max}$ represent the maximum risk factor value. When X was the minimum value, the numerator became zero ($X_{\min}$ - $X_{\min}$ the normalized value was 0. When X was the maximum value, the numerator is equal to the denominator ($X_{\max}$ - $X_{\min}$) and the normalized value was 1 (Margaret, 2023).

Feature Selection Method

The feature selection method refers to the process of reducing the number of risk factors when using ML models. This study used a feature selection method called Pearson correlation ranked based to select the significant risk factors for the prediction of Glaucoma, because the method focused on selecting the most relevant risk factors that are essential in building models for more accurate prediction (Bustamante-Arias *et al.,* 2021). Pearson correlation $\left(\rho_{X,Y}\right)$ is given by

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \quad (4)$$

where $\operatorname{cov}(X,Y)$ is the covariance between X and Y, and σ is the standard deviation (SDs) on X and Y. $\rho_{X,Y}$ value lies between -1 and +1, where -1 means a negative correlation between X and Y, 0 indicates no correlation between X and Y, and +1 shows a positive correlation between X and Y. The closer the $\rho_{X,Y}$ value is 1, the higher the correlation between X and Y. Then, the class-feature relationship was employed to calculate class-feature correlation values for all the risk factors and ranked them by their correlation values from high to low. For the selection of SRFs, Vidhya (2024) and Mc Elduff *et al*. (2002) criteria were implemented, that is, risk factors with a P-value less than 0.05 (P < 0.05) were chosen as SRFs, and those with a P-value greater than 0.05 (P > 0.05) are not significant.

Training of Machine Learning Models and Hyperparameter Tuning

To train the ML model and perform hyperparameter tuning, the study used supervised learning algorithms and the L1 Regularization Technique to prevent overfitting and improve generalization using Scikit – learn in Python. An instance of the models was created, and the models are trained using the model. fit(x_train, y_train).

Hyperparameter tuning was carried out during training using Grid search because this search defines a set of parameter values to search over, and the algorithm tries all possible combinations. Similarly, the study used a model-centric approach because this approach searches for the optimal combination of hyperparameters

Science World Journal Vol. 20(No 4) 2025 https://dx.doi.org/10.4314/swj.v20i4.14
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

within a predefined set of possible values.

Evaluation of the Model's Performance

The performance of the models was evaluated using measurement performance indices, namely accuracy, sensitivity, specificity, and the ROC curve. Accuracy measures the proportion of cases correctly classified, sensitivity measures the fraction of positive cases that are classified as positive, specificity measures the fraction of negative cases that are classified as negative, and the ROC curve measures the discriminatory ability of the models in distinguishing between eye disease and non-eye disease patients.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

$$Specificity = \frac{TN}{TN + FP} \quad (7)$$

$$AUROC = \int_0^1 TPR(FPR)dFPR \quad (8)$$

where TP is the true positive, TN is the true negative, FP is the false positive, FN is the false negative, TPR is the true positive rate, and FPR is the false positive rate (Fogarty & Bamber, 2005). To ensure higher performance of the models in the light of a limited sample (n = 200), the study makes sure that relevant risk factors were selected and used min-max data normalization to prevent risk factors with larger scales from dominating the learning process.

Cross-Validation of the Models

This study employed cross-validation to validate the performance of the models used for the prediction of Glaucoma. There are several types of cross-validation techniques, such as k-fold cross-validation, leave-one-out cross-validation, holdout cross-validation, stratified cross-validation, but this study used k–fold cross-validation because it maximizes the use of limited data, provides a more robust and reliable performance estimate, and minimizes the risk of overfitting to a particular data split (Mc Elduff *et al.,* 2002).

To use k-fold cross-validation, the sampled data of Glaucoma was randomly partitioned into five and 8 equal-sized sub-samples (i.e., k = 5 and 8). For k = 5, one sub-sample was used for testing and the remaining four (4) equal sub-samples for training. Likewise, for k = 8, one sub-sample was used for testing and the remaining seven sub-samples for training. Then, Erickson & Kitamura's (2021) decision was applied to interpret the cross-validation result. That is, 70%-80% good cross-validation and above 80% perfect cross-validation.

RESULTS

Significant Risk Factors of Glaucoma

The results in Table 3 showed that all fourteen (14) risk factors of Glaucoma were significant (P < 0.05). The fourteen SRFs were two (2) DRFs, six (6) CRFs, and six (6) LRFs.

Table 3: Significant Risk Factors of Glaucoma

Risk Factors	Correlation Value	P-Value
Age	0.3394	8.8375e-07
Sex	0.2612	1.8681e-04
Intraocular pressure level	0.4792	7.0872e-13
Family history of Glaucoma	0.5238	1.7426e-15
Previous eye injury	0.2951	2.2157e-05
Decrease corneal thickness	0.8917	4.2987e-70
Presence of diabetes	0.6687	2.7342e-27
High body mass index	0.6239	5.6391e-23
Low diet	0.3318	1.5842e-06
Smoking cigarette	0.2625	1.7299e-04
Too much coffee consumption	0.4837	3.9965e-13
Medication with a steroid	0.1903	1.0645e-02
Inadequate exercise	0.5019	3.6744e-14
Too much alcohol	0.0873	2.1918e-02

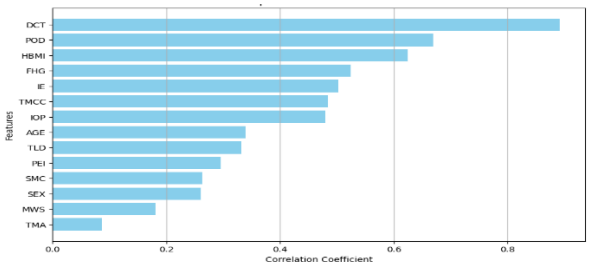

Figure 1: Plot of Pearson Correlation Coefficients for Glaucoma Risk Factors

Splitting of the Glaucoma Dataset into Training and Test Sets

The glaucoma dataset was split into training and test sets, the training set to train the models and the test set to evaluate the models, as done in the work of Lenz (2018), instead of dividing the datasets into training, validation, and test sets as done in other studies, because of the limited dataset. Also, previous studies such as Zhang *et al.* (2022), Saju (2024), and Malik *et al.* (2019) used the ratio 70:30 for data split, but this study employed 80:20 because it gives the best result in terms of model stability and accuracy. The training set contained 80% (160) patients, and the test set contained 20% (40) patients.

Table 4: Splitting of the Glaucoma Dataset into Training and Test Sets

	TRAINING SET			TEST SET		
	GLAUCOMA STATUS			GLAUCOMA STATUS		
	Glaucoma	Non-Glaucoma	Total	Glaucoma	Non-Glaucoma	Total
Count	102	58	160	27	13	40
Percentage	63.8	36.2	100.0	67.5	32.5	100.0

Science World Journal Vol. 20(No 4) 2025 https://dx.doi.org/10.4314/swj.v20i4.14
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Out of 160 patients in the training set, 102(63.8%) were Glaucoma and 58 (36.2%) are Non-Glaucoma. Similarly, in the test set out of 40, 27(67.5%) were Glaucoma and 13(32.5%) are Non-Glaucoma as presented in Table 4.

Estimated Hyperparameters of the Models

The five ML models initially used their default settings so that as each model was adjusted to the data in the training process, the hyperparameters were also adjusted and at the end of the training, the hyperparameters obtained were:
`SVM`: SVC (kernel = `rbf`, C = 1, probability = true, random_state = 42, gamma=scale, degree = 1); ``DT``: DTC (`max_depth = 5, max_ features= none, max_leaf_nodes = 10, min_samples_leaf = 1, min_weight_fraction_leaf = 0.1, splitter = `best`, random_state = 42); ``K-NN``: KNC (leaf_size: 1, n_ neighbours = 5, weights = uniform);``NB``Gaussian NB (var_smoothing = $5.8202481000574e^{-0.8}$); ``MLP``: MLPC (activation = sigmoid, alpha =0.05, hidden_layer_size = (50, 25), learning rate: 0.8, momentum rate = 0.7, max_iter = 300, random_state = 42).

Training Accuracy and Loss of the Models

After estimating the hyperparameter of the models, training accuracy and training loss are also computed. The training accuracy and loss were 0.96, 1.00, 0.96, 0.96 and 1.00; 0.1040, 0.000, 0.1008, 0.5765 and 0.0112 for SVM, DT, K-NN, NB and MLP models respectively.

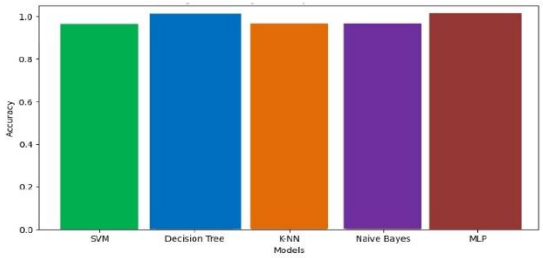

Figure 2: Training Accuracy of the Models

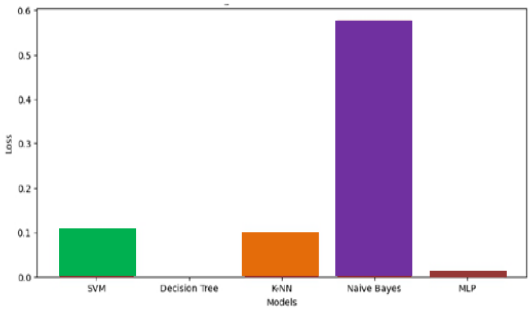

Figure 3: Training Loss of the Models

The plots of training accuracy and loss are presented in Figures 1 and 2, respectively. Similarly, according to Lenz *et al.* (2018), the models have perfect training accuracy; DT, K-NN, and MLP have perfect training loss, SVM has good training loss, and NB has poor training loss.

Confusion Matrix for Prediction of Glaucoma and Non-Glaucoma Patients

The trained ML models are used to predict Glaucoma status in the training and test sets, and the results obtained were shown in Table 5.

Table 5: Confusion Matrix for Prediction of Glaucoma Status

Model	Observed		Classification of Eye Status			
			Glaucoma	**Non-Glaucoma**	**Total**	**Percentage of Predicted Patients**
SVM	*Training set*	*Glaucoma*	*99*	*3*	*102*	*97.1*
		Non-Glaucoma	*3*	*55*	*58*	*94.8*
		Total	***102***	***58***	***160***	
	Glaucoma Test set	*Glaucoma*	*25*	*2*	*27*	*92.6*
		Non-	*2*	*11*	*13*	*84.6*
		Total	***27***	***13***	***40***	
DT	*Training set*	*Glaucoma*	*101*	*1*	*102*	*99.0*
		Non-Glaucoma	*1*	*57*	*58*	*98.3*
		Total	***102***	***58***	***160***	

Science World Journal Vol. 20(No 4) 2025 https://dx.doi.org/10.4314/swj.v20i4.14
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

		Glaucoma	27	0	27	100.0
		Non-Glaucoma	1	12	13	92.3
	Test set	**Total**	**28**	**12**	**40**	
K-NN	Training set	Glaucoma	98	4	102	96.1
		Non-Glaucoma	**3**	**55**	**58**	94.8
		Total	**101**	**59**	**160**	
	Test set	Glaucoma	25	2	27	92.6
		Non-Glaucoma	3	10	13	76.9
		Total	**28**	**12**	**40**	
NB	Training set	Glaucoma	98	4	102	96.1
		Non-Glaucoma	4	54	58	93.1
		Total	**102**	**58**	**160**	
	Test set	Glaucoma	26	1	27	96.3
		Non-Glaucoma	2	11	13	84.6
		Total	**28**	**12**	**40**	
MLP	Training set	Glaucoma	100	2	102	98.0
		Non-Glaucoma	2	56	58	96.6
		Total	**102**	**58**	**160**	
	Test set	Glaucoma	26	1	27	96.3
		Non-Glaucoma	1	12	13	92.3
		Total	**27**	**13**	**40**	

SVM model predicted 97.1% and 92.6% Glaucoma patients, 94.8% and 84.6% Non-Glaucoma patients in the training and test sets. DT model predicted 99.0% and 100% Glaucoma patients, 98.3% and 92.3% Non-Glaucoma patients in the training and test sets. K-NN model predicted 96.1% and 92.6% Glaucoma patients, 94.8% and 76.9% Non-Glaucoma patients in the training and test sets. NB model predicted 96.1% and 96.3% Glaucoma patients, 93.1% and 84.6% Non-Glaucoma patients in the training and test sets. MLP model predicted 98.0% and 96.3% Glaucoma patients, 96.6% and 92.3% Non-Glaucoma patients in the training and test sets, respectively.

Result of Model's Evaluation

The models are evaluated using measurement performance indices, namely accuracy, sensitivity, specificity, and the ROC Curve. SVM model achieved 97.9% accuracy, 96.1% sensitivity, 93.0% specificity, and a 1.00 AUROC curve. The DT model achieved 99.3% accuracy, 99.2% sensitivity, 97.2% specificity, and a 0.95 AUROC curve. The accuracy, sensitivity, specificity, and AUROC curve of the K-NN model were 97.9%, 95.3%, 91.5% and 1.00. The accuracy, sensitivity, specificity, and AUROC curve of the NB model were 96.8%, 96.1%, 91.5% and 1.00. MLP model achieved 98.6% accuracy, 97.7% sensitivity, 95.8% specificity, and 1.00 AUROC. According to Analytic (2021) and TAPS (2005), the models have better accuracy, sensitivity, specificity, and AUROC curve because their performances were above 90% and AUROC curve of 0.95 to 1.00.

Result of 5-Folds and 8-Folds Cross-Validation of the Model's Performance

The models are validated using 5-fold and 8-fold cross-validation. The results showed that for 5-folds the models have 97.7%, 99.1%, 97.6%, 96.6% and 98.5% accuracy; 96.0%, 99.1%, 95.1%, 96.0% and 97.5% sensitivity; 92.8%, 97.0%, 91.3%, 91.3% and 95.6% specificity, and 1.00, 0.95, 1.00, 1.00 and 1.00 AUROC curve. For 8-folds SVM achieved 97.6% accuracy, 96.0% sensitivity, 92.7 specificity, 96.0% and 1.00 AUROC curve; DT achieved 99.0% accuracy, 99.0% sensitivity, 97.0% specificity, and 0.95 AUROC curve; K-NN had 97.5% accuracy, 95.0% sensitivity, 91.2% specificity, and 1.00 AUROC curve; NB achieved 96.6% accuracy, 96.0% sensitivity, 91.2% specificity, and 1.00 AUROC curve; MLP achieved 98.6% accuracy, 97.6% sensitivity, 95.5% specificity, and 1.00 AUROC curve. According to Erickson & Kitamura (2021), the models have better cross-validation results. The results of 5-fold and 8-fold cross-validation are almost consistent with the results of model evaluation, and this shows that the models are adequate

DISCUSSION

This study explored the use of the Pearson Correlation Rank-Based Feature Selection Method (PCRBFSM) to identify SRFs of Glaucoma. The use of PCRBFSM served as an alternative to other

feature selections, such as the variance method, principal component, information gain, and backward stepwise feature selection method in the works of Malik *et al*. (2019), Elsharif & Naser (2022), and Marouf *et al.* (2023). The method of PCRBFSM used in this study computes the correlation values, which could help to know the type of relationship that exists between the SRFs and outputs. The correlation values in Table 3 were positive, and this implies that as the risk factors increase, the risk of having the diseases also increases and vice versa.

Different measures were applied to evaluate models' performance in the previous works. For instance, Malik *et al.* (2019) used nine (9) measures of accuracy, precision, recall, F-Measure, Kappa statistics, mean absolute error, root mean squared error, relative absolute error, and root relative squared error to evaluate the performance of DT, NB, RF, and NN models. Hassan *et al.* (2021) employed three (3) measures, namely accuracy, sensitivity, and specificity to evaluate the performance of the CNN model. Leite *et al.* (2022) applied four (4) measures, which are accuracy, recall, precision, and F1-score. But this study used four (4) measures, namely accuracy, sensitivity, specificity, and ROC curve to evaluate SVM, DT, K-NN, NB, and MLP models' performance because they are the most widely used measures for supervised ML models. The models have perfect accuracy, sensitivity, specificity, and an AUROC curve. Their performance ranged between 92.0% - 99.5% for the prediction of Glaucoma, as well as an excellent AUROC curve in discriminating between Glaucoma and Non-Glaucoma. However, previous studies' performance, such as Leite *et al.* (2022), Hassan *et al.* (2021) ranged between 65.0% - 92.0% with an AUROC curve of less than or equal to 0.95. This demonstrated that the use of LRFs improves the performance of the measures used in this study, unlike previous studies that used DRFs and CRFs.

The results of 5-folds and 8-folds indicated that all five (5) ML models have perfect cross-validation. However, among the models, DT had the best cross-validation then followed by MLP SVM, K-NN, and NB, respectively. Previous studies, such as An *et al.* (2019), used 10-fold cross-validation, Hassan *et al.* (2021) applied 5-fold cross-validation but this study used both 5-folds and 8-folds cross-validation in order to have a robust evaluation of the model's performance.
The 5-fold cross-validation result in this study ranged between 89.1% - 99.1% and 8-folds between 89.0% – 99.4% but previous works, such as An *et al.* (2019), Hassan *et al.* (2021), cross-validation results fall between 85.2% - 94.5% respectively. This showed that these models are adequate and more robust and could be deployed on a new Glaucoma dataset.

Conclusion

This study aimed to use ML models for the prediction of Glaucoma using significant DRFs, CRFs, and LRFs. It was found that all fourteen (14) risk factors of Glaucoma were significant. Five ML models are trained using the SRFs to predict Glaucoma and Non-Glaucoma patients in the training and test sets, evaluated using measurement performance indices, and cross-validated using 5-fold and 8-fold cross-validation. The models have better accuracy, sensitivity, and specificity; an excellent AUROC curve, and perfect cross-validation. The study demonstrated that the use of significant DRFs, CRFs and LRFs could help to predict Glaucoma and Non-Glaucoma patients effectively.

Limitation of the Study

A small sample size n = 200 is the limitation of the study. There is need to check the performance metrics using larger and independent datasets, this would justify the performance of the models using limited sample size (n = 200). Also, there is need to cross-validate the models using larger, independent datasets.

REFERENCES

Anshul, T., Micheal, G., & Siamak, Y. (2020). Predicting Glaucoma Before Onset using Machine Learning. *Journal of American Academy of Ophthalmology*, 21(9):410-419. https://doi.org/10.1136/jaao. 219002. 410-9.

An, G., Omodaka, K., Hushimota, K., Tsuda, S., Shiga, Y., Takada, N., Kikawa, T., Yokota, H., Akiba, M., & Nakazawa, T. (2020). Glaucoma Diagnosis with Machine Learning Based on Optical Coherence Tomography and Colour Fundus Images. *Hindawi Journal of Health Care Engineering*, 120(18), 251-260. Doi: 10. 115/2020/4061313.

Bustamante – Arias, A., Cheddad, A., Jimenez – Prezez, J.C., & Rodriguez – Garcia, A. (2021). Digital Image Processing and Development of Machine Learning Models for the Discrimination of Corneal Pathology: An Experimental Model. *Phonics Journal,* 8(4): 118 – 124. https://doi. org / 10. 3390 / phonics. 8040118.

Cook, D.A., & Beckman, T.J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *American Journal of Medicine*, 119: 166-175.

Elsharif, A.A.E.F., & Naser, S.S.A. (2022). Retinal Diseases Diagnosis using Deep Learning. *International Journal of Academic Engineering Research(IJAER)*, 6(2): 11-37. ISSN: 2643-9085. https://doi.org/ 10.1109/ijaer.2017.8261463.

Erickson, J.B., & Kitamura, F. (2021). Nine Performance Metrics for Machine Learning Models. *Journal of Radiol Artif intell,* 3(3). Retrieved from https://doi. Org/10.1148/ryai.2021200126.

Fei, L., Yuandong, S., Fengbin, L., Zhihuan, L., Yunhe, S., & sheng, N. (2022). Artificial Intelligence Predict Glaucoma Incidence and Progression using Retinal Photographs. *Journal of Clinical Investigation*, 132(11), 960-968. Doi: 10.1172|jci 157968.

Fogarty, E., & Bamber, D. (2005). Area Above the Ordinal Dominance Graph and Area Below the Receiver Operating Characteristic Curve. *Journal of Maths Psychology*, 12, 387-415.

Hassan, K., Tanha, M.D., Amin, T., Faruk, R.M.D., Khan, O., Aljahdaili, M.M., & Masud, S. (2021). Cataract Diseases Detection by using Transfer Learning – Based Intelligent Methods. *Handawi Journal of Computational and Mathematical Methods in Medicine*, 202(14): 121-132. https://doi.org/ 10. 1155/2021hjcmmm/7666365.

Krawczyk, B. (2016). Learning from Imbalanced Data: Open Challenges and Future Directions. *Journal of Progress in Artificial Intelligence*, 5(4), 221-32. Doi : 10.1310/jpai.54221-32.

Leite, D., Campelos, M., Fernandes, A., Batista, P., Beirao, J.,

Menetres, P., & Cunha, A. (2021). Machine Leaning Automatic Assessment for Glaucoma and Myopia Based on Corvis ST Data. *Journal of Procedia Computer Science,* 196(20): 454-460. Doi: 10.1016/jpcs2021. 12. 036.

Lenz, A.L.(2018). Development and Comparison of Machine Learning Methods for Subjective Refraction Prediction. Work Presented in Partial Fulfilment for the Degree of Bachelor in Computer Science, University of Do Rio Grande Do Sul.

Mahyar, S., Toktan, K., Mohammad, H. E, Somaveh, S., Hassan, H., & Akbar, F. (2021). Development of Glaucoma Predictive Model. *Journal of BioData Minning*, 48(14), 754 -759. https://doi. org/ 10.33900/ jbiodata . 102354. 754 -9.

Malik, S., Kanwal, N., Asghar, N.M., Sadiq, A.A.M., Karamat, I., & Fleury, M. (2019). Data Driven Approach for Eye Diseases Prediction with Machine Learning. *Journal of Applied Science*, 202(9), 2789-2793. Doi: 10. 3390/japps 914278993.

Margaret, R. (2023). Machine Learning. Retrieved from https://www. technopedia.com

Marouf, A.A., Mottalib, M.M., Alhaji, R., Rokne, J., & Jafarullah, O., (2023). An Efficient Approach to Predict Eye Diseases from Systems using Machine Learning and Ranker Based Feature Selection Methods. *Journal of Bioengineering*, 10 (25), 390 – 39. Doi: 10 3390/jbioergin. 202310010025.

Mc Elduff, P., Attia, J., Ewald, B., Cockburn, J., & Heller, R. (2002). Estimating the Contribution of Individual Risk Factors to Disease in a Person with More Than One Risk Factor. *Journal of Clinical Epidemiology*, 55(6): 588-592. https://doi.org/10.1016/SO895-4356(02) 00388-8.

National Institute of Health (2023). Eye Diseases. Retrieved from https://www.ncbi.nlm.nih.gov www.ncbi.nlm.gov

National Institute of Health (2024). Refractive Error Eye Disease. Retrieved from https://www.ncbi.nlm.nih.gov www.ncbi.nlm.gov

Pascolini, D., & Mariotti, S.P. (2010). Global Estimates of Visual Impairment. *British Journal of Ophthalmology*, 96(5), 614-618. https://doi.org/ 10.1136/bjophthamol- 20011-300539.

Ravindranath, R., Naor, J. & Wang, S.Y. (2025). Artificial Intelligence Models to Identify Patients at High Risk for Glaucoma using Self-Reported Health Data in a U.S. National Cohort. *Journal of Ophthalmology,* (2025): *664-685.* https://doi. org/ 10.1016/j.XOPS.2024.100685.

Saju, B., & Rajesh, R. (2024). Cataract Risk Factors Prediction using Deep Learning Models. *Research Square Journal*, 1(1): 1234-1242. https://doi.org/10.21203/rs.3.rs3081019/v1.

Santos, D.F., Dosul, R.C., & Hamburgo, N.R. (2023). Classifying Glaucoma using Machine Learning Techniques. *Journal of Medical Research,* 28(4), 2328–2336.Doi: 10.1101/2023.05.02.23289378.

Shuldiner, R.S., Boland, V.M., Ramalu, Y.P., De Moreies, G.C., Elze, T., Myers, J., Pasquale, L., Wellik, S., & Yohannan, J. (2021). Predicting Eye at Risk for Rapid Glaucoma Progression Based on an Initial Visual Field Test using Machine Learning. *Journal of Pone,* 16(4): 258-266. https://doi. org/ 10. 10. 1371/journal. pone. 0249856.

Traditional Academic Point System. (2005). Interpretation of Kappa Statistic Value and Receiver Operating Characteristic Curve. Retrieved form https://www.medium.com

Vidhya, A. (2024). Feature Selection in Machine Learning. Retrieved from https://www.analytic vidhya.com

Wang, W., Hang, X., Zhang, J., Shang, X., Ha, J., Liu, Z., Zhang, L., Luo, L., & He, M. (2020). Predicting the 10- year Risk of Cataract Surgery using Machine Learning Techniques on Questionnaire Data: Findings from the 45 and up Study. *British Journal of Ophthalmology*, 106(28): 318- 329. https:// doi . org/ 10.1136/ bjophthalmol – 138609.

Western, D. & Rosenthal, R. (2003). Quantifying Construct Validity: Two Sample Measures. *Journal of Personality and Social Psychology, 84:* 608-620.

Zhang, X-Q., Hu, Y., Xiao, Z-J., Fang, J-S., Higashita, R., & Liu, J. (2022). Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey. *Journal of Machine Intelligence*, 19(2), 184-208. Doi: 10.1007/s11633-022-1329-0.