International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 228–241

Article

Effect of Hyperparameter Optimization on Artificial Learning Model Performance: Detection of Diabetes Retinopathy

Fatma Önay KOÇOĞLU

Artificial Intelligence and Data Engineering Department, Faculty of Computer and Information Technologies, İstanbul University, İstanbul 34134, Türkiye; fonayk@istanbul.edu.tr; Tel.: +90-212-4400000; [0000-0002-1096-9865]

Received date: 11 July 2024; Accepted date: 4 March 2025; Published online: 27 March 2025

Abstract: Smart systems gain importance when both the importance of early diagnosis and treatment for the patients and the costs associated with the use and maintenance of medical devices are considered. In this direction, it is necessary to obtain high-performance models for more effective smart systems. It is aimed to propose an effective hybrid prediction model in which hyperparameters are automatically determined and to lay the groundwork for a clinical decision support system to be developed for detection of Diabetes Retinopathy. Prediction models have been developed with Extreme Learning Machine and Support Vector Machine algorithms, and parameter optimization has been carried out by hybridizing these algorithms with Genetic Algorithm. ELM, which is thought to be open to development and give effective results in terms of application, has been especially preferred. ELM achieved the highest accuracy with a value of 74.49%. The parameters of this model are as follows: the number of neurons is 180, the activation function used is tan-sigmoid, and the threshold value used to determine the class label is 0.427. According to different performance evaluation measures, the developed models are more successful than the studies carried out with the same data set in the literature.

Keywords: diabetes; extreme learning machine; genetic algorithm; hyperparameter optimization; support vector machine

1. Introduction

Diabetes is a chronic disease that occurs because of insufficient production of the insulin hormone that regulates blood sugar by the pancreas or inability to use the produced insulin effectively [1]. According to the 10th International Diabetes Federation (IDF) Diabetes Atlas [2] published on December 6, 2021, 537 million adults worldwide live with diabetes, and this number is predicted to rise to 643 million by 2030. This disease caused 6.7 million deaths in 2021 and an increase of at least 966 billion dollars in health expenditures in the last 15 years. Over time, diabetes damages many systems in the body, especially the nerves and blood vessels, and causes permanent damage in advanced dimensions. Some complications seen in patients with diabetes can be counted as macrovascular stroke, ischemic heart disease and peripheral arterial disease and/or microvascular retinopathy, neuropathy and nephropathy [3]. These complications significantly affect the quality of patient's life. Diabetes retinopathy (DR) is one of the most important complications that can result in vision loss. Early detection of DR is essential to prevent blindness caused by this disease as vision may not be affected in early retinopathy [4,5]. A strategy used to prevent disease progression in pre-DR or early-stage DR is to strictly monitor and control various risk factors [6].

In addition to its importance for human health and quality of life, the importance of early diagnosis and treatment in health systems should be addressed. Considering the number of patients in the system, the use and maintenance of medical devices, and the costs, this whole cycle can be considered as an industrial engineering problem. Considering the possibility of increasing chronic diseases such as diabetes, cancer, cardiovascular diseases, dementia, and obesity, it is necessary to keep the health expenses at a certain level while ensuring the normal course of life standards of the patients. Therefore, the correct diagnosis and treatment methods should be determined quickly. On the other hand, since it will be important to develop medical decisions, treatments and various health practices in a way that is specific to individuals, it is also necessary to model the complexity of the patients well [7]. Knowledge and experience are one of the most important factors in making effective decisions in a successful patient diagnosis and treatment process. The concept of experience and knowledge is also central to understanding the effects of artificial intelligence on health [8]. Various predictive and descriptive models can be developed using previous patient data and appropriate algorithms with machine learning, where learning occurs based on historical data. These models are transformed into applications for early diagnosis and treatment of disease, smart systems for patient follow-up, and wearable technologies with programming languages and software tools, and are widely used in the field of health. In this respect, it is quite clear how beneficial it will be to introduce AI-based decision support systems, which integrate the data collected from patients and the models obtained with the machine learning process, to the use of physicians and speed up the decision processes.

This study has been handled from two different perspectives. In this context, it is aimed to develop a machine learning-based prediction model for the prediction of DR disease in terms of health sciences and to perform hyper parameter optimization to increase the success of the prediction model in terms of computer sciences. In this direction, the remainder of this study is organized as the literature review, the aim of the study and the sharing of the model development process, the comparative analysis of the performance indicators of the developed models and the interpretation of the results obtained.

2. Background of the Study

There are many studies based on machine learning in the literature for detecting DR.

Vijayan et al. [9] developed models to predict DR disease using decision trees and sample-based classifiers with retinal images. In the analyses performed by WEKA, performance was measured using accuracy and weighted average ROC values. Reddy et al. [10] developed an ensemble-based model consisting of Random Forest, Decision Tree, Adaboost, K-Nearest Neighbor, Logistic Regression machine learning algorithms for the detection of DR and revealed that the developed model outperformed individual machine learning algorithms. Ali et al. [11] measured the performance of machine learning algorithms for DR segmentation and classification in their study using two-dimensional retinal fundus images. A fused hybrid-feature dataset was obtained using the data fusion approach to increase the classification performance. Gadekallu et al. [12] used the standard scalar method for normalization, Principal Component Analysis (PCA) for feature selection, and the Firefly algorithm for size reduction on the data set obtained from UCI. The classification model has been created with Deep Learning. With the developed model, successful results were obtained according to various performance measures. Gadekallu et al. [13] used a PCA-based Deep Neural Network model using the Grey Wolf Optimization (GWO) algorithm for DR diagnosis. Machine learning algorithms such as Support Vector Machine (SVM), Naive Bayes Classifier, Decision Tree and XGBoost have been used to compare the model performance. David [14] used Adaptive Histogram Equalization to convert colour images to grayscale images and Hybrid Colour and Structure Descriptor (HCSD) for feature selection. Finally, various machine learning methods such as Hybrid Radial Basics Kernel based Support Vector Machine have been used to classify images for DR detection. Afian et al. [15] proposed a prediction model in which Recursive Feature Elimination (RFE) is used for the early detection of DR and a Deep Neural Network (DNN) is used for classification, based on individual risk factors.

Hasan et al. [16] measured the performance of machine learning-based DR detection and classification systems. The machine learning models in these systems were trained and tested using retinal fundus and thermal images in open datasets. As a result of the study, the ResNet50 deep convolutional neural network was determined to be the most effective algorithm. In the study of Math and Fatima [17], a segment-based learning approach was proposed for the detection of diabetic retinopathy. In the analyses performed using the dataset in Kaggle, better results were obtained than the existing models according to the sensitivity and specificity values. Mahmoud et al. [18] proposed a hybrid inductive machine learning algorithm (HIMLA) for automatic DR diagnosis. In the study using the CHASE datasets, the algorithm classifies color fundus images as healthy and unhealthy. Normalization and segmentation were carried out to increase the quality of the images used. Nagi et al. [19] combined machine learning methods for DR detection and proposed the Two Stage Classifier method as an ensemble learning method. Gayathri et al. [20] performed feature extraction from fundus images and proposed an automated DR grading method using deep learning, Support Vector Machine (SVM), Random Forest, and J48 machine learning methods. Emon et al. [21] aimed to compare the success of Naive Bayes, Sequential Minimal Optimization (SMO), Logistic Regression, Stochastic Gradient Descent (SGD), bagging, Decision Tree-based algorithms in DR detection. In the study, feature extraction analysis was also performed to determine the features that are important for detecting DR. Liu et al. [22] performed DR detection in Optical Coherence Tomography Angiography (OCTA) images using four machine learning methods, namely Logistic Regression, Elastic Net Penalty, Support Vector Machine and XGBoost. A discrete wavelet transform was also applied to extract texture features from images. Alabdulwahhab et al. [23] investigated the success of Linear Discriminant Analysis, Support Vector Machine, k-Nearest Neighbor, Random Forest and Ranger Random Forest machine learning methods in DR detection using socio-demographic and clinical data collected by systematic random sampling method. This study presents a new methodology and a computerized diagnostic system. Murthy and Arunadevi [24] emphasized that the development of DR can be detected by its growth in retinal blood vessels, and conducted a study to automatically detect this growth with a machine learning-based approach. In the study, in which image data was used, Kinetic Gas Molecule Optimization based on centroid initialization was used for the Fuzzy C-means Clustering in the segmentation phase and Convolution Neural Network with Bidirectional-Long Short-Term Memory (CNN with Bi-LSTM) hybrid method was used for classification. Odeh et al. [25] developed an ensemble-based learning strategy that combines a significant selection of well-known classification algorithms into a single complex diagnostic model to improve the prediction performance for the dataset frequently used in the literature. While developing the models, analyses were carried out for different attribute groups selected from among all attributes. Devi et al. [26] used the Optimized Backpropagation Neural Network (Op-BPN) algorithm to increase the model success for DR diagnosis. Besides predicting success, it is also aimed to shorten the time taken for disease diagnosis. Abreu et al. [27] aimed to comparatively measure the performance of various data mining algorithms for DR detection under different scenarios and data sampling methods. Algorithms such as Logistic Regression, autoMLP, SVM, JRip, NN were used in the analysis. Shankar et al. [28] proposed the Automatic Hyperparameter Tuning Inception-v4 (HPTI-v4) model for the detection and classification of DR using color fundus images. The model incorporates various sub-processes such as preprocessing, segmentation, feature extraction, and classification. It uses the Inception-v4 architecture for feature extraction and a multilayer perceptron (MLP) for classification. The dataset used in the HPTI-v4 model for DR fundus image classification is the MESSIDOR dataset. Modaresnia et al. [29] used three different image enhancement techniques, including contrast enhancement, color constancy, and contrast-limited adaptive histogram equalization (CLAHE), to determine DR and its severity. Three different pre-trained convolutional neural networks, namely AlexNet, GoogLeNet and SqueezeNet, were investigated on the binary and multiple classification of DR fundus images. The Bayesian optimization method was used for hyperparameter tuning. To enhance the robustness and sensitivity of the detection process of DR, Bilal et al. [30] used Hierarchical Block Attention (HBA) and HBA-U-Net architecture, which advances attention mechanisms, especially in image segmentation. This model improves image processing without imposing excessive computational demands by focusing on individual pixel complexities, spatial relations, and channel-specific attention. The method proposed by Luo et al. [31] incorporates the correlations between long-range patches into the deep learning framework to improve DR detection. Since DR lesions usually appear as plaques, patch-based relationships are used to enhance local patch features. The Long-Range unit with a residual structure in the proposed network can be flexibly embedded in other trained networks. Mehedi Shamrat et al. [32] proposed a convolutional neural network model (DRNet13) for automatic diabetic retinopathy classification. They particularly used image preprocessing methods such as Median filter and Gamma correction for noise reduction and image enhancement.

Problem Definition

The literature review can be extended with similar studies. When the studies in the literature are examined, it is seen that especially image data is used and proposals are made regarding the pre-processing process of the data for a better classification. On the other hand, deep neural network architectures are focused on, but the hyperparameter tuning process remains in the background. In this study, a structural data set related to the disease is used instead of unstructured image data. Advanced machine learning models, such as artificial neural networks and support vector machines, have a large number of hyperparameters. Therefore, determining these hyperparameters is also a separate and important research topic. Within the scope of this study, both the success of a different neural network structure in solving the problem is examined and the focus is on hyperparameter optimization to increase the performance.

As can be seen from the literature, although similar algorithms are used in many studies, researchers aimed to increase the performance of the machine learning model with different model designs. Because developing a more successful model means producing more effective and faster solutions for the patient in the field of health sciences, and better management of time and financial resources. The size and structure of the dataset, the design of the training, test, and validation datasets, data pre-processing, etc. affect the performance of the machine learning model. For this reason, even if the same algorithms or data sets are used in different models, different performance indicator values can be obtained due to the design of the model. Model performance can also be affected by the hyperparameter of the learning algorithm. Therefore, one of the research problems is hyperparameter tuning for artificial learning. In the field of engineering, different problems are considered as optimization problems and solutions are produced. In machine learning studies, solutions to hyperparameter tuning problems can be produced using optimization algorithms.

This study focused on the effect of determining the parameter values to increase the model performance. Also, it is aimed to lay the groundwork for a clinical decision support system to be developed by presenting a machine learning model with higher performance for the detection of diabetes retinopathy disease, which causes serious complications. Accordingly, prediction models have been developed using the Extreme Learning Machine (ELM) and Support Vector Machine (SVM) algorithms. Parameter optimization has been carried out by hybridizing these algorithms with the Genetic Algorithm (GA). In addition to the main objectives mentioned above, investigating the effect of ELM on problem solving was determined as a sub-target. ELM is an Artificial Neural Network (ANN)-based algorithm that is easy to implement, produces effective results, and the hyperparameters can be adjusted more easily than ANN. Recently, it has been tried to increase its success by establishing hybrid models with different optimization algorithms, as well as taking part in the solution of different problems. In this direction, the study generally covers computational and comparative analyses to increase the success of the prediction model.

To summarize, within the scope of the study, the following is aimed:

- Developing a model with higher performance than similar studies in the literature for DR detection,
- Investigating the usability of the Extreme Learning Machine (ELM) algorithm for solving the problem
- Performing hyperparameter tuning with a hybrid approach using the GA, ELM and SVM algorithms,
- To perform comparative performance analysis.

3. Materials and Methods

3.1. Data

The Diabetic Retinopathy Debrecen Data Set was used to train and test the models, and the data set was obtained from the UCI Machine Learning Repository. The reason why this dataset is preferred is to increase the comparability of the results of this study with the literature and to ensure that the proposed model can be tested by different users. A total of 19 attribute fields, including 1 output and 18 input attributes, and 1151 records in the dataset. The data set is suitable for binary classification as 0: no signs of DR and 1: contains signs of DR (Accumulative label for the Messidor classes 1, 2, 3). 46.92% (540 records) of the records belong to the class 0, 53.08% (611 records) belong to the class 1, and there is a balanced class label distribution in the data set. The attribute fields are briefly described in Table 1.

Table 1. Description of the attributes.

Attribute Number	**Attribute Data Type**	**Explanation**
0	binary	The binary result of the quality assessment. 0 = bad quality, 1 = sufficient quality.
1	binary	The binary result of pre-screening, where 1 indicates severe retinal abnormality and 0 indicates its lack.
2-7	numerical	The results of microaneurysms (MAs) detection. Each feature value stands for the number of MAs found at the confidence levels alpha = 0.5, . . . , 1, respectively.
8-15	numerical	Contain the same information as Attribute 2-7 for exudates.
16	numerical	The euclidean distance of the center of the macula and the center of the optic disc.
17	numerical	The diameter of the optic disc.

18	binary	Class label. 1 = contains signs of DR (Accumulative label for the Messidor classes 1, 2, 3), 0 = no signs of DR.

In this study, all values were subjected to linear data transformation using Equation (1).

$$x_{normal\ value} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

3.2. Performance Validation

To train the model and measure the performance of the developed model, the dataset is divided into training and testing. This separation process was performed using the Hold out method, as the model performance would be affected. The hold-out method allows the dataset to be divided into a specified number of independent data sets within certain proportions and randomly. In this method, while 2/3 of the data set is separated as the training data set and 1/3 as the test data set, the training data set is used for creating the model and the test data set is used for measuring the performance of the model. In the scope of this study, the hold-out performance validation method was used (70% as training set-30% as test set).

3.3. Modelling with Extreme Learning Machine

The Extreme Learning Machine was developed by Huang et al. [33] and is an Artificial Neural Network-based algorithm. In this method, input weights and threshold values are randomly generated, and analytical methods are used to calculate the output weights. In this way, it is aimed to accelerate the learning process. It provides advantages in terms of computational ease and performance due to the absence of back propagation and repetitive steps, the analytical calculation of output weights, and the reduced intervention of the designer with fewer parameters compared to ANN [34,35].

(x_i, y_i) are different from each other and randomly selected samples, the input is represented by $x_i = \{x_{i1}, x_{i2}, \dots, x_{in}\}^T$ and output by $y_i = \{y_{i1}, y_{i2}, \dots, y_{im}\}^T$.

$$\sum_{i=1}^{\tilde{N}} \beta_i g\left(w_i x_j + b_i\right) = y_j \qquad (j = 1,2,\dots,N) \quad (2)$$

is a mathematical model for Single Hidden Layer Feed-Forward Neural Network (SLFN). Here Ñ is the number of neurons, Ñ<N, $g(x)$ is the activation function, $w_i = \{w_{i1}, w_{i2}, \dots, w_{in}\}^T$ is weight vector associated with the input and hidden neuron *i*, $\beta_i = \{\beta_{i1}, \beta_{i2}, \dots, \beta_{im}\}^T$ is weight vector associating the hidden neuron *i* with output, b_i is threshold value for the hidden neuron *i*.

The initial input weights and biases are randomly assigned during the model training phase and an *H* matrix is obtained as follows. Equation (2) is expressed as follows, where β is the weight vector and *T* is the output vector.

$$H\beta = T$$

$$H = \begin{bmatrix} g(w_1 x_1 + b_1) & \cdots & g(w_{\tilde{N}} x_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & \cdots & g(w_{\tilde{N}} x_N + b_{\tilde{N}}) \end{bmatrix}_{N\times\tilde{N}} \quad (3)$$

$$\beta = \begin{bmatrix} \beta_1^{\,T} \\ \vdots \\ \beta_{\tilde{N}}^{\,T} \end{bmatrix}_{\tilde{N}\times m} \qquad T = \begin{bmatrix} t_1^{\,T} \\ \vdots \\ t_N^{\,T} \end{bmatrix}_{N\times m}$$

It is aimed to maximize the neural network's performance, so the error is expected to be 0 (zero) or minimum.

$$\left\|H\bar{\beta} - T\right\| = min\|H\beta - T\| \quad (4)$$

At the end of the training phase, the vector that minimizes the error is obtained.

$$\bar{\beta} = H^{t}T \quad (5)$$

3.4. Modelling with Support Vector Machine

SVM works on the principle of Structural Risk Minimization, which minimizes the upper bound of errors due to generalization [36]. SVM is used for linear or nonlinear classification problems. In the linear classification problem with SVM, it is aimed to find an optimal hyperplane to distinguish the classes best. The distance between the points located in the parallel planes of this hyperplane and referred to as the support vectors must be maximum.

In Figure 1, the optimal separation plane is H_0, with the boundary planes H_1 and H_2, which are considered to have the maximum distance between the two classes. For the boundary and optimal separation planes specified with the weight vector $W = \{w_1, w_2,...,w_n\}$ and constant b, the region under the optimal separation hyperplane is represented by the inequality (6) and the region above it by the inequality (7).

$$W^T X + b < 0 \quad (6)$$

$$W^T X + b > 0 \quad (7)$$

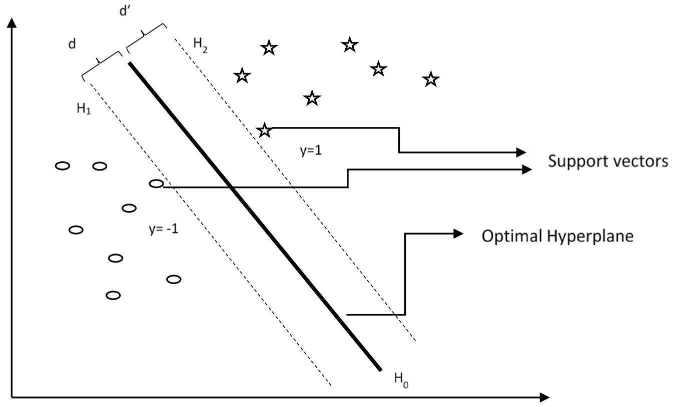

Figure 1. Support Vector Machine.

The distance between the support vectors of the different class attributes in the boundaries is indicated by (10).

$$d=\frac{|W^T X-b|}{||w||} \quad (8)$$

$$d'=\frac{|W^T X+b|}{||w||} \quad (9)$$

$$d+d'=\frac{2}{||w||} \quad (10)$$

Since the distance between the support vectors of the two classes specified in equation (10) is aimed to be maximum, the denominator in this equation should be minimized. Based on this, the problem must be solved under the specified objective function (11) and constraints (12).

$$z_{min} = \frac{1}{2} w^T w \quad (11)$$

$$y_i\,(w^T x_i + b)\ \geq 1 \quad (12)$$

Accordingly, the objective function and constraints to be obtained are remodeled by the Lagrangian function and the Karush-Kuhn-Tucker (KKT) conditions and solved as an optimization problem. The relevant calculations can be examined in [37].

In nonlinear classification problems, to achieve this goal, the dataset is carried to a higher dimensional data space where it can be separated linearly by nonlinear mapping methods [38]. To simulate the projection of the data in the conversion process, the kernel functions are used [39].

3.5. Genetic Algorithm Approach for Hyperparameter Optimization

While the machine learning model is being developed, it has been stated in the previous sections that the model performance may change depending on many factors. One of these factors is the parameter tuning of the algorithms. Many algorithms used to develop models in machine learning have some specific parameters. The algorithm parameters that can be determined by the model developer are called hyperparameters. One of the important fields of study in the field of machine learning is hyperparameter optimization. Hyperparameter optimization is expressed as [40]:

Let A be a machine learning algorithm with N (1,2,3,…,n) different hyperparameters. The hyperparameter combination space of the algorithm is $\Lambda = 1 \times 2 \times \ldots \times n$. Let A_λ be the model of algorithm A obtained using λ hyperparameter values, λ is a hyperparameter vector in Λ combination space.

D is a data set with hyperparameter optimization. The goal is to find the hyperparameter vector λ* that will minimize the loss in model performance:

Here, V is the function in which the performance loss in the model is calculated according to the determined indicators.

One of the methods used in hyperparameter optimization is the grid search approach. With this approach, the entire combination space is scanned to obtain the optimum parameter combination. However, the grid search approach can be disadvantageous in terms of processing capacity and time due to the high number of parameters, dependency, and wide parameter value range. On the other hand, meta-heuristic methods are used in hyper-parameter optimization as in various optimization problems. One of these methods is the Genetic Algorithm (GA). Genetic algorithm is a widely used, high-performance algorithm based on the random search approach, which is useful in complex optimization problems where the number of parameters is large or the analytical solution is difficult [41]. According to the principles of the theory of evolution, this algorithm is based on the application of the rules of genetics and natural selection to keep living things adapting to the environment and to eliminate those that cannot. The detailed information can be examined in [42]

GA steps can be expressed in general as follows:

- The initial population P containing the q solutions is chosen as the current population.
- A fitness value is calculated for each solution in the P population.
- To create a matching pool, solutions are selected from P using the fitness value.
- Crossover, mutation, and other selected operations are applied to randomly selected solutions in the matching pool to generate a new population.
- It is checked whether the obtained solution is the best, and if it is not the best solution, all steps are returned.

The operators used for the selection, crossover, and mutation mentioned above are the main operators of GA, one of the population-based stochastic algorithms [42]. GA-based hyperparameter optimization has been used to determine the necessary hyperparameter values in order to optimize the model success in models developed with ELM and SVM.

3.6. Parameter Tuning for The Algorithms

The developed machine learning models and the hyperparameter value ranges used for their optimization are given in Table 2.

Table 2. Used parameter values.

Algorithm	Parameters
ELM	Neuron number: Integers in [1,1000] Activation function: Hard-limit, radial basis, sigmoid, sine, symmetric hard-limit, tan-sigmoid, triangular basis, linear Threshold value: [0.4, 0.8]
SVM	Kernel function: Radial-based and sigmoid Cost value: $[2^{-12}, 2^{12}]$ (The exponential values of 2) Gamma value: $[2^{-9}, 2^{9}]$ (The exponential values of 2)
GA	Type: Real valued Population size: 10, 20, 30, 50 Max Iteration: 20, 30, 50 Crossover: 0.6, 0.7, 0.8, 0.9, 1.0 Mutation: 0.0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1

3.7. Performance Evaluation

For the model performance evaluation, the accuracy and F-score values have been considered. To calculate these measures, the confusion matrix is used. This matrix has been given in Table 3.

Table 3. Confusion matrix.

		Actual Class		
		Positive	Negative	Evaluation Measure
Predicted Class	Positive	True Positive (TP)	False Positive (FP)	Precision: $\frac{TP}{TP+FP}$
	Negative	False Negative (FN)	True Negative (TN)	Negative Prediction Value: $\frac{TN}{FN+TN}$
	Evaluation Measure	Sensitivity: $\frac{TP}{TP+FN}$	Specificity: $\frac{TN}{TN+FP}$	Accuracy: $\frac{TP+TN}{TP+FP+FN+TN}$

Within the scope of the study, the highest values of accuracy and F-score were calculated separately. The fitness value used in the GA has been developed for this purpose. Models were run first to find the best accuracy and then for the best F-score.

Figure 2 shows a graphical summary of the modelling process.

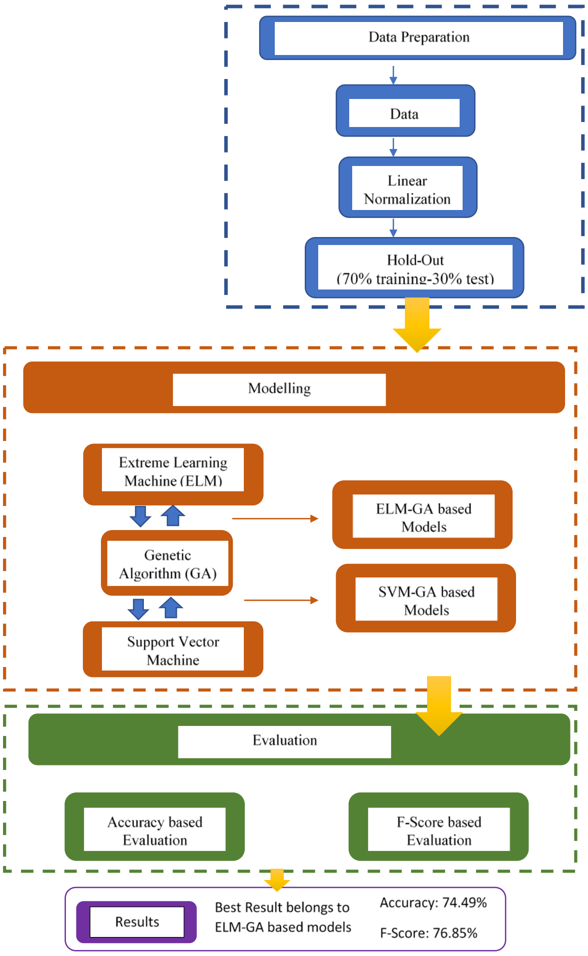

Figure 2. Graphical abstract of the proposed model.

4. Results

To model development and for visualization R programming language and RStudio editor have been preferred [43,44]. Within the scope of the study, the packages needed are as follows.: For reading and printing data set - xlsx [45], for data transformation – clusterSim [46], for hold-out performance validation method- caret [47], for Support Vector Machine - e1071 [48], for Extreme Learning Machine - elmNN [49], and for Genetic Algorithm – GA [50].

The models were evaluated according to the Accuracy and F-score measures. The highest performance indicator values and related parameters are given in Table 3 for Accuracy and in Table 4 for the F-score.

Table 4. Results of the models according to accuracy.

Classification Algorithm	**Genetic Algorithm Parameters**	**Classification Algorithm Parameters**	**Accuracy**
ELM	Population size: 50 Max Iteration: 50 Crossover: 0.7 Mutation: 0.0	Neuron number: 180 Activation function: Tan-sigmoid Threshold value: 0.427	74.49%
SVM	Population size: 10 Max Iteration: 20 Crossover: 0.6 Mutation: 0.0	Kernel function: Radial Cost value: 128 (2^7) Gamma value: 0.015625 (2^{-6})	73.91%

According to Table 4 , even though the accuracy values achieved with SVM and ELM are very close to each other, ELM achieved the highest accuracy with a value of 74.49%. In the model obtained with ELM, the number of neurons was 180, the activation function used was tan-sigmoid, and the threshold value used to determine the class label was 0.427.

According to Table 5, the difference between the success rates of SVM and ELM increased slightly. The model obtained using the ELM achieved the highest success with the F-score of 76.85% and surpassed the model performance obtained using the SVM by approximately 4%. In the model obtained with ELM, the number of neurons was 165, the activation function used was tan-sigmoid, and the threshold value used to determine the class label was 0.444.

Table 5. Results of the models according to the F score.

Classification Algorithm	**Genetic Algorithm Parameters**	**Classification Algorithm Parameters**	**F-score**
ELM	Population size: 30 Max Iteration: 30 Crossover: 0.5 Mutation: 0.04	Neuron number: 165 Activation function: Tan-sigmoid Threshold value: 0.444	76.85%
SVM	Population size: 30 Max Iteration: 20 Crossover: 0.6 Mutation: 0.0	Kernel function: Radial Cost value: 512 (2^9) Gamma value: 0.0078125 (2^{-7})	72.94%

The comparative performance indicator values of the developed artificial learning models are given in Figure 3.

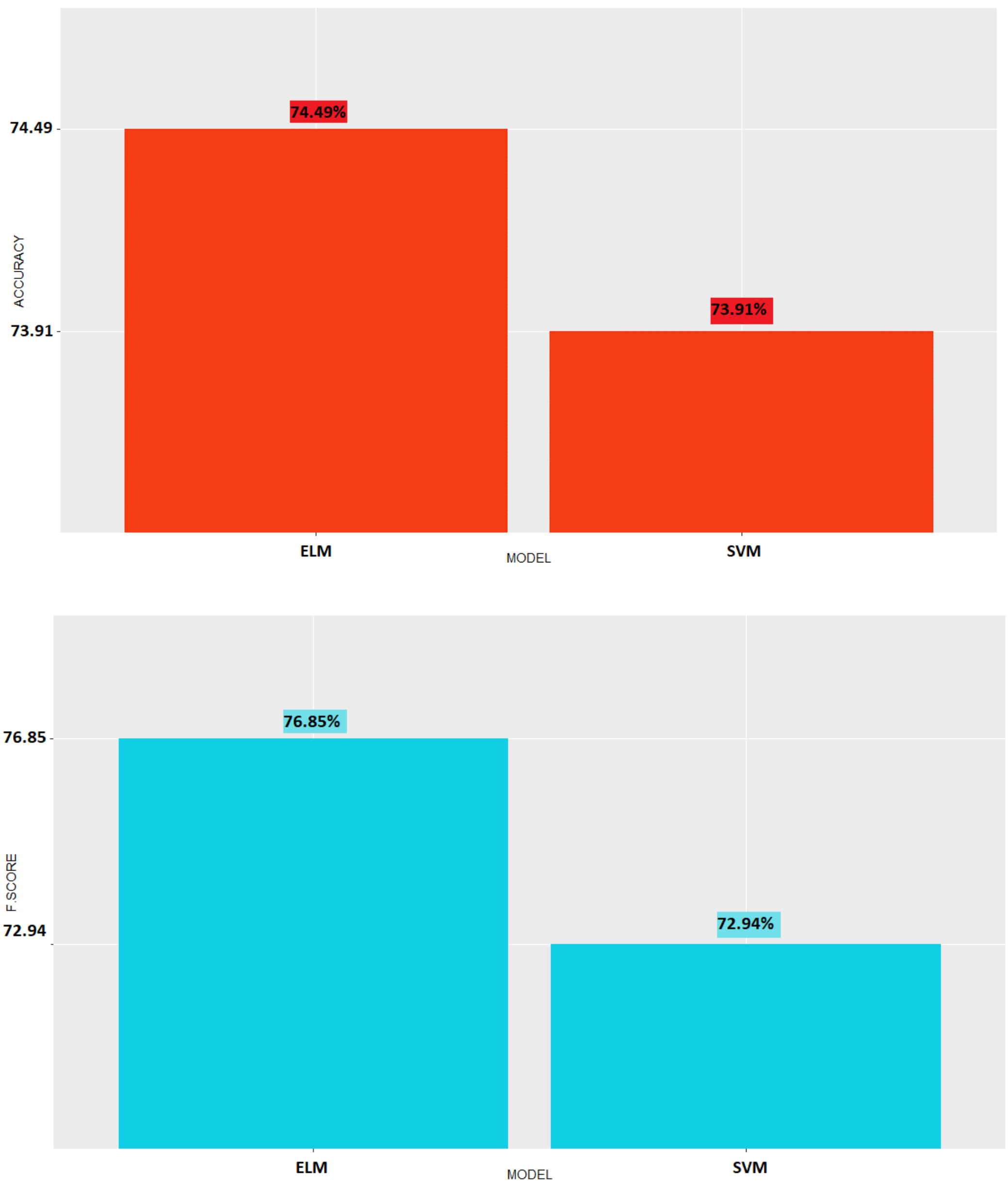

Figure 3. Performance of the models for different measures.

The best results according to Accuracy and F-score obtained for each algorithm separately are given in Table 6. The table separately includes the model performances with the highest F-score and accuracy values among the ELM-based models and the model performances with the highest F-score and accuracy values among the SVM-based models. The Recall value is 85.25% in the model with the highest F-score value when the ELM algorithm is used. In the model with the highest accuracy value when the SVM algorithm is used, the Precision value is 81.21% and the Specificity value is 82.72%. When evaluated in general, it is seen that ELM's performance is better than SVM according to Accuracy and F-score performance indicators.

Table 6. Other performance indicator values of the models.

Classification Algorithm	Accuracy	F-score	Recall	Precision	Specificity
ELM (Best Accuracy Model)	0.7449	0.7528	0.7322	0.7746	0.7593
ELM (Best F-score Model)	0.7275	0.7685	0.8525	0.6996	0.5864
SVM (Best Accuracy Model)	0.7391	0.7289	0.6612	0.8121	0.8272

SVM (Best F-score Model)	0.7333	0.7294	0.6776	0.7898	0.7963

5. Discussion

When the model performances are examined, it is seen that the models obtained have achieved a prediction success of approximately 75%. The value specified is at a level that can be considered successful for a forecasting model. It is important to compare the results of this study with those of other studies using the same data set to better evaluate the success of the prediction models. Oladele et al. [51] obtained the highest accuracy value of 73.07% with the model of the Multi-Layer Perceptron algorithm and the sensitivity value of 75.92% with the model of the J48 algorithm. In these models, future selection has also been conducted. Emon et al. [21] obtained accuracy values between 57% and 75% and F-score values between 35% and 75%. The best results belong to the logistic regression algorithm. Kumar et al. [52] reached the highest accuracy value of 72.29% with the model they developed based on the deep learning architecture. Similar to this study, in the hybrid models developed with Artificial Neural Networks and Particle Swarm Optimization Algorithms, Herliana et al. [53] achieved the highest accuracy with 76.11%, while the precision value was 83.81% and the recall value was 69.22%. Nagi et al. [19] reached an accuracy value of 76.40% with the model obtained using the Two Stage Classifier algorithm. In the same study, the accuracy value obtained using SVM was 60.80%. The comparative results are given in Table 7.

Table 7. Comparative results.

	Accuracy	F-score
ELM-GA model	74.5%	76.9%
[51]	72.0%	-
[21]	75.0%	75%
[52]	72.3%	-
[53]	76.1%	-

When the developed models are compared with the studies performed with the “same data set” in the literature, it has been seen that hyperparameter optimization increases the success of the model according to different performance evaluation criteria. Therefore, this study provided an improvement in the model performance. This shows that the developed models are preferable for DR prediction, especially for a decision support system based on a structural data set, compared to their counterparts in the literature.

6. Conclusions

Within the scope of the study, it is aimed to develop a prediction model using artificial learning algorithms for the detection of diabetes retinopathy and to perform parameter optimization to increase the performance of these models. Genetic Algorithm has been used in parameter optimization, and models have been run with various parameter values for GA, ELM, and SVM. The results obtained were evaluated in terms of accuracy and F-score. According to both performance evaluation measures, the values reached with ELM are higher than those with SVM. Therefore, it can be stated that the models to be developed with ELM are more successful than those with SVM in detecting DR. In addition to the accuracy measure, the F-score has been considered. The reason for choosing the F-score is that this measure is calculated based on two different performance indicators, Recall and Precision.

No reorganization has been made while the dataset was used in the analysis. The purpose of this study is to see what might be the highest prediction success that can be achieved with the basic form of the data. The study is limited in this direction. However, the model can be extended by incorporating various data pre-processing techniques into the hybrid model. However, it should be considered that there are 18 prediction factors (feature space). At this point, even if reducing this number with feature selection increases the success, it may cause the prediction model to be determined by very few features. In addition, when it is considered a real-world problem, it is a separate issue that more factors should be examined together. In addition to feature selection, the performance effect of cross-validation can be examined in the performance validation section. However, due to the presence of parameter optimization, this process will significantly extend the processing time. The results obtained should be evaluated separately in terms of performance increase and negligible processing time.

In the literature studies, it is seen that the tendency toward non-structural, that is, image data sets, has increased. In this direction, it is aimed to measure the success of the developed hybrid models (GA-ELM, GA-SVM) in the processing of image data in future studies. Determining the performance of ensemble learning classifiers, which has gained importance today, is also among the future studies. Developing a decision support system with the best model obtained through additional studies in which different hybrid approaches are discussed is also among the targeted studies.

Funding
For this study, no funding was received.

Conflict of Interest Statement
The author declare that they have no conflicts of interest.

Data Availability Statement
The Diabetic Retinopathy Debrecen Data Set was used to train and test the models, and the data set was obtained from the UCI Machine Learning Repository. Data can be reached from; https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen+Data+Set.

References

1. “World Health Organization,” World Health Organization. Accessed: Nov. 02, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes
2. “IDF Diabetes Atlas,” International Diabetes Federation, 10th Edition, 2021. Accessed: Nov. 08, 2021. [Online]. Available: https://diabetesatlas.org/
3. E. J. Duh, J. K. Sun, and A. W. Stitt, “Diabetic retinopathy: current understanding, mechanisms, and treatment strategies,” *JCI Insight*, vol. 2, no. 14, p. e93751, 2017, doi:10.1172/jci.insight.93751.
4. M. U. Akram, S. Khalid, and S. A. Khan, “Identification and classification of microaneurysms for early detection of diabetic retinopathy,” *Pattern Recognition*, vol. 46, no. 1, pp. 107–116, Jan. 2013, doi: 10.1016/j.patcog.2012.07.002.
5. E. Reichel and D. Salz, “Diabetic retinopathy screening,” in *Managing Diabetic Eye Disease in Clinical Practice*, R. P. Singh, Ed., Cham: Springer International Publishing, 2015, pp. 25–38. doi: 10.1007/978-3-319-08329-2_3.
6. H. Safi, S. Safi, A. Hafezi-Moghadam, and H. Ahmadieh, “Early detection of diabetic retinopathy,” *Survey of Ophthalmology*, vol. 63, no. 5, pp. 601–608, Sep. 2018, doi: 10.1016/j.survophthal.2018.04.003.
7. A. Holzinger, “Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning,” *The IEEE intelligent informatics bulletin*, vol. 15, no. 1, pp. 6–14, 2014.
8. Y. Mintz and R. Brodie, “Introduction to artificial intelligence in medicine,” *Minimally Invasive Therapy & Allied Technologies*, vol. 28, no. 2, pp. 73–81, Mar. 2019, doi: 10.1080/13645706.2019.1575882.
9. T. Vijayan, M. Sangeetha, A. Kumaravel, and B. Karthik, “Gabor filter and machine learning based diabetic retinopathy analysis and detection,” *Microprocessors and Microsystems*, p. 103353, 2020, doi: 10.1016/j.micpro.2020.103353.
10. G. T. Reddy *et al.*, “An Ensemble based Machine Learning model for Diabetic Retinopathy Classification,” in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, Feb. 2020, pp. 1–6. doi: 10.1109/ic-ETITE47903.2020.235.
11. A. Ali *et al.*, “Machine Learning Based Automated Segmentation and Hybrid Feature Analysis for Diabetic Retinopathy Classification Using Fundus Image,” *Entropy*, vol. 22, no. 5, Art. no. 5, May 2020, doi: 10.3390/e22050567.
12. T. R. Gadekallu *et al.*, “Early Detection of Diabetic Retinopathy Using PCA-Firefly Based Deep Learning Model,” *Electronics*, vol. 9, no. 2, Art. no. 2, Feb. 2020, doi: 10.3390/electronics9020274.
13. T. R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. K. R. Maddikunta, and G. Srivastava, “Deep neural networks to predict diabetic retinopathy,” *J Ambient Intell Human Comput*, Apr. 2020, doi: 10.1007/s12652-020-01963-7.
14. D. S. David, “A Novel Specialist System Based on Hybrid Colour and Structure Descriptor and Machine Learning Algorithms for Early Diabetic Retinopathy Diagnosis,” *Artech Journal of Effective Research in Engineering and Technology*, vol. 1, no. 2, pp. 50–56, 2020.
15. G. Alfian *et al.*, “Deep Neural Network for Predicting Diabetic Retinopathy from Risk Factors,” *Mathematics*, vol. 8, no. 9, Art. no. 9, Sep. 2020, doi: 10.3390/math8091620.
16. D. A. Hasan, S. R. M. Zeebaree, M. A. M. Sadeeq, H. M. Shukur, R. R. Zebari, and A. H. Alkhayyat, “Machine Learning-based Diabetic Retinopathy Early Detection and Classification Systems- A Survey,” in *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*, Apr. 2021, pp. 16–21. doi: 10.1109/BICITS51482.2021.9509920.
17. L. Math and R. Fatima, “Adaptive machine learning classification for diabetic retinopathy,” *Multimed Tools Appl*, vol. 80, no. 4, pp. 5173–5186, Feb. 2021, doi: 10.1007/s11042-020-09793-7.
18. M. H. Mahmoud, S. Alamery, H. Fouad, A. Altinawi, and A. E. Youssef, “An automatic detection system of diabetic retinopathy using a hybrid inductive machine learning algorithm,” *Pers Ubiquit Comput*, Jan. 2021, doi: 10.1007/s00779-020-01519-8.
19. A. T. Nagi, M. Javed Awan, R. Javed, and N. Ayesha, “A Comparison of Two-Stage Classifier Algorithm with Ensemble Techniques On Detection of Diabetic Retinopathy,” in *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, Apr. 2021, pp. 212–215. doi: 10.1109/CAIDA51941.2021.9425129.

20. S. Gayathri, V. P. Gopi, and P. Palanisamy, "Diabetic retinopathy classification based on multipath CNN and machine learning classifiers," *Phys Eng Sci Med*, vol. 44, no. 3, pp. 639–653, Sep. 2021, doi: 10.1007/s13246-021-01012-3.
21. M. U. Emon, R. Zannat, T. Khatun, M. Rahman, M. S. Keya, and Ohidujjaman, "Performance Analysis of Diabetic Retinopathy Prediction using Machine Learning Models," in *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, Jan. 2021, pp. 1048–1052. doi: 10.1109/ICICT50816.2021.9358612.
22. Z. Liu, C. Wang, X. Cai, H. Jiang, and J. Wang, "Discrimination of Diabetic Retinopathy From Optical Coherence Tomography Angiography Images Using Machine Learning Methods," *IEEE Access*, vol. 9, pp. 51689–51694, 2021, doi: 10.1109/ACCESS.2021.3056430.
23. K. M. Alabdulwahhab, W. Sami, T. Mehmood, S. A. Meo, T. A. Alasbali, and F. A. Alwadani, "Automated detection of diabetic retinopathy using machine learning classifiers," *European Review for Medical and Pharmacological Sciences*, vol. 25, pp. 583–590, 2021.
24. N. S. Murthy and B. Arunadevi, "An effective technique for diabetic retinopathy using hybrid machine learning technique," *Stat Methods Med Res*, vol. 30, no. 4, pp. 1042–1056, Apr. 2021, doi: 10.1177/0962280220983541.
25. I. Odeh, M. Alkasassbeh, and M. Alauthman, "Diabetic Retinopathy Detection using Ensemble Machine Learning," in *2021 International Conference on Information Technology (ICIT)*, Jul. 2021, pp. 173–178. doi: 10.1109/ICIT52682.2021.9491645.
26. R. M. Devi *et al.*, "Detection of Diabetic Retinopathy using Optimized Back-Propagation Neural Network (Op-BPN) Algorithm," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Apr. 2021, pp. 1695–1699. doi: 10.1109/ICCMC51019.2021.9418433.
27. A. Abreu, D. Ferreira, C. Neto, A. Abelha, and J. Machado, "Diagnosis of Diabetic Retinopathy Using Data Mining Classification Techniques," in *Advances in Digital Science: ICADS 2021*, Springer International Publishing, 2021, pp. 198–209.
28. K. Shankar, Y. Zhang, Y. Liu, L. Wu, and C.-H. Chen, "Hyperparameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification," *IEEE Access*, vol. 8, pp. 118164–118173, 2020, doi: 10.1109/ACCESS.2020.3005152.
29. Y. Modaresnia, F. Abedinzadeh Torghabeh, and S. A. Hosseini, "Enhancing multi-class diabetic retinopathy detection using tuned hyper-parameters and modified deep transfer learning," *Multimed Tools Appl*, vol. 83, no. 34, pp. 81455–81476, Oct. 2024, doi: 10.1007/s11042-024-18506-3.
30. A. Bilal *et al.*, "Improved Support Vector Machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification," *PLOS ONE*, vol. 19, no. 1, p. e0295951, Jan. 2024, doi: 10.1371/journal.pone.0295951.
31. X. Luo *et al.*, "A deep convolutional neural network for diabetic retinopathy detection via mining local and long-range dependence," *CAAI Transactions on Intelligence Technology*, vol. 9, no. 1, pp. 153–166, 2024, doi: 10.1049/cit2.12155.
32. F. M. J. Mehedi Shamrat *et al.*, "An advanced deep neural network for fundus image analysis and enhancing diabetic retinopathy detection," *Healthcare Analytics*, vol. 5, p. 100303, Jun. 2024, doi: 10.1016/j.health.2024.100303.
33. G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, Jul. 2004, pp. 985–990. doi: 10.1109/IJCNN.2004.1380068.
34. L. Qin, Z. Yi, and Y. Zhang, "Enhanced surface roughness discrimination with optimized features from bio-inspired tactile sensor," *Sensors and Actuators A: Physical*, vol. 264, pp. 133–140, Sep. 2017, doi: 10.1016/j.sna.2017.07.054.
35. M. Rafiei, T. Niknam, and M.-H. Khooban, "Probabilistic Forecasting of Hourly Electricity Price by Generalization of ELM for Usage in Improved Wavelet Neural Network," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 71–79, Feb. 2017, doi: 10.1109/TII.2016.2585378.
36. A. Shankar, "Face Detection in images: Neural networks & Support Vector," Indian Institute of Technology, Kanpur, 2002.
37. A. Mammone, M. Turchi, and N. Cristianini, "Support vector machines," *WIREs Computational Statistics*, vol. 1, no. 3, pp. 283–289, 2009, doi: 10.1002/wics.49.
38. C. Cortes and V. Vapnik, "Support-vector networks," *Mach Learn*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.
39. B. Dixon and N. Candade, "Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?," *International Journal of Remote Sensing*, vol. 29, no. 4, pp. 1185–1206, Feb. 2008, doi: 10.1080/01431160701294661.
40. F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., *Automated Machine Learning: Methods, Systems, Challenges*. in The Springer Series on Challenges in Machine Learning. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-05318-5.
41. F. H. F. Leung, H. K. Lam, S. H. Ling, and P. K. S. Tam, "Tuning of the structure and parameters of a neural network using an improved genetic algorithm," *IEEE Transactions on Neural Networks*, vol. 14, no. 1, pp. 79–88, Jan. 2003, doi: 10.1109/TNN.2002.804317.
42. S. Mirjalili, "Genetic Algorithm," in *Evolutionary Algorithms and Neural Networks: Theory and Applications*, S. Mirjalili, Ed., in Studies in Computational Intelligence. , Cham: Springer International Publishing, 2019, pp. 43–55. doi: 10.1007/978-3-319-93025-1_4.

43. R Core Team, *R: A language and environment for statistical computing*. (2018). R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www. R-project.org/
44. RStudio Team, *RStudio: Integrated Development Environment for R*. (2020). RStudio, PBC, Boston, MA. [Online]. Available: http://www.rstudio.com/
45. A. Dragulescu and C. Arendt, *xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files*. (2020). [Online]. Available: https://CRAN.R-project.org/package=xlsx
46. M. Walesiak and A. Dudek, "The Choice of Variable Normalization Method in Cluster Analysis," in *Education Excellence and Innovation Management: A 2025 Vision to Sustain Economic Development During Global Challenges*, K. S. Soliman, Ed., International Business Information Management Association, 2019, pp. 325–340.
47. M. Kuhn, *caret: Classification and Regression Training*. (2020). [Online]. Available: https://CRAN.R-project.org/package=caret
48. D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch, *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly:E1071) R package*. (2015). TU Wien. Accessed: May 27, 2022. [Online]. Available: https://CRAN.R-project.org/package=e1071
49. A. Gosso, *elmNN: Implementation of ELM (Extreme Learning Machine) algorithm for SLFN (Single Hidden Layer Feedforward Neural Networks)*. (2012). Accessed: May 27, 2022. [Online]. Available: https://CRAN.R-project.org/package=elmNN
50. L. Scrucca, *GA: Genetic Algorithms*. (Oct. 15, 2021). Accessed: May 27, 2022. [Online]. Available: https://CRAN.R-project.org/package=GA
51. T. O. Oladele, R. O. Ogundokun, A. A. A. Kayode, and M. O. Adebiyi, "Application of Data Mining Algorithms for Feature Selection and Prediction of Diabetic Retinopathy," in *Computational Science and Its Applications – Lecture Notes in Computer Science*, vol. 11623, Springer, Cham, 2019, pp. 716–730. Accessed: May 26, 2022. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-24308-1_56
52. R. R. Kumar, G. K. Kumar, M. K. Koushik, and M. Chennakesavulu, "A Deep Learning Neural Network for Detecting the Diabetic Retinopathy," *IJITEE*, vol. 8, no. 10, pp. 3522–3525, 2019, doi: 10.35940/ijitee. J9746.0881019.
53. A. Herliana, T. Arifin, S. Susanti, and A. B. Hikmah, "Feature Selection of Diabetic Retinopathy Disease Using Particle Swarm Optimization and Neural Network," in *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, Aug. 2018, pp. 1–4. doi: 10.1109/CITSM.2018.8674295.