SSRG International Journal of Computer Science and Engineering
ISSN: 2348–8387 / https://doi.org/10.14445/23488387/IJCSE-V10I5P105

Volume 10 Issue 5, 33-39, May 2023

Original Article

Investigating Tree-Based Classifiers and Selected Ensemble Learning on Iris Flower Species Classification

Ramoni Tirimisiyu Amosa[1*], Adekiigbe Adebanjo[1], Fabiyi Aderanti Alifat[1], Olorunlomerue Adam Biodun[1], Oni Esther Kemi[1], Adejola Aanu Adeyinka[1], Adigun Olajide Israel[1], Joseph Babatunde Isaac[1]

[1]*Department of Computer Science, School of Applied Sciences, Federal Polytechnic Ede, Osun State, Nigeria.*

Received: 05 April 2022 Revised: 09 May 2022 Accepted: 19 May 2023 Published: 31 May 2023

Abstract *- Eloquence, hope, knowledge, the ability to communicate effectively, and faith are some of the meanings associated with the iris flower in the language of flowers. Iris has different species types, and each type has its own medicinal purpose. Classifying the flower has become a serious task for researchers due to the high volume of datasets (big data), hence the introduction of machine learning algorithms for accurate and reliable classification. This paper focuses on the classification of the Iris flower using five tree-based algorithms; Best First Tree (BFTree), Least Absolute deviation Tree (LADTree), Cost-Sensitive Decision Forest (CSForest), Functional Tree (FT) and Random Tree (RT). Three selected ensemble learning (Bagging, Dagging and cascade generalisation) were equally implemented in the algorithm. The dataset that was utilised in this investigation is open source and may be downloaded without cost from a public repository (kaggle.com). The result of the classification showed that the FT classifiers outperform other tree-based classifiers with an accuracy of 96.67% and an AUC of 1.00. The ensemble algorithm has a significant impact on the performance of single classifiers (tree-based). Outperform tree based. AUC/ROC (Area Under Curve/Receiver Operating Characteristics) was used to evaluate the algorithm's performance. Bagging ensemble outperforms other ensembles (Dagging and Cascade) with an accuracy of 96.00% and AUC of 1.00.*

Keywords *- Algorithm, Classification, Dataset, Ensemble, Performance.*

1. Introduction

A flower is a component of a plant that typically has vivid colors, develops at the end of a stalk, and only lasts a brief period. Unlike trees, bushes, and vegetables, flowers are small plants that are grown only for their blossoms [1, 7]. In terms of ecosystems, flowers are the most significant producers on earth and may flourish in a wide range of climatic conditions. They continue to play a significant role in the food chain by providing food for practically all insect species worldwide, and numerous medications can be made using their medicinal characteristics. For these reasons, understanding flowers and their species is crucial for identifying new or endangered plant species. If not, many plants might suffer damage or be sold for incredibly low rates because they are thought to be dangerous to one's agriculture.

Moreover, all of this happens because the plant species are not properly recognised. However, it is a proven fact that many of the plants seen in nature can be cultivated. [6, 13, 14, 17] classified Iris flowers into different types, which include; Bulb Iris, Dutch Iris, Reticulata Iris, Rhizome Iris, Bearded Iris, and Beardless Iris.

The iris flower is seen as a symbol of discernment, grace, faith, and regal events in life. The flower is renowned for its immaculate beauty, lovely hues, and various variants. Irises are also associated with beauty, innocence, ecstasy, and compassion. Due to its prevalence in a range of temperate regions throughout the world, its meanings have been modified to suit different cultural contexts. IRIS flower is one of the biological plants that may be used for a variety of things, including medical uses as a skin salve, treating kidney and bladder issues, reducing nausea, and using it as perfume. Numerous sorts of research have been done for pattern identification of IRIS flowers. When bioinformatics analysts used traditional methods to analyse IRIS flowers, it could take days or weeks to find significant information, there is often some information hidden in the data, and much of the data is never analysed. Machine learning techniques are used to solve these kinds of problems [9]. It is biology data containing about 300 species distributed in temperate regions across the Northern Hemisphere, occurring mostly in Eurasia and North America. Although certain iris species can be found in wetland habitats, the majority of iris species live in dry, stony, or semi-arid conditions. The distinctive characteristics of iris species include a basal fan of uni-facial leaves, a colorful perianth with three horizontal sepals and three upright petals that are basally joined into tube-shaped branches, and three stamens that are arranged in opposition to the sepals and petals use for medicinal purposes. [2, 5, 8]

The iris flower is a well-liked presenting flower prized for its distinctive structure and gorgeous hues. The iris plant is highly adaptable and ideal for bringing brightness to both indoor and outdoor spaces. Iris flowers also have an intriguing backstory to share, thanks to the special symbolism and meanings associated with these prized blooms. Irises make wonderful gifts for all different types of individuals in your life because of their wide range of colors and symbolic meanings. You can offer a parent or mentor purple irises, a significant other yellow irises, and someone who needs a little more motivation blue iris. [3, 11, 21] The color of the flower affects a particular iris' symbolism: Irises are said to represent both love and monarchy in purple and wisdom in yellow. Irises that are blue stand for optimism and faith. White iris represents innocence. Data mining is an important area for computer scientists. Data mining techniques help in gaining knowledge so that the data may be sorted, organised, and used to make future predictions. The key challenge nowadays is how to extract useful information from the enormous number of data [4].

Classification is one of the data mining techniques; it supervises learning, in which different data map to the predetermine classes or groups, an algorithm that implements classification known classifier. The goal of classification, a supervised machine learning technique, is to predict the appropriate label for some input data. In classification, the model is thoroughly trained using the training data before being evaluated using the test data and then used to make predictions on fresh, untainted data. The Classification algorithm is a Supervised Learning method used to categorise fresh observations based on training data. In classification, a program makes use of the dataset or provided observations to learn how to classify fresh observations into different categories or groups. For example, yes or no, 0 or 1, red or blue, spam or not spam, etc. Classes can be denoted by targets, labels, or categories. Because the Classification method is an unsupervised learning technique that includes input and output data, it makes use of labeled input data. Machine Learning is a program that learns from past data sets to perform better with experience. It is a tool and technology we can utilise to answer questions with our data. Machine Learning works on two values these are discrete and continuous. The goal of machine learning is to create computer programs that can learn to improve and adapt as they encounter new information. Weather forecasts, spam detection, biometric attendance, computer vision, pattern recognition, sentiment analysis, identification of diseases in the human body, and many more are just a few examples of the various uses and applications of machine learning. Machine learning techniques fall into three categories: supervised, unsupervised, and reinforcement learning. Examples of a training data set with various input properties and anticipated outputs can be found in supervised learning. The classification process is a subset of supervised learning in which a computer program picks up knowledge from the data it receives and applies it to categorise fresh observations. There are several classification techniques: Decision Trees, Bayes Classifiers, Nearest Neighbor, Support Vector Machines, Neural Networks, and many more.

2. Literature Review

[6, 23] applied Artificial Neural Network (ANN) and Support Vector Machine (SVM) for iris flower species classification. The classification was obtained by first searching for patterns in the sizes of the iris flower's sepals and petals and then determining how the pattern was predicted to form the class of the iris flower. This process was repeated until the classification was complete. The results of the experiments showed that ANN and SVM are both capable of accurately classifying iris flowers, with ANN achieving 98.66% accuracy and SVM achieving 97.79% accuracy, respectively. SVM and Genetic algorithm (GA) is also better classifiers when it comes to iris classification. According to [16, 20], SVM produced an accuracy of 98.70%, while GA accuracy was 97.78%.

The execution of the identification of the Iris flower is carried out using the Naive Bayes classifier (NBC) and K closest neighbour (KNN) in [26] research work. The NBC operates on the basis of two fundamental tenets: first, it determines the posterior probability for a given likelihood function, and second, it uses the prior probabilities already known for each class. The model's performance is tested by dividing the dataset into two distinct sets, one of which is used for training and the other is used for testing, which results in an average testing accuracy of 100%. Second, the K-nearest neighbour strategy is just a majority vote amongst the K most comparable occurrences to a given "unseen" observation, and it produced a model accuracy of 98.57%. This method was used to get this result. [16] in his research work titled "A study of pattern recognition of Iris flower based on Machine Learning", in this research K-means algorithm was used to achieve clustering and species classification, and the evaluation of the result achieved an accuracy of 94.7%.

Ahmadi et al. [27], in their research titled "An intelligent method for iris recognition using supervised machine learning techniques", adopted the Two-dimensional Gabor kernel (2-DGK), step filtering (SF), and polynomial filtering (PF) approach and accuracy of 99.99%. This accuracy is considered to be better when compared with [15] with an accuracy of 86.00%; although the algorithm implemented in this case are Random Forest and Logistic Regression (LR), both cases employed the feature extraction techniques of principal component analysis (PCA). [4] developed Flower Classification using Supervised Learning using random forest and LR and as well achieved an accuracy of 96.67% which is an improvement over [10]. Sepal and petal sizes of the flower are popular features in classifying Iris flower, and accuracy can be better improved with algorithms such as SVM and KNN [12, 18, 25]

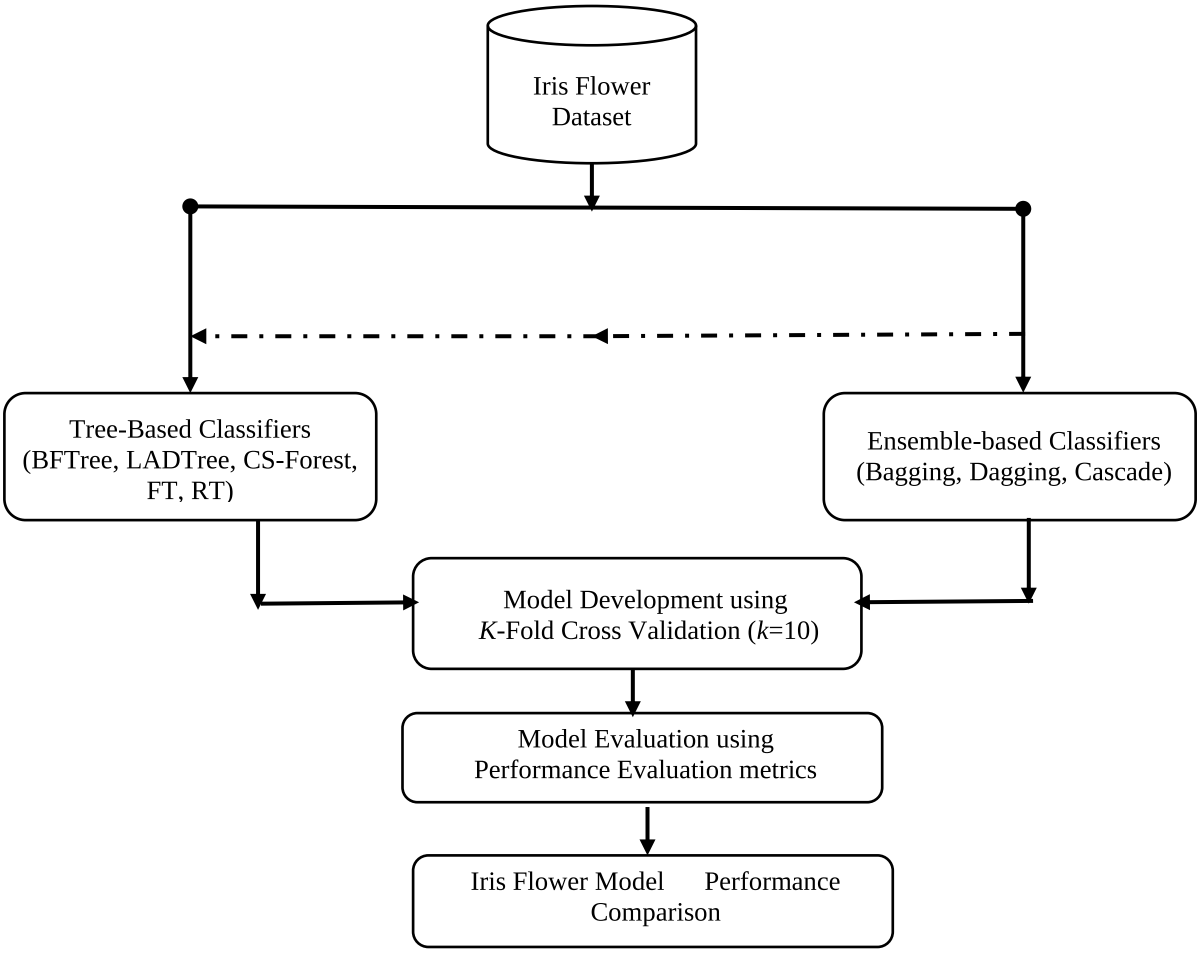

Fig. 1 Research framework

3. Methodology

The datasets used in this research were acquired from a public data repository (Kaggle.com)—the tree-based classifiers such as BFTree, LADTree, CSForest, FT and RT. Three selected ensemble learning used are Bagging, Dagging and cascade generalisation. The combined models increase the accuracy of the results significantly [19, 27]. Weka (version 3.8.5) Machine learning tool was used to analyse and classify the dataset and formulate the model. Finally, the evaluation of the models were implemented using the Accuracy and ROC/AUC (Receiver Operating Characteristics/Area Under Curve) curve.

3.1. Best First Tree (BFTree)

In the field of machine learning, classification and regression work can be accomplished with the help of a decision tree technique known as the best-first tree. A greedy algorithm is one that chooses the optimal attribute to split the data at each node based on a quality metric, such as information gain or gain ratio. This algorithm is an example of a greedy algorithm. The Best-first tree algorithm works by constructing a tree from the root node downward, iteratively selecting the attribute that maximises the quality measure, and establishing a new branch for each conceivable value of the attribute. This process is repeated until the tree has been completed. The algorithm will keep running until one of its stopping criteria is satisfied. These criteria could include achieving the maximum depth, the minimum number of instances per leaf, or the minimal reduction in impurity (Rana et al., 2020)

3.2. Least Absolute Deviation (LADTree)

For applications involving regression, the machine learning method known as LADTree is typically utilised. Least Absolute Deviation is a loss function that is utilised by LADTree in order to estimate the model parameters. The acronym "LAD" refers to this loss function. The most typical loss function in conventional regression is called the Mean Squared Error, or MSE for short. This function penalises large errors more severely than it makes minor errors.

On the other hand, LADTree makes use of the Least Absolute Deviation loss function, which is more resistant to outliers and does not place a heavy emphasis on substantially penalising huge errors. In order to make an accurate prediction of the target variable, the LADTree method constructs a decision tree by recursively partitioning the data based on the input attributes. At each split, the algorithm chooses the feature that uses the LAD loss function to achieve the greatest possible separation of the data. The tree structure that was produced as a result makes it possible to automatically identify significant input features and the interactions that exist between them.

3.3. Cost-Sensitive Decision Forest (CS-Forest)

The method known as the decision tree is useful for producing accurate classifications. The tree branches are divided at specific points so that there is as little overlap between the two sets of examples as possible. When both groups are dominated by instances from one class, the criterion used to identify a split point will see good separation when, in reality, the examples from the minority class are being disregarded. As an alternative to the pruning method utilised by C4.5, CS-Forest employs a cost-sensitive method of tree cutting [24, 26]. C4.5 will prune a tree provided that the likelihood of future records being incorrectly classified does not rise in an expressive manner as a direct result of the pruning. Nevertheless, CS-Forest will prune a tree only if it determines that the likely classification cost for upcoming records will not grow expressively as a result of the pruning. In addition, in contrast to the Cost-Sensitive Decision Tree (CS-Tree), the Cost-Sensitive Forest (CS-Forest) allows a tree to reach its full potential before it is cut initially.

3.4. Functional Tree (FT)

The Functional Tree (FT) algorithm is a machine learning type used to solve regression and classification problems. Jerome H. Friedman presented it for the first time in his paper titled "Multivariate Adaptive Regression Splines", published in 1991. In order to make an accurate prediction of the target variable, the FT algorithm is a tree-based method that constructs a piecewise linear function depending on the features fed into it. It does so by constructing a binary tree, each node of which comprises a test on one of the input features. The data is split up into two subsets by the test, and the tree then develops a function in a recursive fashion for each of those subgroups [15].

3.5. Random Tree (RT)

RT is a method of supervised and collective learning that results in a large number of individuals learning on their own. It makes use of a flimsy concept to assemble a collection of random data in order to form a tree. Within the standard tree, each node is partitioned utilising the split that yields the greatest results considering all criteria. In the RF, each node is partitioned using a predictor that performs the best among a selection of predictors that was randomly chosen at that node [7]

3.6. Bagging

An early ensemble approach, known as bootstrap aggregating or bagging classifier for short, bootstrap aggregating's primary purpose is to reduce the variance (overfitting) across a training set. As a type of bagging classifier, the random forest model is one of the variants that is used the most commonly. Intuitively, the bagging model selects the class for a classification problem based on the major votes estimated by the number of trees to lower the overall variance, while the data for each tree is selected by random sampling with replacement from the overall dataset. This is done to ensure that the overall dataset is as representative as possible. The bagging model, on the other hand, takes an average of several estimates when applied to regression situations.

3.7. Dagging

Dagging, which is an abbreviation that stands for "Decorrelated Aggregating," is a form of ensemble learning approach. This method entails training many models on distinct subsets of the training data, with the training set for each model being randomly picked from the original training data with replacement. Dagging is a type of ensemble learning. The fact that the models are trained to be as independent from one another as is practically possible while still being aggregated in an ensemble is where the term "decorrelated aggregating" originates from. Dagging is an ensemble method that focuses on constructing varied models with as little connection between them as possible. This is in contrast to other ensemble methods, such as bagging and boosting, which merely sample subsets of the available data. To achieve this goal, each model is trained using a unique subset of the training data, and the sequence in which the features are utilised during training is randomised.

3.8. Cascade Generalization (CG)

CG is a form of ensemble learning approach in which many models are trained in a succession of stages, with each step using the output of the previous stage as an extra input. Each stage uses the output of the stage before it as additional input. This method is sometimes referred to as cascade generalisation and stacked generalisation. The first step in CG is training a series of base models using the first training data. This is done using the original data. After that, the results of these models are put to use as additional input features for the training of a second set of models, the results of which are then put to use as input features for a third set of models, and so on.

3.9. Model Evaluation Metrics

3.9.1. Accuracy

When analysing classification models, accuracy is one parameter that can be used. In a more colloquial sense, accuracy refers to the proportion of correct predictions made by our model. According to the accepted definition, accuracy is defined as follows:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Calculating accuracy in terms of positives and negatives is another method that can be utilised when dealing with binary classification.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Where True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) are denoted, respectively.

3.9.2. AUC

Area Under Curve (also known as AUC) is one of the measures utilised most frequently in evaluation. It is applied to problems involving binary classification. The area under the curve (AUC) of a classifier is equal to the likelihood that the classifier will rank an example that was picked at random as positive higher than an example that was chosen at random that was considered negative. Suppose the area under the curve (AUC) is less than one.

In that case, it indicates that the model is predicting more classes accurately and is moving in the right direction i. e The closer the AUC to 1, the better the model and if AUC < = 0.5, tis means that the classifier cannot predict the larger percentage of data correctly.

4. Result and Discussion

This section discusses the various results of the experiment.

Table 1. Tree-based classification result

Algorithm/ Classifier	Accuracy (%)	AUC	F-Measure	MCC
BFTree	94.67	0.97	0.95	0.92
LADTree	94.00	0.96	0.94	0.91
CSForest	92.67	0.99	0.93	0.89
FT	96.67	1.00	0.97	0.95
RT	92.00	0.94	0.92	0.88

The five tree-based classifiers were applied for the classification, as shown in Table 1. The classification result shows that FT is the best classifier here, having produced an accuracy of 96.67%, 1.00 AUC and MCC of 0.95. CSForest also performs better with an AUC of 0.99. However, BFTree has better accuracy when compared with that of CSFores, but the accuracy of the classifier cannot be relied on as a metric for evaluating the algorithm.

Table 2. Bagging ensemble classification result

Bagging+ Classifier	Accuracy (%)	AUC	F-Measure	MCC
Bagging+ BFTree	94.67	0.99	1.00	1.00
Bagging+ LADTree	96.00	0.99	0.96	0.94
Bagging+ CSForest	94.00	0.99	0.94	0.91
Bagging+FT	96.00	1.00	0.96	0.94
Bagging+RT	94.67	0.99	0.95	0.92

Bagging is applied to improve the performance of the tree-based classifiers. The result shows that the ensemble has a significant impact on the performance of the classifiers. It improves the accuracy of the classifiers, and also the AUC of all the classifiers improve greatly. However, bagging combined with FT has the best AUC of 1.00, while other classifiers also performed excellently with the same AUC of 0.99. It is also observed that the FT as a single classifier has the best AUC (1.00), which is the same result when combined with the Bagging ensemble except for the difference in MCC. This shows that the bagging ensemble has no significant effect on FT, while it has an impact on other classifiers.

Table 3. Dagging ensemble classification result

Dagging+ Classifier	Accuracy (%)	AUC	F-Measure	MCC
Dagging +BFTree	92.00	0.98	0.92	0.88
Dagging +LADTree	94.00	0.99	0.94	0.91
Dagging +CSForest	33.33	0.50	0.32	0.00
Dagging +FT	33.33	0.50	?	?
Dagging +RT	92.67	0.99	0.93	0.89

Table 3 shows the result of Dagging ensemble combined with the five tree-based classifiers. Dagging combined with LADTree produces the best result with an AUC of 0.99 and accuracy of 94.00%, followed by dagging+RT and Dagging+BFTree. Dagging+CSForest and Dagging+FT were observed to perform poorly, with an AUC of 0.50. This shows that the ensemble with the two classifiers has average performance, which means the possibility of correct classification is 50%. This result also shows that CSForest and FT would perform better when applied as single classifiers than combined with an ensemble-like dagging.

Table 4. Cascade ensemble classification result

Cascade+ Classifier	Accuracy (%)	AUC	F-Measure	MCC
Cascade +BFTree	94.67	0.92	0.95	0.92
Cascade +LADTree	94.00	0.96	0.94	0.91
Cascade +CSForest	92.67	0.99	0.93	0.89
Cascade +FT	96.67	1.00	0.97	0.95
Cascade +RT	94.00	0.96	0.94	0.91

The Cascade ensemble is also one of the popular ensembles used for classification. Here, the cascade is also combined with all the tree-based classifiers. Cascade+FT produced the best classification result with an accuracy of 96.67% and AUC of 1.00, followed by cascade+CSForest. As shown in Table 2, the result of the bagging ensemble outperforms the two other ensembles (Dagging and Cascade). Bagging performs excellently well with all the tree-based classifiers, producing the highest accuracy and AUC, followed by the performance of the Cascade algorithm. Dagging seems not to be a good ensemble for Iris flower classification because it performed averagely and is difficult to predict accurately in some instances, as shown in Table 3.

5. Conclusion

The outcome of the research shows that the performance of the ensemble learning algorithm outperforms the single classifier. Although this may not be applicable in all circumstances, as shown in Table 5.

Table 5. Summary of the classification result

Algortihm/Classfifier	Accuracy	AUC	MCC
FT	96.67	1	0.95
Bagging+FT	96	1	0.94
Dagging +LADTree	94	0.99	0.91
Cascade +FT	96.67	1	0.95

Table 5 shows the summary of the classification; the best classifier in each of Table 1, table 2, table 3 and Table 4 were selected and combined to make Table 4. The output of the table shows that FT and Cascade + FT produces the same result. However, the ensemble produced is better, more accurate and more reliable than the result of some of the previous researchers because some of these research-based the algorithm's performance on accuracy only. However, using accuracy only to measure the model's performance is unreliable, as stated earlier, because most classifiers are always biased towards the majority class. Hence, AUC/ROC and MCC were also used to compare the performance of the classification.

References

[1] Neda Ahmadi et al., “An Intelligent Method for Iris Recognition using Supervised Machine Learning Techniques,” *Optics & Laser Technology*, vol. 120, p. 105701, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[2] B. Angkita, and V. Deval, “Recognition of Iris Flower Species using Spatial Features,” *NeuroQuantology*, vol. 20, no. 14, pp. 880-886, 2022.

[3] Anju Gulia, Rajan Vohra, and Praveen Rani, “Liver Patient Classification using Intelligent Techniques,” *International Journal of Computer Science and Information Technologies*, vol. 5, no. 4, pp. 5110-5115, 2014. [Google Scholar] [Publisher Link]

[4] T. Shashidhar Halakatti, and Shambulinga T. Halakatti, “Identification of Iris Flower Species using Machine Learning,” *IPASJ International Journal of Computer Science (IIJCS)*, vol. 5, no. 8, pp. 59-69, 2017. [Google Scholar]

[5] Zahra Faiz Hussain et al., “A New Model for Iris Data Set Classification Based on Linear Support Vector Machine Parameter's Optimization,” *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, pp. 1079-1084, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Amar Yehya Hussien, “Flower Species Recognition Using Machine Learning Classifiers,” *Academic Journal of Nawroz University*, vol. 11, no. 4, pp. 469-475, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] G.A. Vijayalakshmi Pai, and S. Rajasekaran, *Neural Networks, Fuzzy Logic and Genetic Algorithms Synthesis and Applications*, 6th Edition, Prentice Hall of India Pvt. Ltd., 2006.

[8] Swati Rath, and Santanu Rath, “An Evolutionary Approach for Protein Classification Using Feature Extraction by Artificial Neural Network,” *International Conference on Computer and Communication Technology (ICCCT)*, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[9] Asmita Shukla et al., “Flower Classification Using Supervised Learning,” *International Journal of Engineering Research and Technology*, vol. 9, no. 5, pp. 757-762, 2020. [Google Scholar] [Publisher Link]

[10] Ahmed Hussein Ali, “A Survey on Vertical and Horizontal Scaling Platforms for Big Data Analytics,” *International Journal of Integrated Engineering*, vol. 11, no. 6, pp. 138-150, 2019. [Google Scholar] [Publisher Link]

[11] Shahreen Kasim et al., “A Comparative Study of Different Template Matching Techniques for Twin Iris Recognition,” *International Journal on Advanced Science, Engineering and Information Technology*, vol. 7, no. 4-2, pp. 1581-1588, 2017. [Google Scholar] [Publisher Link]

[12] Shahreen Kasim et al., “A Study on Facial Expression Recognition Using Local Binary Pattern,” *International Journal on Advanced Science, Engineering and Information Technology*, vol. 7, no. 5, pp. 1621-1626, 2017. [Google Scholar] [Publisher Link]

[13] Nur Ariffin Mohd Zin et al., “Contact Lens Classification by Using Segmented Lens Boundary Features,” *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 11, no. 3, pp. 1129-1135, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[14] Debaraj Rana, Swarna Prabha Jena, and Subrat Kumar Pradhan, “Performance Comparison of PCA and LDA with Linear Regression and Random Forest for Iris Flower Classification,” *PalArch's Journal of Archaeology of Egypt/Egyptology*, vol. 17, no. 9, pp. 2353-2360, 2020. [Google Scholar] [Publisher Link]

[15] Madhusmita Swain et al., “An Approach for Iris Plant Classification using Neural Network,” *International Journal on Soft Computing*, vol. 3, no. 1, pp. 79-89, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[16] Lulin Ma et al., “Transcriptome Analysis of Key Genes Involved in Color Variation between Blue and White Flowers of Iris Bulleyana,” *BioMed Research International*, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Shivam Vatshayan, “Performance Evaluation of Supervised Learning for Iris Flower Species,” *International Journal of Innovative Research in Technology*, vol. 5, no. 11, 2019. [Google Scholar] [Publisher Link]

[18] Joylin Priya Pinto, Soumya Kelur, and Jyothi Shetty, “Iris Flower Species Identification using Machine Learning Approach,” *2018 4th International Conference for Convergence in Technology (I2CT)*, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] Sajida Perveen et al., “A Systematic Machine Learning based Approach for the Diagnosis of Non-alcoholic Fatty Liver Disease Risk and Progression,” *Scientific Reports*, vol. 8, pp. 1-12, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[20] Yu Yang, “*A Study of Pattern Recognition of Iris Flower based on Machine Learning*,” Bachelor’s Thesis, 2013. [Google Scholar] [Publisher Link]

[21] Anshuman Singh, Rohan Akash, and V. Gokul Rajan, "Flower Classifier Web App Using Ml & Flask Web Framework," *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] N. Pavitha, and Shounak Sugave, "A Novel Multi-Stage Stacked Ensemble Classifier using Heterogeneous Base Learners," *International Journal of Engineering Trends and Technology*, vol. 71, no. 4, pp. 65-71, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Pranshul Agarwal et al., "Optimised Ensemble Machine Learning Model for IRIS Classification," *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Adnan Mohsin Abdulazeez et al., "Leaf Identification Based on Shape, Color, Texture and Vines using probabilistic Neural Network," *Computacion y Sistemas*, vol. 25, no. 3, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Mohd Arfian Ismail et al., "Newton Competitive Genetic Algorithm Method for Optimization the Production of Biochemical Systems," *Advanced Science Letters*, vol. 24, no. 10, pp. 7481-7485, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[26] Mohd Arfian Ismail et al., "Multi-objective Optimization of Biochemical System Production Using an Improve Newton Competitive Differential Evolution Method," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 7, no. 4-2, pp. 1535-1542, 2017. [Google Scholar] [Publisher Link]

[27] Mohd Arfian Ismail et al., "Optimisation of Biochemical Systems Production using Hybrid of Newton Method, Differential Evolution Algorithm and Cooperative Coevolution Algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 8, no. 1, pp. 27-35, 2017. [CrossRef] [Google Scholar] [Publisher Link]