IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
onal Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

IRIS Flower Classification Using Machine Learning

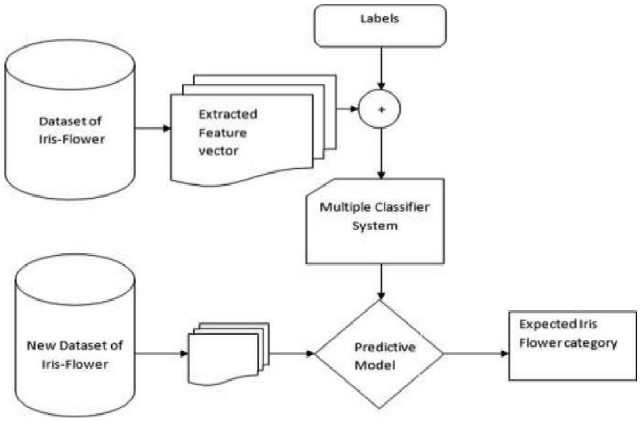
Arjun Patil[1] and Suyog Thakur[2]
Assistant Professor and Head Department of IT[1]
Student, P.G. Department of IT[2]
Veer Wajekar ASC College, Phunde, Uran

Abstract: *This study employs machine learning techniques, specifically the Random Forest classifier, to accurately identify different species of Iris flowers based on key morphological attributes. By training the model on measurements of petal and sepal lengths and widths, the research aims to uncover patterns that enable the classification of unseen Iris samples into their respective species. The primary objective is to develop a reliable prediction model that simplifies and automates the process of species identification. This work highlights the effectiveness of data-driven approaches in botanical classification and showcases the practical application of machine learning in real-world categorization tasks. The model not only enhances the accuracy and efficiency of plant species recognition but also serves as a valuable example of machine learning's potential in the field of biological sciences.*

Keywords: Machine Learning, Web Application, Random Forest Classifier, Real-Time Prediction

I. INTRODUCTION

AI, or machine learning, allows computers to learn without explicit programming, is a crucial area in computer science. This idea was initially proposed by Arthur Samuel in 1959, and since then, it has developed into a significant field of artificial intelligence where The main goal is to develop algorithms that are capable of learning from data. and make predictions. In contrast to conventional programming paradigms, Algorithms for machine learning construct models from sample inputs by using patterns found in pattern recognition and computer learning theory as inspiration. Because these models base their decisions or predictions on data inputs rather than strict, static instructions, they are extremely flexible and efficient for a variety of complex and variable computer tasks. The use of machine learning many uses, which highlights its revolutionary potential. It is essential to several domains, including network intrusion detection, email filtering, computer vision, and learning to rank. Due to its adaptability and capacity to manage intricate data-driven tasks, machine learning is an essential tool in modern technology.



This project illustrates a typical machine learning workflow sing the Iris dataset. A classification model is constructed using measurements of an iris blossom attributes and their species from the Iris dataset. In orderto achieve of

ISSN 2581-9429 IJARSCT

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
onal Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

facilitating model evaluation, the dataset is split into sets for testing and training. Next, using the training data, the Random Forest Classifier, an ensemble learning algorithm, is started and trained.During the

In this study, a machine learning model is trained to classify *Iris* flower species based on petal and sepal dimensions. The process includes data preparation, model training, prediction generation, and accuracy evaluation. Once trained, the model is saved for future use and is capable of predicting the species of *Iris* flowers using new input measurements.

To enhance usability, a web application is developed featuring two distinct user roles: Administrators and Users. Administrators can view user-generated predictions, manage frequently asked questions (FAQs), and edit the inventory of registered users. Users, on the other hand, can register, log in, access the dataset, ask or view FAQs, and use the model to predict the species of *Iris* flowers based on their measurements.

This application ensures efficient operation and seamless interaction by assigning specific functionalities to each user role. The project showcases the practical application of machine learning in building user-friendly, role-based systems for real-time biological classification tasks.

Technologies Used

Python

Python is a popular high-level, interpreted programming language known for its simplicity and readability, making it easy to learn and understand. Unlike languages such as C++ or Java, Python was developed by Guido van Rossum with an emphasis on code clarity and ease of use. also initially released in 1991, places an a focus on readable code and syntax, enabling programmers to express concepts in fewer lines of code. Python is a adaptable vocabulary that can be applied to a variety of tasks, such as artificial intelligence, data analysis, and web development, machine learning, automation, and scientific computing. Its vast standard library and sizable ecosystem of third-party packages further contribute to its versatility. The language's design philosophy, which is reinforced by features like automated memory management, dynamic typing, and a large collection of pre-built data structures, encourages producing clear and maintainable code. Another factor supporting Python's popularity is its vibrant community, which consistently advances the language.

HTML

HyperText Markup Language, or HTML, is the popular markup language used to produce and organize content on the web. It acts as the cornerstone of all webpages, defining their layout and style with the use of different elements and tags. A series of nested tags divides an HTML document's content parts, which include headings, paragraphs, links, images, and other multimedia components. Programmers can design complex web pageswith interactive forms, text, and graphics with ease thanks to HTML's flexibility and simplicity. HTML has evolved over time, and HTML5 is the greatest recent iteration that includes new features and components for handling multimedia, graphics, and more robust document architecture.

Libraries and Frameworks:

Scikit-Learn

Python's Scikit-learn library is a robust and intuitive machine learning tool. It makes a range of machine learning methods, like the Random Forest Classifier, easier to deploy by offering straightforward and efficient data mining instruments analysis. Due to the fact that it is founded on NumPy, SciPy, and Matplotlib, it has excellent speed and is easy to integrate with other libraries for scientific computing.

Pandas

The popular Python toolbox Pandas is used for data analysis and manipulation. It offers data structures that facilitate the handling and manipulation of structured data, such as Data Frames. Pandas has a large number of functions to collaborate with many data kinds and formats, making it a popular instrument for data translation, cleansing, and analysis.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
onal Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

NumePy

A core Python numerical operations library is called NumPy. Large multi-dimensional arrays and matrices are supported, and several mathematical operations on these arrays can be performed with its help. Scientific computing requires NumPy as it forms the foundation of numerous data analysis and machine learning tools, such as Scikit-learn and Pandas.

Flask

Python has a lightweight and adaptable web framework called Flask. Its purpose is to facilitate the rapid and minimal setup creation of web applications. Flask is a popular choice for constructing small to medium-sized web apps and APIs since it manages backend processes and allows extensions to add functionality as needed.

II. LITERATURE SURVEY

H. D., Y. T. Reddy, and L. K. (2022) conducted a study aimed at classifying different species of *Iris* flowers using various machine learning models, including K-Nearest Neighbors (KNN), Decision Trees, Gaussian Naive Bayes, and Support Vector Machines (SVM). The primary objective was to evaluate the performance of these models based on key metrics such as accuracy, precision, recall, and F1 score. The study sought to identify the most effective model for predicting *Iris* species based on physical features like petal length, sepal length, and sepal width. Additionally, the research demonstrated how the selected model could be applied to new data to generate accurate predictions.

J. P. Pinto, S. Kelur 2018,This research uses the Iris dataset to investigate and contrast various classification techniques for iris flower species prediction. We specifically concentrate on Assistance Vector Systems(SVM), Naive Bayes, and Decision Trees. The objective is to apply these techniques with the Scikit- learn tool, evaluate each one's accuracy-based performance, and identify the optimal approach. To determine the best classification strategy for this assignment, we test the algorithms using the Iris dataset, which contains measurements of sepal and petal diameters.

3. Y. Pachipala and H. C. Maddipati (2022) explored the application of artificial intelligence techniques, particularly neural networks and classification algorithms, for identifying different species of Iris flowers using the well-known Iris dataset. The study focused on achieving accurate classification based on sepal and petal measurements, employing the Random Forest algorithm as the primary model. The objective was to deepen the understanding of how machine learning, especially neural networks, can be leveraged to precisely identify and categorize Iris species. By using a pre-processed dataset and evaluating key performance metrics such as error rate and training time, the authors aimed to develop a robust flower classification approach. This work contributes to the field of data mining and supports biologists by improving prediction accuracy in plant classification tasks.

A.ELDEM,H.ELDEM2018, ThisworkaimstouseDeepNeuralNetworks(DNN)toclassifyirisblossomsinto three classes. The measurements of the petals and sepals are employed as input attributes. The optimal configuration is determined by experimenting with various epoch counts and activation functions. High classification accuracy is the desired outcome as well as on the Iris dataset, the top-performing model achieved a 96% success rate.

T. Gupta, P. Arora 2022, This research compares and assesses three classification models for iris flower species prediction: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Logistic Regression. The Iris dataset is preprocessed using Exploratory Data Analysis (EDA) in this research, and then each classification model is applied to provide predictions. Finding the model with the best accuracy for iris flower classification based on characteristics such petal width, petal length, and sepal width is the aim. The study provides a thorough analysis of the models' performancesandreportsthemodels'accuracyas96.43% for Logistic Regression, 98.21% for SVM, and 94.64% for KNN.

6. V. Arya and R. K. Rathy (2014) aimed to enhance the classification of the Iris dataset using a neuro-fuzzy approach. The dataset categorizes Iris flowers into three species—Setosa, Virginica, and Versicolor—based on floral measurements. However, it often presents challenges due to non-linear classification boundaries. To achieve more accurate results, this study explores the application of neuro-fuzzy classification algorithms, which integrate fuzzy logic with neural networks, considering four key parameters. The primary objective is to improve overall classification

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
onal Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

performance and address the limitations of earlier models, thereby achieving greater accuracy in distinguishing between different Iris species.

Hussien, A. Y. (2022) conducted a study with the primary goal of identifying distinct flower species using a variety of machine learning classifiers. The objective was to accurately classify and differentiate between species based on their physical attributes by applying and evaluating multiple classification techniques. This research highlights the effectiveness of machine learning methods in enhancing the accuracy of botanical classification.

Gupta, Tina, and colleagues (2022) proposed a method that utilizes three classification models—Support Vector Machine (SVM), Logistic Regression, and K-Nearest Neighbors (KNN)—along with Exploratory Data Analysis (EDA) for data preprocessing. When evaluated on the Iris dataset, the models achieved maximum accuracy scores of 96.43% for SVM, 98.21% for Logistic Regression, and 94.64% for KNN, demonstrating the effectiveness of these techniques in flower species classification.

Pawar, Lokesh,andcolleaguesconcentrated oncreating adistinctiveclassificationstrategy todeterminethe plant's iris To be able to classify the floral pattern. It is advised that an ideal ensemble model become accustomed to identify and classify the pattern. To enhance, an ensemble model is suggested. performance compared to the original models of Bayesnet, Random Forest, Adaboost, OneR, and Decision Tree. (Lokesh and others, 2022).\

Singh, Anshuman, and Rohan Akash (2022) developed a machine learning model for flower classification using the *Iris* dataset and the Flask web framework. The dataset includes 50 samples each of *Iris virginica*, *Iris versicolor*, and *Iris setosa*, with recorded measurements of petal and sepal length and width, expressed in inches. Their work focused on building a user-friendly web application that utilizes these floral features to accurately classify *Iris* species, demonstrating the practical integration of machine learning and web technologies for botanical identification tasks.

III. METHODOLOGY

To establish a web application based on machine learning in relation to the classification of iris flower species, this study follows a defined process. The building of web applications, model training, and data preprocessing are some one of the essential components of the methodology.

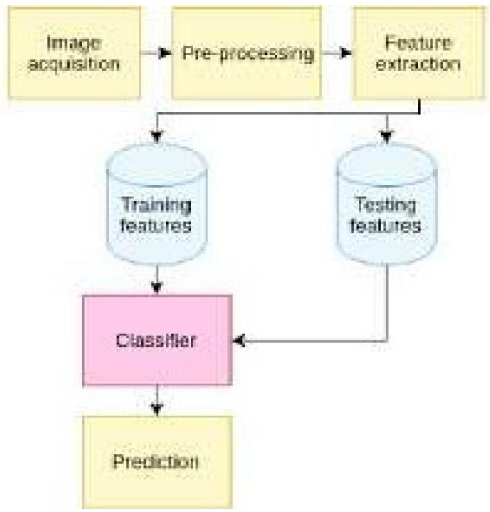


Figure 2: Flow Diagram of Data Processing

Data Preprocessing

During the data preprocessing phase, we begin with the well-known Iris dataset, which contains measurements of various Iris flower characteristics such as petal length, petal width, sepal length, and sepal width, along with the corresponding species labels. To ensure data quality and reliability, the dataset is first cleaned by removing any missing values or inconsistencies. After cleaning, the data is split into two subsets: a training set comprising 70–80% of the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

data, used to train the machine learning model, and a testing set comprising the remaining 20–30%, used to evaluate the model's performance and generalization capability.

Model Training

We employ the Random Forest Classifier, a potent a well-known ensemble learning algorithm accuracy in classification tasks, for the training of the model phase. By utilizing the training dataset, the classifier is trained to identify patterns and connections between the input attributes and the target variable (iris species). Metrics like accuracy are computed to assess the model's performance once it has been trained on the testing dataset. The completed model is stored to disk for later usage once the parameters are adjusted to improve accuracy.

Web Application Development and Deployment

The backend of the application is developed using Flask, a lightweight Python web framework that handles HTTP requests and processes input data. Flask facilitates seamless interaction with the trained machine learning model, enabling users to generate predictions. The frontend is built using HTML, CSS, and JavaScript, providing a user-friendly interface where users can enter flower measurements, register, log in, and view predictions.

The application supports two user roles:

- Administrators: Can manage user accounts and update the FAQ section.
- Users: Can make predictions based on input data and view FAQs.

Deployment

For deployment, the web application is hosted on a cloud platform such as Heroku, making it accessible to users online. This phase includes configuring the environment, setting up the web server, and ensuring smooth operation in a production setting. Post-deployment, the application undergoes rigorous testing to identify and fix any bugs or issues. Regular updates and maintenance are performed to ensure security, functionality, and continued user satisfaction.

IV. ALGORITHM USED

Random Forest Classifier

The Random Forest Classifier is an ensemble learning technique commonly used for classification tasks. It builds multiple decision trees using different subsets of the training data and selects random features at each split. Each tree generates its own prediction, and the final classification is determined by a majority vote across all trees. This approach reduces the risk of overfitting and improves accuracy compared to using a single decision tree. In this study, the Random Forest Classifier is employed to identify the species of Iris flowers based on sepal and petal measurements. Once trained, the model is integrated into a web application to provide real-time predictions based on user inp

Advantages

- High Accuracy: Random Forest often outperforms individual decision trees by combining the results of multiple trees, leading to higher predictive accuracy.
- Robustness to Overfitting: The use of techniques like feature bagging and bootstrapping reduces the risk of overfitting, making the model more generalizable.
- Handles High Dimensionality: It is well-suited for datasets with a large number of features, efficiently managing complex data.
- Feature Importance: Random Forest provides insights into which features are most influential by offering a built-in measure of feature importance.

V. RESULT

As part of this study, a web application was developed to identify various species of Iris flowers using the Random Forest Classifier. After training the model with the Iris dataset and conducting accuracy testing, the classifier

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, February 2025 Impact Factor: 7.67

demonstrated high performance and reliability. The web application enables users to input floral measurements and receive real-time predictions of the flower species. Additionally, it offers administrative functionalities for managing user accounts and content, while also providing users with access to the dataset and frequently asked questions (FAQs). Overall, the system effectively integrates machine learning with a user-friendly interface, delivering accurate flower classifications along with practical and accessible features..

VI. CONCLUSION

The Iris flower classification project serves as a prime example of how machine learning techniques can be applied to automate species identification. This system features a web-based application with dedicated modules for administrators and users, offering an efficient and structured approach to managing and predicting flower species. The machine learning algorithms—trained extensively on a dataset containing measurements of Iris flower attributes—deliver highly accurate predictions based on user-provided input. Key features such as advanced administrative controls, user participation in the FAQ section, and direct access to the dataset contribute to a robust and user-friendly experience. This work not only highlights the practical applications of machine learning in simplifying complex classification tasks but also showcases its effectiveness in building interactive, real-world solutions.

REFERENCES

[1] H. D., Y. T. Reddy, L. K., and S. C. Davusaheb Saajid, "Iris Flower Species Detection using Machine Learning Technique," *2023 3rd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, 2023, pp. 1–4, doi: 10.1109/CONIT59222.2023.10205566.

[2] J. P. Pinto, S. Kelur, and J. Shetty, "Iris Flower Species Identification Using Machine Learning Approach," *2018 4th International Conference for Convergence in Technology (I2CT)*, Mangalore, India, 2018, pp. 1–4, doi: 10.1109/I2CT42659.2018.9057891.

[3] Y. Pachipala, H. C. Maddipati, P. S. Udimudi, V. Thota, V. N. Sree, and L. R. Burra, "Iris Flower Classification by Using Random Forest in AWS," *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2022, pp. 1772–1777, doi: 10.1109/ICICCS53718.2022.9788415.

[4] A. Eldem, H. Eldem, and D. Üstün, "A Model of Deep Neural Network for Iris Classification With Different Activation Functions," *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, Malatya, Turkey, 2018, pp. 1–4, doi: 10.1109/IDAP.2018.8620866.

[5] T. Gupta, P. Arora, R. Rani, G. Jaiswal, P. Bansal, and A. Dev, "Classification of Flower Dataset Using Machine Learning Models," *2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST)*, Delhi, India, 2022, pp. 1–6, doi: 10.1109/AIST55798.2022.10065178.

[6] V. Arya and R. K. Rathy, "An Efficient Neuro-Fuzzy Approach for Classification of Iris Dataset," *2014 International Conference on Reliability Optimization and Information Technology (ICROIT)*, Faridabad, India, 2014, pp. 161–165, doi: 10.1109/ICROIT.2014.6798304.

[7] A. Y. Hussien, "Flower Species Recognition Using Machine Learning Classifiers," *Academic Journal of Nawroz University*, vol. 11, no. 4, pp. 469–475, 2022. doi: 10.25007/ajnu.v11n4a1636.

[8] T. Gupta et al., "Classification of Flower Dataset Using Machine Learning Models," *2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST)*, IEEE, 2022.

[9] L. Pawar et al., "Optimised Ensembled Machine Learning Model for Iris Plant Classification," *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, IEEE, 2022.

[10] A. Singh and R. Akash, "Flower Classifier Web App Using ML & Flask Web Framework," *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, IEEE, 2022.