International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Iris Classification Using Machine Learning

Gonti Suchitra

B. Tech Student, Department of IT, GMR Institute of Technology, Rajam-532127, Andhra Pradesh, India
Email: 21341A1245@gmrit.edu.in

ABSTRACT

Classifying iris flowers is a classic problem in machine learning and pattern recognition. In the current system, interest distribution is made with the help of KNN and Logistic regression. However, identifying flowers by their characteristics has some disadvantages. To overcome this drawback, this paper proposes a new method using the support vector machine (SVM) algorithm to solve this problem. SVM is known for its results in binary and multi-class operations; This makes it ideal for the iris dataset, which has three different species: Setosa, Versicolor and Virginica. Our work begins with preprocessing of the Iris dataset, including data normalization and feature selection to improve the performance of SVM. Through extensive testing and cross-validation, SVM hyperparameters have been fine-tuned to improve the accuracy and capability of the model. The results show that the SVM algorithm has the best performance in correctly classifying iris flowers and achieves high accuracy. Additionally, we gain insight into the importance of feature selection in developing accurate classification. Our findings support the continued development of SVM-based classification studies and provide important guidance to machine learning practitioners.

Keywords: Iris flower classification, Support Vector Machine (SVM), machine learning, pattern recognition, multi-class classification, feature selection, kernel functions, hyperparameter tuning, data preprocessing.

INTRODUCTION

The iris flower classification problem provides a hands-on introduction to machine learning, enabling practitioners to grasp essential concepts, data preprocessing, model training, evaluation, and deployment. It's a stepping stone toward more complex classification tasks and a foundation for understanding various machine learning algorithms. Iris flower classification using machine learning serves as an educational tool for understanding fundamental concepts in machine learning and classification. It has real-world applications in botany, agriculture, and environmental sciences for automating the classification of iris species based on their features. Furthermore, the iris flower classification problem serves as a benchmark for testing and comparing the performance of different machine learning models, fostering a deeper comprehension of algorithmic strengths and weaknesses. Its simplicity makes it an ideal starting point for beginners while providing a robust foundation for more advanced studies. The knowledge gained from this exercise goes beyond iris classification, providing transferable skills that can be applied to many fields. As machine learning continues to evolve, iris datasets remain a timeless resource for understanding and understanding the interactions between data and algorithms in distributed operations. Celebrating its first anniversary, the iris classification problem remains one of the most difficult and valuable moments in machine learning.

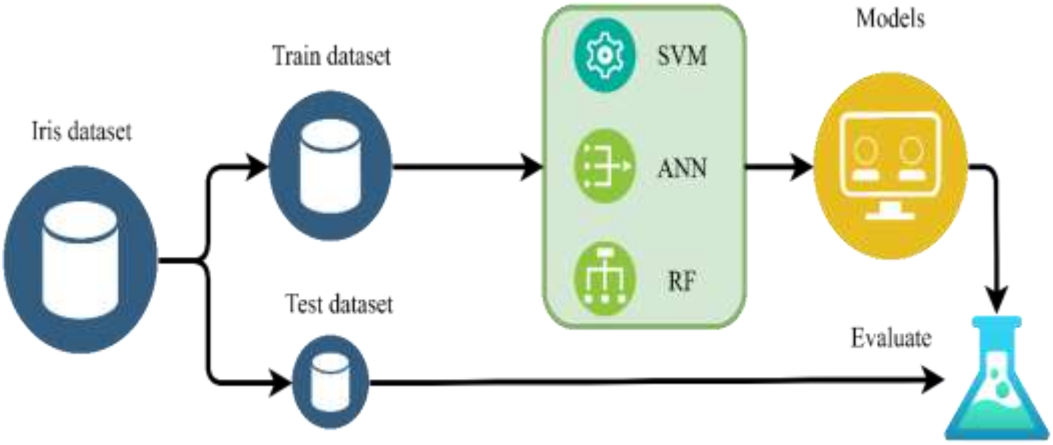
RESEARCH APPROACH

The paper employs The iris classification task was performed using three different learning algorithms: Random Forest (RF) algorithm, Support Vector Machine (SVM), and Artificial Neural Network (ANN). The RF algorithm implemented in Python turned out to be the most successful among the training examples with 97% accuracy. There are three different groups of iris data (setose, versicolor and virginica) and as the basis of this study, the RF algorithm has been cleverly used to model the data.

In addition to RF, SVM algorithms, especially those using the Support Vector Classification (SVC) class, are also included in the classification work. The design of the ANN algorithm, which also plays an important role in the classification process, is shown in Figure 1. In this study, the performance evaluation method calculated with the receiver function based on the ROC curve (AUC) characteristic curve (ROC) is used.), sensitivity, accuracy, precision and F1 score. These metrics help evaluate the effectiveness of the training model across various dimensions, providing a strong benchmark for classification algorithms.

The iris dataset contains three different iris types to test the performance of machine learning classifiers. With the successful application of RF, SVM and ANN algorithms, this study not only demonstrates their performance in iris classification but also demonstrates the importance of various method benchmarks in evaluating the effectiveness of such models.

METHODOLOGY



The iris classification task was performed using three different learning algorithms: Random Forest (RF) algorithm, Support Vector Machine (SVM), and Artificial Neural Network (ANN). The RF algorithm implemented in Python turned out to be the most successful among the training examples with 97% accuracy. There are three different groups of iris data (setose, versicolor and virginica) and as the basis of this study, the RF algorithm has been cleverly used to model the data.

In addition to RF, SVM algorithms, especially those using the Support Vector Classification (SVC) class, are also included in the classification work. The design of the ANN algorithm, which also plays an important role in the classification process, is shown in Figure 1. In this study, the performance evaluation method calculated with the receiver function based on the ROC curve (AUC) characteristic curve (ROC) is used.), sensitivity, accuracy, precision and F1 score. These metrics help evaluate the effectiveness of the training model across various dimensions, providing a strong benchmark for classification algorithms.

The iris dataset contains three different iris types to test the performance of machine learning classifiers. With the successful application of RF, SVM and ANN algorithms, this study not only demonstrates their performance in iris classification but also demonstrates the importance of various method benchmarks in evaluating the effectiveness of such models.

RESULTS

The article on "Using machine learning for iris classification" explores the field of iris classification from a machine learning perspective, specifically using scikit-learn tools. An integrated semi-cognitive method for data extraction to identify iris flower species. The machine learning model is carefully trained using existing data and the next prediction model will be used to predict the type of invisible object based on the received model. > Rigorous modeling of overall usage using techniques such as cross-matching and regularization is important to address limitations such as overfitting and bias. Highlighting the subtle differences between predicting unseen data and clarifying the labels of learning data is done by emphasizing showing the performance pattern for both groups, increasing the transparency of the evaluation process.

In order to reduce dependence on a single library such as Scikit-Learn, this research is encouraged to explore a variety of machine learning tools in different libraries, including the integration of deep learning such as TensorFlow or PyTorch. Additionally, issues of ethics and bias need to be discussed to ensure transparency and accountability in research. Together, these proposed improvements aim to promote a better understanding of technology learning systems, improve model details, clarify evaluation methods, differentiate research and raise awareness of ethics, ultimately leading to greater scholarship and research.

CONCLUSION

In summary, it is a valuable learning experience for new operators when solving iris classification problems through machine learning. This exercise provides a way to master the fundamentals of machine learning, including preprocessing, model training, evaluation, and even deployment. The simplicity and ease of use of the Iris dataset make it an excellent starting point for students to expertly interpret data and understand the nuances of various machine learning algorithms. Apart from its important teaching, the iris classification model is also a simple task for further classification as the basis for responding to complex challenges in many fields. Skills gained in selecting and interpreting patterns can become transferable assets for solving real-world problems, from medical diagnosis to image recognition. Moreover, the practical use of iris classification is remote. Automation of iris type identification in botany, agriculture and environmental science streamlines the research process and leads to deeper understanding of plant biodiversity, precision agriculture and ecological studies. This utility demonstrates the suitability and suitability of the model in solving today's problems in this area. More importantly, using machine learning to classify iris not only provides a better understanding of the underlying concepts, but also allows doctors to apply their skills to real situations. It shows the relationship between learning and practical application and demonstrates the difference of machine learning in understanding and solving complex problems in different domains.

References

1. Bayrakçı, H. C., Keşkekçi, A. B., & Arslan, R. (2022). Classification of Iris Flower by Random Forest Algorithm.
2. Rao, T. S., Hema, M., Priya, K. S., Krishna, K. V., & Ali, M. S. (2021). Iris Flower Classification Using Machine Learning.
3. Prathima, P., & Ranjith Kumar, T. (2021). A Model Identifying Iris Species using Machine Learning.
4. Sogir, S. Mithy, S. A., Hossain, S., Akter, S., Honey, U., & B. (2022). Classification of Iris Flower Dataset using Different Algorithms.
5. Prathima, P., & Ranjith Kumar, T. (2021). Comparison on Iris Dataset Using Classification Techniques.
6. S. T. Halakatti and S. T. Halakatti, "Identification Of Iris Flower Species Using Machine Learning," vol. 5, no. 8, pp. 59–69, 2017.
7. Asmita Shukla, Ankita Agarwal, Hemlata Pant, and Priyanka Mishra, "Flower Classification using Supervised Learning," Int. J. Eng. Res., vol. V9, no. 05, pp. 757–762, 2020.