

# Iris Flower Classification

[Za Neih Cinzah, Lukas Gurvin, Zuheib Alasoow],

[04.11.24].

## PROBLEM DESCRIPTION

Our machine learning project is about classifying different species of iris flowers into 3 different categories setosa, versicolor and virginica. We do this by using four different measurements which is inputted by the user. These four measurements are sepal length, sepal width, petal length and petal width. The model is trained on these four features to predict which type of iris flower the user has.

## SCOPE

The goal of this project is to improve the species identification process by creating an interactive, web-based iris classification system. This will reduce the errors in botany while automating a traditionally manual task.

The existing solutions are only manually measurement. It will take time and there will be risk of human error, while in machine learning there will be less time-consuming and objective solutions based on data.

## METRICS

Business Metric: If a model accuracy of at least 80% is reached the project is considered a success.

Accuracy: Measures the percentage of correct predictions.

Classification Report: Includes precision and recall for each class to evaluate performance for each flower type.

Latency: Fast response time is required, especially for real-time predictions on the web application.

The project is done using multiple components, we collected the dataset from Kaggle and used python. We also used Flask web application for users' interaction. We used a pickle file to be accessible for predictions for the users within the web application.

The stakeholders will be botanists, environmental researchers, and educational institutions that benefit from fast, reliable iris flower classification.

Our timeline consists of both preparation and working with the project. This is the timeline from week to week.

Week 1: Data collection and preparation

Week 2: Model development and testing

Week 3: Web application development

Week 4: Integration and final evaluation

The resources we used were a computer where we could both use python and flask. We used VS Code to write our code. We also had to use Google and YouTube for information and ideas for the project.

There was also the usage of Ai and GPT for the project to help us understand the project.

## DATA

Dataset: The project uses the well-known Iris dataset, which contains 150 samples divided across the three species.

Data Source and Labels: The data is publicly available and pre-labeled. Labels are used to train the model by associating the correct species with their respective features.

Data Processing: The data consists of four features and one label. Preprocessing includes scaling methods to enhance model performance.

## **MODELING**

Model Choice: The undertaking makes use of a logistic regression version, which is thought to carry out properly on category responsibilities of this nature.

Baseline Performance: Expected baseline accuracy for iris category with logistic regression is round 90%.

This version used a test dataset which it had not seen before and scored brilliantly with an accuracy of 92% of predicting the outcome.

Error Analysis and Feature Importance: With the use of the category report misclassification can be observed. To recognize its function each features significance may be estimated within the category process.

## **DEPLOYMENT**

Model Deployment: The model is deployed with the use of the Flask net software where users can obtain predictions immediately after inputting the flower measurements.

Monitoring and Maintenance: The machine can be monitored for accuracy, and the version may be up to date with new statistics or changes as needed. The net software consists of a homepage and a outcomes web page to offer remarks to the person primarily based totally at the version.

## **REFERENCES**

- Team, D. (2021, October 20). Iris Flower Classification Project using Machine Learning - DataFlair. DataFlair. <https://data-flair.training/blogs/iris-flower-classification/>
- Scikit-learn. (2019). scikit-learn: machine learning in Python. Scikit-Learn.org. <https://scikit-learn.org/stable/>
- Used GPT and Ai for help and information regarding the project