NGF COLLEGE OF ENGINEERING AND TECHNOLOGY, PALWAL (HARYANA)

IRIS FLOWER CLASSIFICATION IN DATA SCIENCE

Submitted By   **Submitted To**

Ishwa   **Ms. Bhawna**

B. tech CSE 7th Sem   **Assistant Professor**
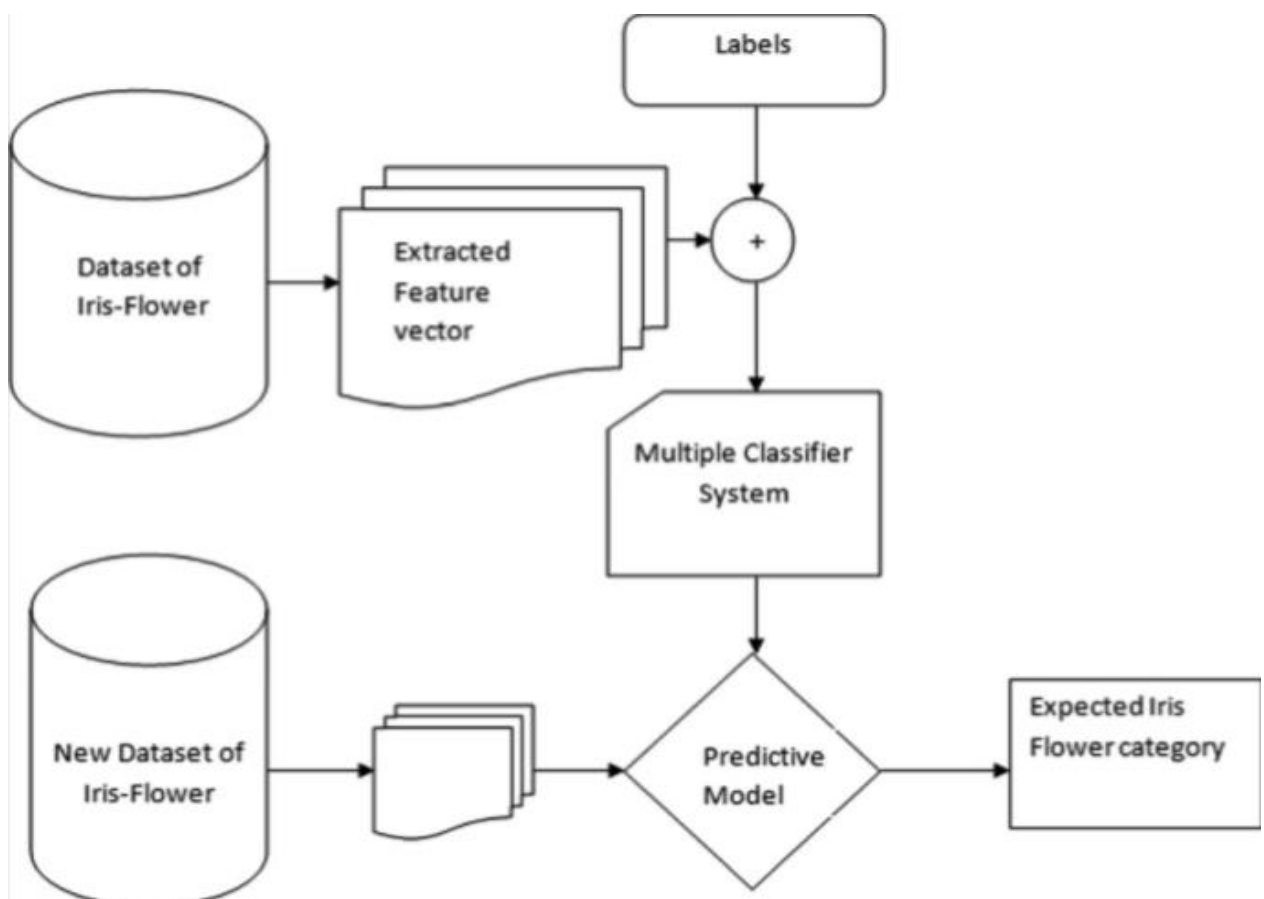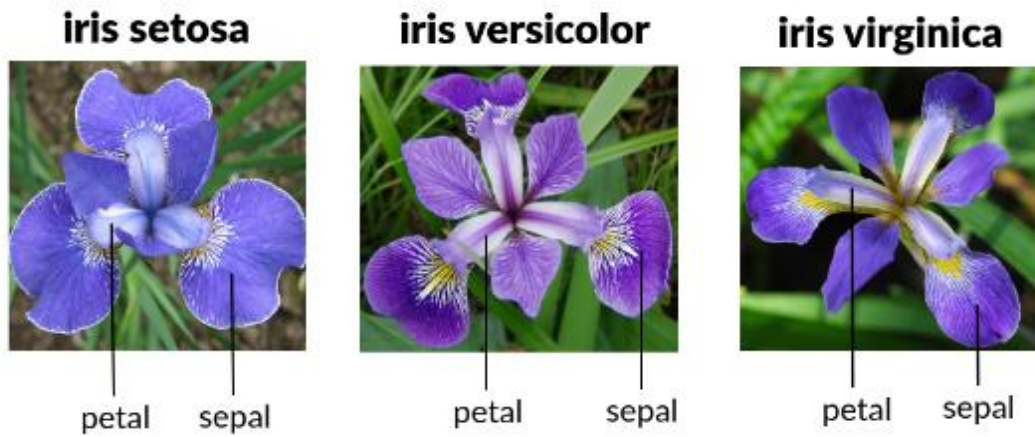
S19CSE022

INTRODUCTION

The iris classification in data science is end to end machine learning project or data science project for iris flower classification. The aim of the iris flower classification is to predict flowers based on their specific features and separate different species and to identify them. It contains five columns namely – Petal Length, Petal Width, Sepal Length, Sepal Width and Species Type. All these lengths were in centimetres. And dependent feature, which will be the output for the model, is species. It contains the name of the species to which that particular flower with those measurements belongs. Iris is a flowering plant, the researchers have measured various features of the different iris flowers and recorded them digitally.

OBJECTIVE

The objective of iris classification in data science is to classify the flowers according to their traits. The central goal is to design a model that makes proper classification for new flowers. The iris data set contains fifty instances of each of three species. Iris Flower Classification is a Machine Learning Project. The iris dataset contains three classes of flowers, Versicolor, Setosa, Virginica, and each class contains 4 features, 'Sepal Length', 'Sepal width', 'Petal length', 'Petal width'. The aim of the iris flower classification is to predict flowers based on their specific features and separate different species and to identify them.

DIAGRAM

iris setosa   iris versicolor   iris virginica

petal   sepal   petal   sepal   petal   sepal

Labels

Dataset of Iris-Flower → Extracted Feature vector → +

Multiple Classifier System

New Dataset of Iris-Flower → Predictive Model → Expected Iris Flower category

REQUIREMENTS

- Ide –Anaconda
- Language used- Python
- Step-1: Load the data
- Step-2: Analyse and Visualize the dataset
- Step-3: Model Evaluation
- Step-4: Testing the Model

PROCEDURE

- Select the hardware and software suitable for ML.
- Acquire the ML dataset.
- Explore the ML dataset.
- Choose and create the ML model.
- Pre-process the dataset for training.
- Carry out training, prediction and testing.
- Evaluate mode's performance.
- Vary parameters of the ML model manually.
- Tune the ML model systematically.

Conclusion

We can easily distinguish Setosa. Due to a lack of data, it is hard to distinguish between Virginica and Versicolor. In this project, we learned to train our own supervised machine learning model using Iris Flower Classification Project. Through this project, we learned about machine learning, data analysis, data visualization, model creation, etc.

REFERENCE

- https://seaborn.pydata.org/generated/seaborn.pairplot.html
- https://seaborn.pydata.org/generated/seaborn.FaceGrid.html
- https://www.kaggle.com/arshid/iris-flower-dataset
- https://www.datacamp.com/community/tutorial/machine-learning-in-r