Khawaja Fared University of Engineering & Information Technology
Rahim Yar Khan

A project on
"Iris Flower Prediction"

Submitted by :
Ayesha Shabbir(cosc222102018)
Sidra Kousar (cosc222102010)
Submitted to:
Sir Imran Bilal
Associate professor of department of data science

Submission Date:

Department of Data Science & Artificial Intelligence Spring 2024 |
Program: BS Data Science

Abstract

The Iris Flower Prediction project encapsulates the fusion of botany and machine learning, leveraging sophisticated algorithms to accurately classify iris flowers based on their distinct characteristics. This initiative aims to streamline the classification process, traditionally a meticulous task for botanists, by training computer models to recognize and differentiate between various iris species. Utilizing the renowned Iris flower dataset from the UCI Machine Learning Repository, which comprises 150 samples across three species (Iris virginica, Iris versicolor, and Iris setosa), this project delves into the analysis of four key features: sepal width, sepal length, petal width, and petal length, measured in centimeters. The documentation delineates the methodology, emphasizing supervised learning and classification techniques. Furthermore, it elaborates on the hardware and software requirements essential for the project's implementation. The ultimate objective is to harness technology to enhance the accuracy and efficiency of iris flower identification, benefiting botanists, horticulturists, and enthusiasts alike, while facilitating deeper insights into the evolutionary and environmental aspects of these captivating plants.

Chapter 1

INTRODUCTION

1.1 Introduction

1.1. Machine Learning

Machine learning is a process of feeding a machine enough data to train and predict a possible outcome using the algorithms. the more the processed or useful data is fed to the machine the more efficient the machine will become. When the data is complicated it learns the data and builds the prediction model. It is state that more the data, better the model, higher will be the accuracy. There are many ways for machine learning i.e. supervised learning, unsupervised learning and reinforcement learning.

1.2. Supervised Learning

In supervised learning machine learning model learns through the feature and labels of the object. Supervised learning uses labeled data to train the model here, the machine knew the features of the object and labels associated with those features or we can say that the supervised learning uses the set of data where the labels or the desired outcomes are already known. It is allowed to prediction about the unseen or future data.

1.3 Classification

Classification is one of the major data mining processes which maps data into predefined groups. It comes under supervised learning method as the classes are determined before examining the data. For applying all approaches to performing classification it is required to have some knowledge of the data. Usually, the knowledge of the data helps to find some unknown patterns. The aim of pattern classification is to building a function that provides output of two or more than two classes from the input feature. The dataset for this project carried out from the UCI Machine Learning Repository. The Iris flower data set introduced by the British statistician and biologist Ronald

Fisher that’s why it 2 is also known by Fisher's Iris data set and it is a multivariate data set. The use of multiple measurements in taxonomic problems as an example of linear discriminant analysis. The expectation from mining iris data set would be discovering patterns from examining sepal and petal size of the iris plant and how the prediction was made from analyzing the pattern to predict the class of iris plant. In upcoming years, using the classification and pattern recognition other flowers can be individually distinguish to each other. It is unmistakably expressed that the sort of relationship that being mined utilizing iris dataset would be a classification model.

1.3. HARDWARE REQUIREMENTS:

• Processor – i3

• 2 GB RAM

• Memory – 5 GB

1.5. SOFTWARE AND LIBRARIES REQUIREMENTS:

• python 3.7.2

• Jupyter Notebook

• sklearn

• csv

• numpy

• pandas

• matplotlib

Iris Flower Prediction project

The Iris Flower Prediction project represents an exciting intersection of botany and machine learning. By leveraging advanced algorithms, we aim to accurately classify iris flowers into different species based on their unique characteristics. This introduction serves to provide context for the project, emphasizing its importance and outlining the objectives we seek to achieve.

Welcome to the Iris Flower Prediction project, where we're blending the worlds of botany and machine learning to do something pretty cool. Basically, we're using fancy computer techniques to figure out exactly what type of iris flower we're looking at, just by checking out its special features. This introduction is like the start button for our project, giving you a heads-up on why it's important and what we're hoping to do.

So, imagine you're out in a garden, surrounded by all sorts of beautiful iris flowers. They come in different shapes, sizes, and colors, and it can be tricky to tell them apart sometimes. That's where our project comes in handy. We're using super-smart computer algorithms to study these flowers up close and learn how to recognize each type accurately.

Why does this matter? Well, it's not just about identifying flowers for fun. Understanding different types of iris flowers helps scientists learn more about nature, how plants evolve, and how they interact with their environment. Plus, it can help gardeners and farmers grow better plants and preserve rare species.

Now, let's talk about what we're aiming to do with this project. Our main goal is to create computer programs that can look at iris flowers and say, "Yep, that's a setosa," or "Nope, that's a virginica." We want these programs to be really good at their job, so they can help us learn more about these beautiful flowers and how to take care of them better.

So, buckle up, because we're about to dive deep into the world of iris flowers and machine learning. Get ready to learn some cool stuff and maybe even see some flowers in a whole new light!

Background Information:

Iris flowers are famous for being really colorful and having lots of different shapes and sizes. They're all over the world, with more than 300 different kinds out there. Imagine a huge family of flowers, each one with its own unique look! People who study plants, called botanists, have always been fascinated by irises. They want to know everything about them: where they came from, how they've changed over time, and why they're important in nature.

But here's the thing: telling all these iris flowers apart isn't easy. It's like trying to pick out specific people in a crowd – they might all look similar, but they're actually different in small ways. For botanists, figuring out which iris is which often means carefully studying the details of each flower, like its shape and size.

So, one of the big challenges in iris research is getting really good at classifying, or sorting, these flowers into the right groups. That's where our project comes in. We're using fancy technology to help us do this job better and faster. Instead of relying solely on human eyes, we're training computers to recognize the unique features of each iris flower. It's like teaching a robot to be a flower expert!

By doing this, we hope to make it easier for botanists to study irises and learn more about their amazing world. Plus, who knows? Maybe we'll discover some cool new things along the way!

Importance of Iris Flower Prediction:

Being really good at figuring out what kind of iris flower we're looking at is super useful for a bunch of different reasons. First off, it helps scientists who study plants (we call them botanists) do their job better. When we can tell exactly which species of iris we're dealing with, it makes it easier to study things like how many different kinds there are, how they've changed over time, and how they interact with their environment.

But it's not just for science nerds – knowing all about iris flowers can actually help people who grow plants for a living, like gardeners and farmers. When they know which species of iris they're working with, they can do cool stuff like breeding new types of flowers, protecting rare ones from dying out, and figuring out the best ways to take care of them.

And hey, even if you're not into science or gardening, iris flower prediction is still pretty neat. It's like solving a fun puzzle, trying to match each flower to its correct species. Plus, it's a great way for students and anyone who loves plants to learn more about the fascinating world of botany. So, whether you're a scientist, a gardener, or just someone who thinks flowers are cool, being able to predict iris flower species is definitely a handy skill to have!

Objectives of the Project:

The main goal of our Iris Flower Prediction project is to use computers to help us figure out exactly what type of iris flower we're looking at. To do this, we've set out some specific things we want to achieve:

Get a bunch of data: First, we need lots of information about iris flowers – like their size, shape, and other characteristics. We're going to gather all this data and make sure it's ready for our computer programs to use.

Take a good look at the data: Once we've got our data, we'll spend some time looking at it closely. We'll use things like numbers and graphs to understand what the data is telling us about the iris flowers.

Teach the computers: Next, we'll train our computer models to recognize patterns in the data. We'll show them examples of different types of iris flowers and teach them how to tell the difference between them.

Test our models: Once our computer models are trained, we need to make sure they're doing a good job. We'll give them new data that they haven't seen before and see how well they can predict the species of iris flowers.

Talk about what we found: Finally, we'll talk about what we learned from our project. We'll discuss what worked well, what challenges we faced, and how our work could be useful for botany and other fields.

Overall, our goal is to use technology to make it easier and more accurate to identify different types of iris flowers. We hope that our project will help both scientists and computer experts learn more about these beautiful plants and how we can use machines to understand them better.

Iris Data:

The dataset for this project originates from the UCI Machine Learning Repository. The Iris flower data set or Fisher's Iris data set is a multivariate data set. The data set consists of 50 samples from each of three species of Iris (Iris virginica, Iris versicolor and Iris setosa).

• Four features were measured from each sample (in centimeters):

- Length of the petals
- Width of the petals
- Length of the sepals
- Width of the sepals

Understanding the data:

Iris flower data set contains the observation data with 150 samples. Since the dataframe has four features (Sepal width, sepal length, petal width and petal length) with 150 samples belonging to either of the three target classes. In this step we going into the mathematics of the dataset to find out the standard deviation, mean, minimum value and the four-quartile percentile of the data. Since the dataframe has four features (Sepal width, sepal length, petal width and petal length) with 150 samples belonging to either of the three target classes. In this step we going into the mathematics of the dataset to find out the standard deviation, mean, minimum value and the four-quartile percentile of the data.

Conclusion:

The Iris Flower Prediction project successfully demonstrates the integration of botany and machine learning, enhancing iris flower identification through data-driven techniques. Utilizing the UCI Iris dataset, computer models were trained to accurately classify iris species based on distinct features. This interdisciplinary approach highlights the potential of technology to streamline botanical research, fostering collaboration and innovation. The project underscores the value of leveraging machine learning for ecological insights and conservation efforts, paving the way for future advancements in the field.