Mini Project

# Import necessary libraries

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

from sklearn import datasets

from sklearn.model\_selection import train\_test\_split

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import classification\_report, confusion\_matrix

# Step 1: Load the Iris dataset

iris = datasets.load\_iris()

# Convert to DataFrame for easier manipulation

df = pd.DataFrame(data=iris.data, columns=iris.feature\_names)

df['species'] = iris.target\_names[iris.target]

# Step 2: Data Exploration

print(df.head()) # Display the first few rows of the dataset

# Visualize using pairplot

sns.pairplot(df, hue='species')

plt.title("Iris Flower Pairplot")

plt.show()

# Step 3: Data Preprocessing (No missing values in this dataset)

# Step 4: Model Selection and Training

# Splitting the dataset into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(iris.data, iris.target, test\_size=0.2, random\_state=42)

# Creating and training the Random Forest model

model = RandomForestClassifier(n\_estimators=100, random\_state=42)

model.fit(X\_train, y\_train)

# Step 5: Making predictions

predictions = model.predict(X\_test)

# Step 6: Model Evaluation

print("Confusion Matrix:")

print(confusion\_matrix(y\_test, predictions))

print("\nClassification Report:")

print(classification\_report(y\_test, predictions))

# Optional: Visualizing the Confusion Matrix

conf\_matrix = confusion\_matrix(y\_test, predictions)

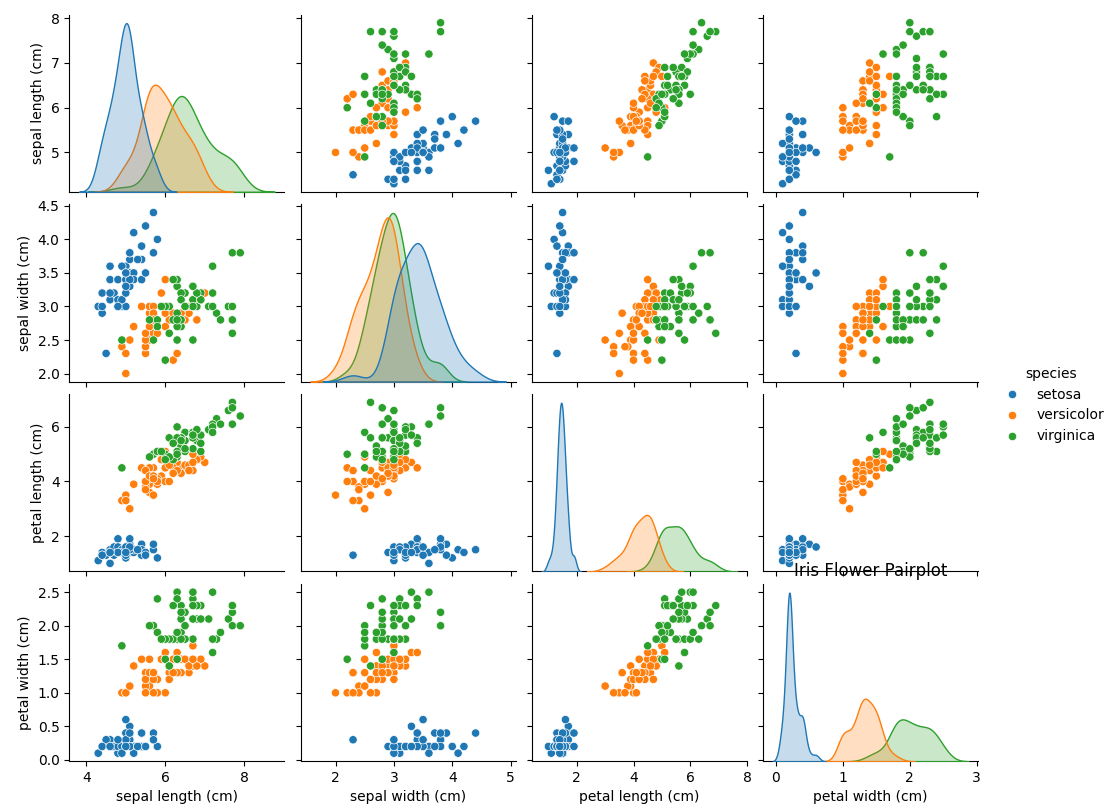
sns.heatmap(conf\_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=iris.target\_names, yticklabels=iris.target\_names)

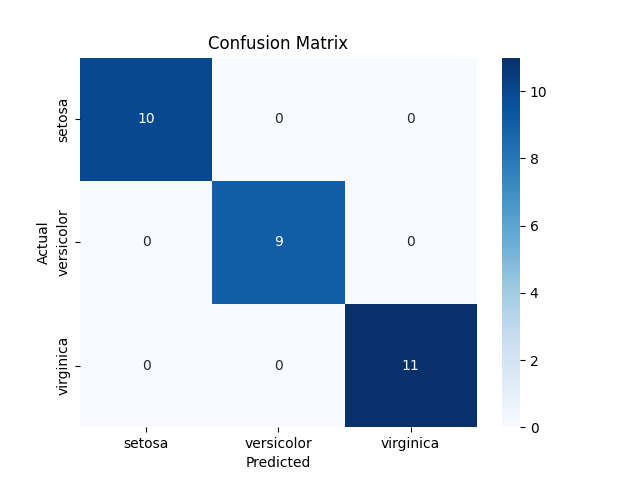
plt.ylabel('Actual')

plt.xlabel('Predicted')

plt.title('Confusion Matrix')

plt.show()

#OUTPUT



sepal length (cm) sepal width (cm) petal length (cm) petal width (cm) species

0 5.1 3.5 1.4 0.2 setosa

1 4.9 3.0 1.4 0.2 setosa

2 4.7 3.2 1.3 0.2 setosa

3 4.6 3.1 1.5 0.2 setosa

4 5.0 3.6 1.4 0.2 setosa

Confusion Matrix:

[[10 0 0]

[ 0 9 0]

[ 0 0 11]]

Classification Report:

precision recall f1-score support

0 1.00 1.00 1.00 10

1 1.00 1.00 1.00 9

2 1.00 1.00 1.00 11

accuracy 1.00 30

macro avg 1.00 1.00 1.00 30

weighted avg 1.00 1.00 1.00 30