Sydney Trout
Dr. Li
CPSC 4430
26 January 2022

Project 1 Report
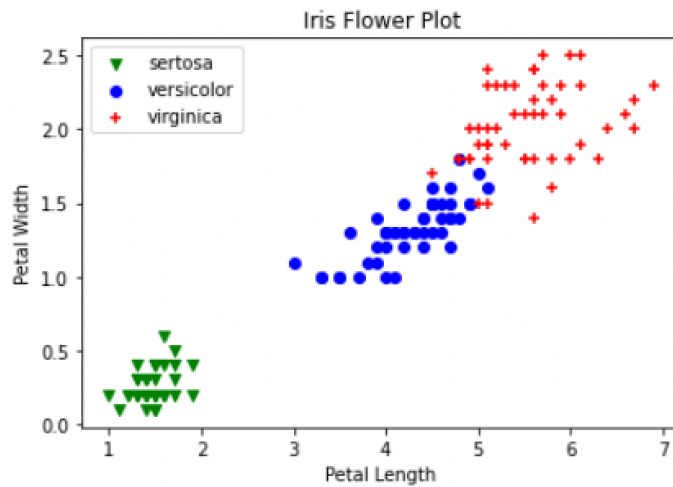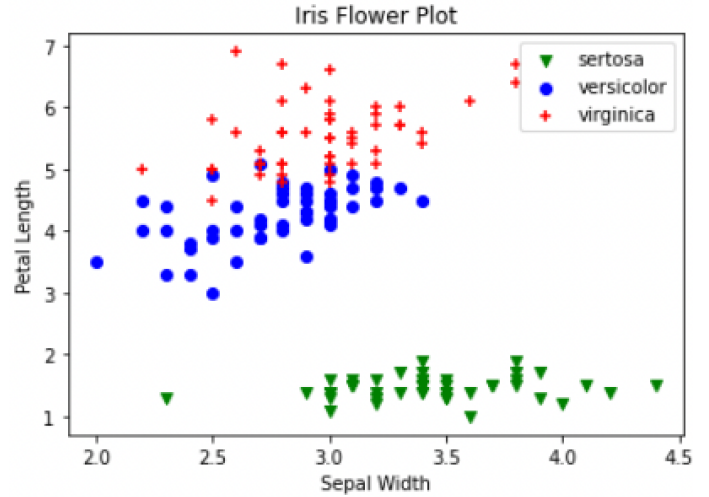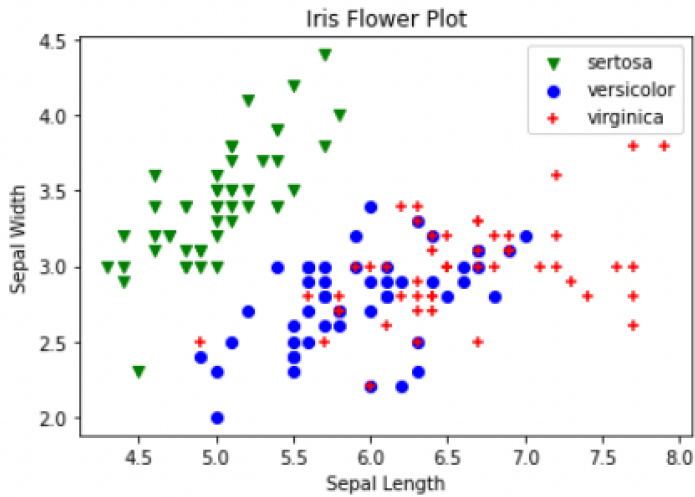
After reviewing all of the different plots that were created given the different iris variables, I believe that the best feature combination to classify the data is petal length and petal width. This plot can be seen above as it is the third image. This given plot stands out to me the most as it very clearly differentiates the three iris types. When looking at the other two plots above it and the rest of the plots that this data is able to create, there are many times where the data values of the different types merge together in the same range. This would make it hard to use these plots to classify an iris if you were measuring two of these values. But with petal length and petal width, the data is much easier to differentiate. The data shows that the sertosa have shorter and skinnier petals, the virginica have very long and wide petals, and the versicolor are right in between. Using this plot, I believe it would be much simpler to classify an iris, making petal length and width the best feature combination.

CODE COPY:

```
# Import libraries
import numpy as np
import matplotlib.pyplot as plt

# Main program
# Receive file name input from user
fileName = input("Enter the name of your file: ")
if fileName is None:
    print('Invalid file name')
    fileName = input("Enter the name of your file: ")

# Read first line info about file and convert to integers
dataFile = open(fileName, "r")
readData = dataFile.readline()
fileInfo = readData.split("\t")
rows = int(fileInfo[0])
features = int(fileInfo[1])

# Creates two-dimensional array of zeros
flowerData = np.zeros([rows,features+1])

# Reads the data from the file and stores it in the array
for i in range(rows):
    readData = dataFile.readline()
    data = readData.split("\t")
```

```python
    for j in range(features):
        flowerData[i,j] = float(data[j])
    if "setosa" in data[features]:
        flowerData[i,features] = 1
    elif "versicolor" in data[features]:
        flowerData[i,features] = 2
    else:
        flowerData[i,features] = 3


# Close the file
dataFile.close()

# Initialize value for the while loop
playAgain = "y"

# Loop until the user no longer wishes to plot
while (playAgain != "n"):

    # Reaceives plot features from the user for the vertical and horizontal axes and converts them
to integers
    print("You can do a plot of any two features of the Iris Data set")
    print("The feature codes are:")
    print("\t0 = sepal length")
    print("\t1 = sepal width")
    print("\t2 = petal length")
    print("\t3 = petal width")
    featCode1 = input("Enter feature code for the horizontal axis: ")
    featCode1 = int(featCode1)
    while(featCode1 != 0 and featCode1 != 1 and featCode1 != 2 and featCode1 != 3):
        print("Invalid feature code!")
        featCode1 = input("Enter feature code for the horizontal axis: ")
        featCode1 = int(featCode1)
    featCode2 = input("Enter feature code for the vertical axis: ")
    featCode2 = int(featCode2)
    while(featCode2 != 0 and featCode2 != 1 and featCode2 != 2 and featCode2 != 3):
        print("Invalid feature code!")
        featCode2 = input("Enter feature code for the vertical axis: ")
        featCode2 = int(featCode2)

    # Based on plot features, plots the data and labels the axes
```

```python
    if(featCode1 == 0 and featCode2 == 0):
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,0], flowerData[i,0], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,0], flowerData[i,0], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,0], flowerData[i,0], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Sepal Length")
        plt.ylabel("Sepal Length")
    elif featCode1 == 0 and featCode2 == 1:
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,0], flowerData[i,1], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,0], flowerData[i,1], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,0], flowerData[i,1], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Sepal Length")
        plt.ylabel("Sepal Width")
    elif(featCode1 == 0 and featCode2 == 2):
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,0], flowerData[i,2], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,0], flowerData[i,2], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,0], flowerData[i,2], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Sepal Length")
        plt.ylabel("Petal Length")
    elif(featCode1 == 0 and featCode2 == 3):
```

```python
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,0], flowerData[i,3], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,0], flowerData[i,3], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,0], flowerData[i,3], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Sepal Length")
        plt.ylabel("Petal Width")
    elif(featCode1 == 1 and featCode2 == 0):
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,1], flowerData[i,0], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,1], flowerData[i,0], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,1], flowerData[i,0], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Sepal Width")
        plt.ylabel("Sepal Length")
    elif(featCode1 == 1 and featCode2 == 1):
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,1], flowerData[i,1], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,1], flowerData[i,1], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,1], flowerData[i,1], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Sepal Width")
        plt.ylabel("Sepal Width")
    elif(featCode1 == 1 and featCode2 == 2):
        for i in range(rows):
```

```python
        if flowerData[i,features] == 1:
            s1 = plt.scatter(flowerData[i,1], flowerData[i,2], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,1], flowerData[i,2], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,1], flowerData[i,2], color = "red", marker = "+", label =
"virginica")
    plt.xlabel("Sepal Width")
    plt.ylabel("Petal Length")
elif(featCode1 == 1 and featCode2 == 3):
    for i in range(rows):
        if flowerData[i,features] == 1:
            s1 = plt.scatter(flowerData[i,1], flowerData[i,3], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,1], flowerData[i,3], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,1], flowerData[i,3], color = "red", marker = "+", label =
"virginica")
    plt.xlabel("Sepal Width")
    plt.ylabel("Petal Width")
elif(featCode1 == 2 and featCode2 == 0):
    for i in range(rows):
        if flowerData[i,features] == 1:
            s1 = plt.scatter(flowerData[i,2], flowerData[i,0], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,2], flowerData[i,0], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,2], flowerData[i,0], color = "red", marker = "+", label =
"virginica")
    plt.xlabel("Petal Length")
    plt.ylabel("Sepal Length")
elif(featCode1 == 2 and featCode2 == 1):
    for i in range(rows):
        if flowerData[i,features] == 1:
```

```python
            s1 = plt.scatter(flowerData[i,2], flowerData[i,1], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,2], flowerData[i,1], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,2], flowerData[i,1], color = "red", marker = "+", label =
"virginica")
    plt.xlabel("Petal Length")
    plt.ylabel("Sepal Width")
elif(featCode1 == 2 and featCode2 == 2):
    for i in range(rows):
        if flowerData[i,features] == 1:
            s1 = plt.scatter(flowerData[i,2], flowerData[i,2], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,2], flowerData[i,2], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,2], flowerData[i,2], color = "red", marker = "+", label =
"virginica")
    plt.xlabel("Petal Length")
    plt.ylabel("Petal Length")
elif(featCode1 == 2 and featCode2 == 3):
    for i in range(rows):
        if flowerData[i,features] == 1:
            s1 = plt.scatter(flowerData[i,2], flowerData[i,3], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,2], flowerData[i,3], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,2], flowerData[i,3], color = "red", marker = "+", label =
"virginica")
    plt.xlabel("Petal Length")
    plt.ylabel("Petal Width")
elif(featCode1 == 3 and featCode2 == 0):
    for i in range(rows):
        if flowerData[i,features] == 1:
```

```python
            s1 = plt.scatter(flowerData[i,3], flowerData[i,0], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,3], flowerData[i,0], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,3], flowerData[i,0], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Petal Width")
        plt.ylabel("Sepal Length")
    elif(featCode1 == 3 and featCode2 == 1):
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,3], flowerData[i,1], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,3], flowerData[i,1], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,3], flowerData[i,1], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Petal Width")
        plt.ylabel("Sepal Width")
    elif(featCode1 == 3 and featCode2 == 2):
        for i in range(rows):
            if flowerData[i,features] == 1:
                s1 = plt.scatter(flowerData[i,3], flowerData[i,2], color = "green", marker = "v", label =
"sertosa")
            elif flowerData[i,features] == 2:
                s2 = plt.scatter(flowerData[i,3], flowerData[i,2], color = "blue", marker = "o", label =
"versicolor")
            else:
                s3 = plt.scatter(flowerData[i,3], flowerData[i,2], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Petal Width")
        plt.ylabel("Petal Length")
    elif(featCode1 == 3 and featCode2 == 3):
        for i in range(rows):
            if flowerData[i,features] == 1:
```

```python
            s1 = plt.scatter(flowerData[i,3], flowerData[i,3], color = "green", marker = "v", label =
"sertosa")
        elif flowerData[i,features] == 2:
            s2 = plt.scatter(flowerData[i,3], flowerData[i,3], color = "blue", marker = "o", label =
"versicolor")
        else:
            s3 = plt.scatter(flowerData[i,3], flowerData[i,3], color = "red", marker = "+", label =
"virginica")
        plt.xlabel("Petal Width")
        plt.ylabel("Petal Width")
    else:
        break

    # Creates the plot title and legend
    plt.title("Iris Flower Plot")
    plt.legend((s1,s2,s3), ("sertosa", "versicolor", "virginica"))
    plt.show()

    # Prompts user for input on whether to create another plot
    playAgain = input("Would you like to do another plot? (y/n): ")
    while (playAgain != "y" and playAgain != "n"):
        print("Invalid input!")
        playAgain = input("Would you like to do another plot? (y/n): ")
```