

Joseph E. Palacios
R11585932

Project 1

3 Classes: Setosa, Versicolor, Virginica
4 Features: Sepal Length/Width, Petal Length/Width

The iris flower is distinguished by both the sepals and petals. The sepals are the outer leaves that appear to be wider and longer whereas petals are the inner leaves that appear to be skinner and shorter. There are variations between the three classes in both sepal/petal lengths and widths, so these features/measurements are a good way to distinguish between classes.

I do note that when plotting sepal length vs width. Setosa has its own cluster apart from Versicolor and Virginica which indicates that the measurements of sepal length and width are similar in Versicolor and Virginica and Setosa have differing measurements. More specifically, Setosa has wider, and shorter sepals compared to that of Versicolor and Virginica.

Now looking at the petal lengths and widths, Setosa again has its own cluster apart from Versicolor and Virginica however, there now seems to be some variation between Versicolor and Virginica. Setosa's petals are shorter and skinner compared to the other two classes. Virginica has petals that are longer and wider in comparison to Versicolor.

	Sepal Length	Sepal Width	Petal Length	Petal Width
Minimum	4.3	2.0	1.0	0.1
Maximum	7.9	4.4	6.9	2.5
Mean	5.843	3.057	3.758	1.199
Variance	0.681	0.189	3.096	0.577
Within-Class Variance	0.257	0.112	0.180	0.041
Between-Class Variance	0.421	0.076	2.914	0.536

From the heatmap, you can see that the Sepal Width has little to no correlation with the other features. The other features, however, have a very high correlation with the other features. The petal length and width have the highest correlation among features. There is a strong correlation with the features and their class as well. The only oddball seen here is the Sepal Width as previously mentioned.

It looks like the features PetL and PetW are best for classification since their measurements are separated by class and there's little to no overlap in measurements between classes. Setosa has the shortest measurements, Versicolor has the middle ground of measurements and Virginica has the longest measurements. However, looking at SepL and SepW, you can see many of the same measurements within the three classes. This would make it difficult to differentiate between classes when classifying. You can see those measurements of SepW varies the least between all three classes.

Joseph E. Palacios
R11585932

Classification Reports:

Using np.random.seed(11585932)
Found at line 163 in .py file.

Setosa vs All:

Batch Perceptron Misclassifications: 0
Converged in 2 epochs
BP Weight vector is: [-0.02913393 0.30445502 -0.42560216 0.17624956 0.03597005]
LS Weight vector is: [0.13205954 0.48569574 -0.44931423 -0.11494546 -0.76355422]
Least Squares Misclassifications: 0

Setosa vs 3 & 4:

Batch Perceptron Misclassifications: 0
BP Converged in 221 epochs
BP Weight vector is: [-0.2917528 0.35440523 0.48527811]
LS Weight vector is: [-0.5026581 0.01966852 1.53206671]
Least Squares Misclassifications: 1

Virginica vs All:

Batch Perceptron Misclassifications: 0
Converged in 10 epochs
BP Weight vector is: [-0.51983176 0.94009115 -0.34258167 0.43837517 0.58210205]
LS Weight vector is: [-0.09175217 0.40553677 0.00797582 1.10355865 -2.39056373]
Least Squares Misclassifications: 11

Virgi vs 3 & 4:

Batch Perceptron Misclassifications: 6
Did Not Converge After 1000 epochs.
BP Weight vector is: [0.17794236 0.55514478 -1.75907251]
LS Weight vector is: [-0.14601222 1.28051055 -1.32037839]
Least Squares Misclassifications: 8

Multiclass LS:

Multiclass Misclassifications: 34
Multiclass LS Weight vector is: [-0.25132905 0.00983426 1.26603335]

Plots!

