

Aaron Shabanian

Assignment 1

Machine Learning

Abstract

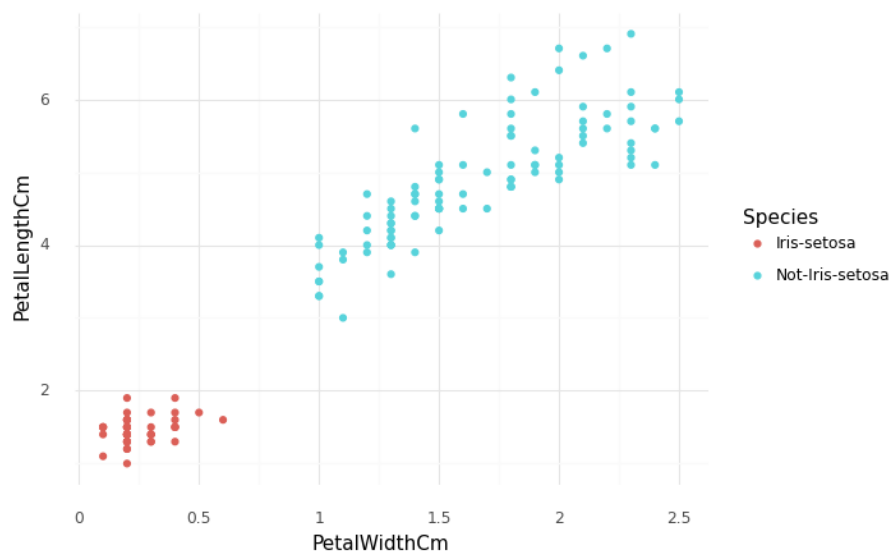
In this assignment, I reviewed a dataset of plants and used support vector machines to classify if the plant was an iris setosa or not.

Background

Support vector machines are a form of supervised learning. SVM's can perform linear classification through the creation of a linear separator and a margin on both sides of the separator. However, the kernel trick can be used to lower the dimensionality of a model when there is not a linear separator that can be used. The kernel trick can get the model to a place where a linear separator with a margin can be used.

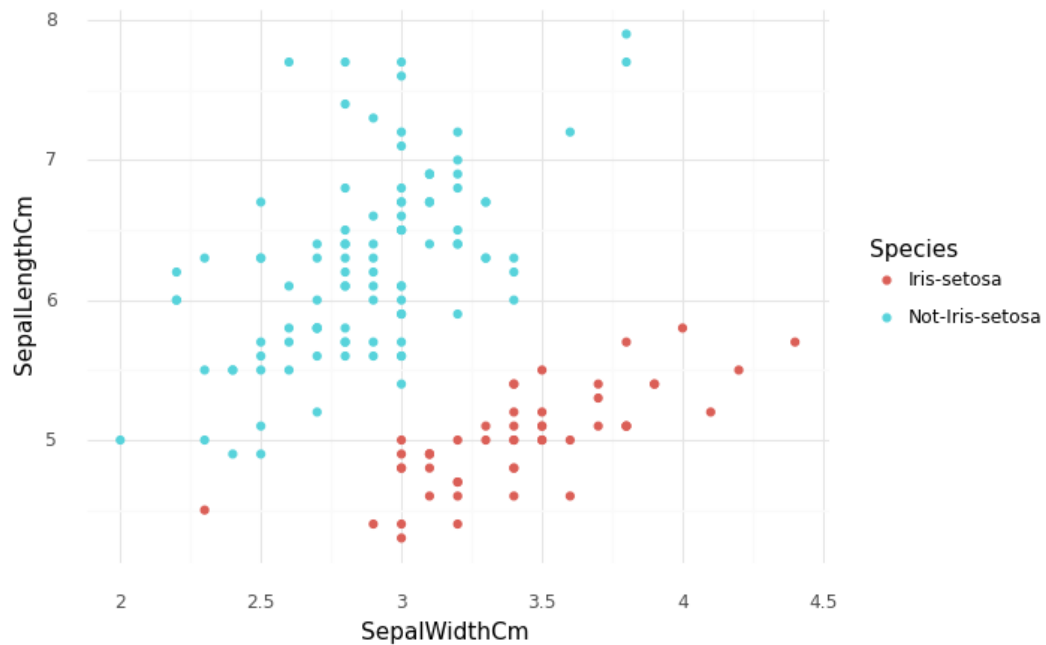
Findings

First, I made a chart comparing the petal length and petal width and whether or not they are Iris setosas.



Looking at this, I can see a big space between the Iris Setosa and the non iris setosas. There is plenty of space for a linear divider with a large margin. There would also be little to no slack in the margin.

Second, I made a chart comparing the sepal length and sepal width and whether or not they are Iris setosas.



Looking at this, there is a much smaller margin between the Iris setosas and non Iris Setosas and a very small amount of slack. Comparing both of these charts, I could estimate that I should get near perfect accuracy since they are linearly separable.

Once I ran the support vector machine, I got the expected accuracy of 100%. This makes sense since it's clearly linearly separable.