Karan Deshmukh

Tableau Documentation:

I have already visualized some of the insights in python. I am carrying out the other interpretations in Tableau.

My idea – Trying to figure out what features are giving the most information by observing their plots. My aim was to figure out – do we need all 4 sepal/petal length/width to predict the species of a flower? Which would be a better indicator for predicting?

Initially I plotted **Sepal Width vs Sepal Length:**

* Here we can observe a mixed pattern and there is no clear interpretation/trend
* Many of Iris Virginica is overlapped with Iris versicolor

A graph with many colored dots

Description automatically generated with medium confidenceA screenshot of a computer

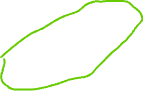
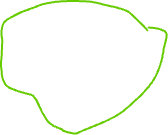
Description automatically generated

I have plotted **Petal Length vs Petal Width**: Here we can observe several insights.

* The data after plotting appears in a cluster fashion.
* Here we can observe clear boundaries and well defined classes
* Interpretations: It can be observed that flowers with **low petal length and width** are clustered in blue and labeled as **Iris Setosa**
* Similarly those with **medium/large petal width and length** are labeled as **versicolor** and **virginica** respectively.

A graph with dots and numbers

Description automatically generated with medium confidence



A screenshot of a computer

Description automatically generated

**Result**: As per above analysis, it can be said that **Petal width and Petal length are most important features when it comes to predicting the flower species**. As observed, plotting Petal Length against petal width is giving us clear plots making it easy to determine species.