LEVEL 1 : Iris Species Prediction

Data Description

Source : Kaggle

There are 150 records with 5 columns each (*sepal length*, *sepal width*, *petal length*, *petal width*, *species*).

sepal length - The length of sepal. (in cm)

sepal width - The width of sepal. (in cm)

petal length – The length of petal. (in cm)

petal width - The width of petal. (in cm)

species - Iris setosa, Iris virginica and Iris versicolor.

Dataset contains 50 observations of each species (setosa, versicolor, virginica).

Problem Statement

To predict the species of the flowers using the four features of sepals and petals.

Accuracy must be used to evaluate the model's predictions.

Features - ***sepal length, sepal width, petal length, petal width.***

Target column – **species**

Notes

The data provided is completely pre-processed.

Hints:

1) For this problem, building a basic model directly without any other ML steps (Data Pre-processing, Feature Engineering, Feature Selection, Cross validation etc.) is sufficient to get a decent accuracy.
2) But feel free to use additional ML techniques and algorithms to achieve better predictions.

References

https://towardsdatascience.com/classification-basics-walk-through-with-the-iris-data-set-d46b0331bf82