Overview of K-Nearest Neighbors Example

Dataset Link:

<https://www.kaggle.com/arshid/iris-flower-dataset>

Dataset Content:

This dataset contains five columns. One column includes three iris species with 50 samples each and the four remaining columns provide some properties (sepal length, sepal width, petal length, petal width) about each flower.

Problem:

We are working as botanologists and he have been provided with three species of flowers but we don’t know how to distinguish them.

Solution:

We want to develop a classification model using K-Nearest Neighbors which allows us to classify with accuracy if each one of the flowers is an iris setosa, iris versicolor or iris virginica based on the combination of some features we have access to.

Variables:

Independent variable 🡪 Sepal width, Petal width.

Dependent variable 🡪 Iris setosa, Iris Virginica, Iris versicolor.