Overview of K-Nearest Neighbors Example

Dataset Link:

<https://www.kaggle.com/jmcaro/wheat-seedsuci>

Dataset Content:

This dataset contains eight columns. The first seven columns contain measurements of geometrical properties (area, perimeter, compactness, length of kernel, width of kernel, asymmetry coefficient, length of kernel) of kernels belonging to three different varieties of wheat. The three classes of wheat are presented on the last column.

Problem:

We would like to classify in which wheat variety each kernel belongs based on the features we have available.

Solution:

We want to develop a classification model using K-Nearest Neighbors which allows us to classify with accuracy if the kernel belongs to wheat variety 1 , wheat variety 2 or wheat variety 3.

Variables:

Independent variable 🡪 Area, perimeter, compactness, length of kernel, width of kernel, asymmetry coefficient, length of kernel.

Dependent variable 🡪 Wheat Variety 1, Wheat Variety 2, Wheat Variety 3.