Data scientist project

Question: Can we predict Iris flowers characteristics by ML?

RIS Dataset:

The iris flower data set Fischer data set is a multivariate data set used by British biologist Ronald Fisher. Using several measurements in problems, for example, linear linear analysis.

The data set contains 50 samples from each of the three species Iris, as out put (Iris setosa, Iris virginica and Iris versicolor). Four characteristics of each sample were measured (as input): length and width of the spoon and petals, in cm. Based on a combination of these four features, Fisher created a linear differential model for species identification.

Abstract:

In this project, we loaded a data set and understand the structure by using different graphs, statistical methods, data visualization and look at machine learnings' elements in Spyder.

Introduction:

The method used are as follow:

1- Class prediction
2- Single regression prediction
3- Class prediction
4- KNeighbors Classifier
5- Splitting the data set
6- Data visualization

Discussion:

Bu using all 4 techniques, features of each flower is predictable.

Result

Result of linear regression analysis:

X=[-1.07296862 -0.52817175], Predicted=-62.218618956076064

X=[-0.61175641 1.62434536], Predicted=-13.418251657176773

X=[-2.3015387   0.86540763], Predicted=-110.5797962861611

References:

https://github.com/cce-bigdataintro-1160/cebd1160_project_template/tree/gkexample

https://stats.stackexchange.com/questions/333485/linear-model-with-iris-dataset

https://towardsdatascience.com/knn-using-scikit-learn-c6bed765be75

https://cce-bigdataintro-1160.github.io/CEBD-1160-spring-2019/8-ml.html

https://rpubs.com/Hgoswami/368890