2023

Assignment 2 - Classic Machine Learning

Bing Li

CEE 9730/4420 Assignment 2

1) We will continue with the dataset contained in "CLAY_10_7490_TC304.csv". Build a classifier for the soil type (CH, CL, MH, ML) based on any parameters **except** Atterberg limits LL, PL, PI, LI. Pick **at least 2** different kinds of machine learning algorithms for this classifier. Evaluate the performance **on the test set** of each quantitatively (Precision, recall, accuracy) and qualitatively (by plotting it). (4 points)
2) Based on any of the classifiers above, which parameters best predict the soil type? (1 point)
3) Pick **at least 2** regression algorithms to fill the missing values of OCR, Remolded su, and su(mob) using any available data. Evaluate the performance **on the test set** quantitatively (MAE, RMSE, R2) and qualitatively (by plotting it). (8 points)
4) Based on any of the regression models above, which parameters best predict the OCR? (2 points)

Note 1: Figures are expected to be plotted in python with all axes labelled and legend included where multiple data series are presented.

Note 2: Some parameters may not be correlated at all, and you may not achieve good performance scores. This is ok, just document what was done and why.