Assignment #1 - Supervised Learning

Firas Bayram
firas.bayram@kau.se

Karlstad University — August 26, 2021

Introduction

Supervised machine learning algorithms are those algorithms where the corresponding output of each training data point is available. The learning task is to find the function that models the relationship between the inputs and outputs. This assignment is about getting familiar with supervised machine learning algorithms and understanding how it works in practice. The goal of the assignment is to give hands-on exposure on the machine learning algorithms we have discussed in class: Linear models and KNN. In this assignment, we will use the Python programming language and Jupyter Notebook to develop the solution.

Part #1 – Classification

1. Generate and plot a 2D training dataset with two classes of 500 points. Similarly, generate a testing dataset of 50 points.
2. Fit Linear SVM and KNN classifiers on the training dataset. Select the best parameters using cross validation method.
3. Apply Linear SVM and KNN algorithms on the testing dataset. Compare the results.
4. Add some noise (10%) to the generated training dataset by flipping the labels. Evaluate the classifiers' performance trained on the noisy dataset.

Part #2 – Regression

1. Generate a regression problem of 500 training data points. Similarly, generate a testing dataset of 30 points.
2. Fit KNN and Linear regressors on the training dataset. Select the best parameters using cross validation method.
3. Apply KNN and Linear regression algorithms on the testing dataset. Compare the results.

Examination

Hand in your solution in Jupyter Notebook on Canvas and book a meeting to demonstrate your solution with one of the lab assistants. Note that lab demonstration meetings are booked through email, not on Slack.

End of Assignment