Elements of Machine Learning
Assigment 2 - Problem 3

Sangeet Sagar(7009050), Philipp Schuhmacher(7010127)

{sasa00001,phsc00003}@stud.uni-saarland.de

December 2, 2021

Problem 3 (T, 2 Points). Cross Validation

The main idea of k fold cross validation is to randomly split the data set into k folds. We start by splitting the data randomly into k parts and using the kth fold as testing data and the rest as training data. We go through this step k times and use every fold as testing data once. At the end, the data gets collected and mean of MSE is calculated (by summing up all individual MSE's and averaging them over k folds). With this method the best estimate of the test error is computed, because every k estimation on its own can have few quality at the end. We get k estimates for the test error and have the best estimation at the end.