Fall 2020 Homework 2-Learning

CS465/565 Homework

Due: October 18, 2020 @ 11:59

**Part 1-Supervised learning**

There exists a nonlinear relationship between input attributes x and y and output target z. The training set consists of 30 noisy samples. You can assume the range of x is -1:3 and the range of y is -3:1.

Choose a parametric supervised training method that predicts the values in z given the attributes in x and y. Train your learner using python skikit-learn or matlab. Be careful not to overfit.

*x.csv*- input attribute 1

*y.csv*- input attribute 2

*z.csv* – target value

Submit a zip file with the following items:

* A python or MATLAB script that reads in attributes from x.csv and y.csv and outputs the predicted values to z-predicted.csv.
* README.md in the root directory that provides running instructions, describes the parameters and architecture of your learner, justification of why your learner is a good choice over other learners and a discussion of how you traded off model complexity with empirical loss over the training set.

**Part 2-Unsupervised learning**

*F20-465-hmwk2-p2* contains a two attribute dataset. Use an unsupervised learning method to identify the number and parameters of clusters of similar data. You are to provide the mean and covariance of clusters of similar data or divide the data points into similar sets. In addition, provide the details of your unsupervised learning method discussing the choice of method and number of clusters. You can provide empirical information on your search for the number of clusters and the parameters.