

ECE 9603/9063b – Data Analytics Foundations

Assignment 1: Forecasting

Deadline:

Wednesday, February 14th, 2018

In this assignment you will experiment with different forecasting approaches and algorithms. As this is your first assignment in this course, you can keep it simple.

1. Select a forecasting problem for which there is an open source dataset or for which you have data from other sources. Links for open data sets are provided on the course site in OWL (in Wiki). Some possible forecasting tasks include Energy Forecasting (<https://www.kaggle.com/c/global-energy-forecasting-competition-2012-load-forecasting/data>) and house prices prediction (<https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data>).
2. Carry out forecasting with at least 3 different approaches/algorithms. The approaches need to be significantly different. For example, support vector regression with two different kernels will be considered one approach. Some possibilities include neural networks, support vector machine/regression, multivariate regression, k-nearest neighbour, regression trees, ARMA, ARIMA, Markov chain and any other.

Note: The objective of this assignment is to experiment with different models. The focus is on applying forecasting approaches and not on optimizing models.

3. Evaluate the three models/algorithms using hold-out or cross-validation and compare their accuracy

Deliverables:

- Report submitted in OWL (total 20 points):
 - Description of the selected forecasting problem (4 points)
 - Description of available data (attributes, context, quantity...). Clearly indicate what attributes and/or parts you have used (4 points)
 - Short overview of the selected algorithms (4 points)
 - Specifics about how algorithms were applied and the evaluation procedure (4 points)
 - Accuracy comparison (4 points)
 - Code. Although there are no marks for the code itself, marks will be deducted if the code does not match the rest of the report.