Cooper Sloan

September 29, 2016

Evaluation Metrics

My goal is to do accurate prediction on bus arrival times, so the key metric I would like to measure is error from the predicted to the actual values. Specifically I will use mean square error which is appropriate in this case. Because this is a machine learning problem, I will measure the training as well as validation error, obviously with the emphasis being on validation error. The error will be a function of several factors, one of which is the model. I want to standardize a training and validation set so that I can compare models. Additionally I would like to measure how different sizes of training data affect the validation accuracy. While there are many other factors I could measure I want to limit my scope to these two factors for now.

In order to measure the error, I will need to frame my models in a way which generate predictions for bus arrival times in a structured way. In order for the error comparisons to be valid I am going to identify a subset of the MBTA data for training and another subset for validation. I want to make these sets rather large in order to avoid overfitting. As long as my models can generate predictions for the input set in a consistent way, it will be simple to measure the error for each model and compare them. Some of my models may not require the training data, and in that case I will compare their error to the validation error of the other models.