Applying machine learning

Prediction of restaurant inspections in NYC

0.1 Framing the problem as a machine learning problem

Our main goal is to predict the time until the next inspection. In a first attempt, I will try to predict the number of days until the next inspection. Furthermore, I would like to use predictors constructed from the variables in the data set. Hence, the main question is framed as a supervised learning problem, in particular, a regression problem

0.2 Main features

The main features will be score, grade, critical flag, inspection type, and number of violations in each violation group.

0.3 Techniques

I am planning to use linear regression and regression trees (CART)

0.4 Evaluation metrics

I will use R squared or RMSE to evaluate the success of the model.