The training data was originally split into train and validation data. 70% of the data were randomly sampled into the training set and the remaining 30% were sampled into the validation set with seed = 123.

For the first task, we were asked to create a model predicting whether something would be a static or dynamic activity. For the baseline model, I used a naïve Bayes prior probability approach of just predicting the most commonly seen outcome in the training data which was “static”. So, by using this strategy and guessing “static” for all of the data in the validation set, regardless of covariates, I had an accuracy of 0.584.