**Optimizing the hyperparameter**

To minimize the mean squared errors of our Lasso and Ridge regression we performed k-fold cross validation to optimize the hyperparameter .

We split the data into a test set and a development set, consisting of respectively 20% and 80% of the total observations. Subsequently, we use k-fold cross-validation to randomly split the development set into *k* folds, where folds are used to train the model. The remaining fold is used to validate the model’s generalizability by calculating the mean squared errors of the trained model’s prediction of the left-out fold (Raschka 2017:191). This process is repeated k times and each time a new fold is left out for validation. Since we are working with a relatively large dataset we chose to split our data into 5 folds, and computed the average MSE for the 5 iterations. By using the k-fold cross-validation method we relieve ourselves of the concern that the estimation of our model’s performance is simply due to a lucky or unlucky split of the data.

We performed this procedure for 12 different values of spanning between and . We chose the value of which yielded the smallest average MSE over the 5 folds.

We both calculated the optimal hyperparameters for a Ridge regression model, a Lasso regression model and an Elastic Net regression model. The following table shows the performance of the different models, when trained with their optimal hyperparameter and predicting the test data.