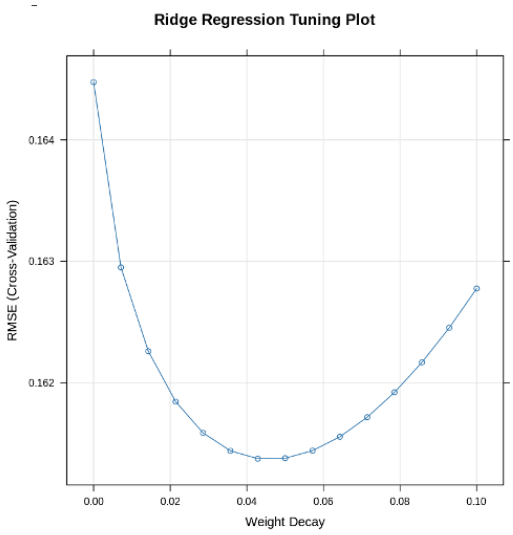
Appendix 1: Supplemental Material for Tuning plots and parameters for the models

Ordinary Least Squares (OLS)

The Ordinary Least Squares model does not require tuning because it is a simple linear regression method that directly fits a linear relationship between the predictors and the target variable. Unlike models such as Ridge, Lasso, or Elastic Net, which involve regularization parameters that need to be optimized to control overfitting, OLS doesn't have any hyperparameters to adjust.

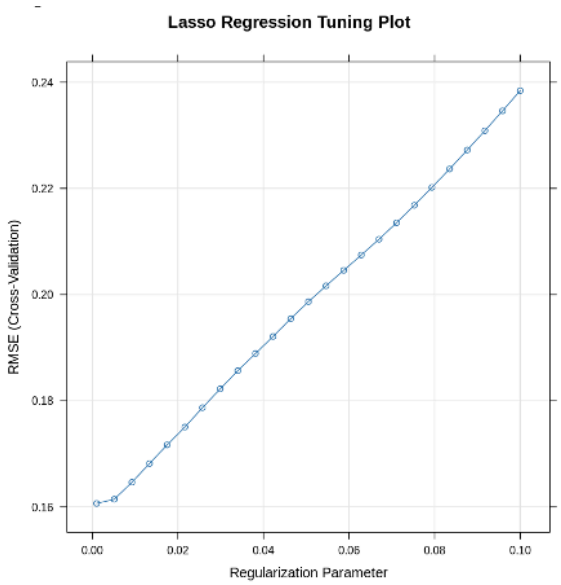
Ridge Regression Model

RMSE was used to select the optimal model using the smallest value. The final value used for the model was lambda = 0.05. RMSE: 0.1616538 R-squared: 0.8404443



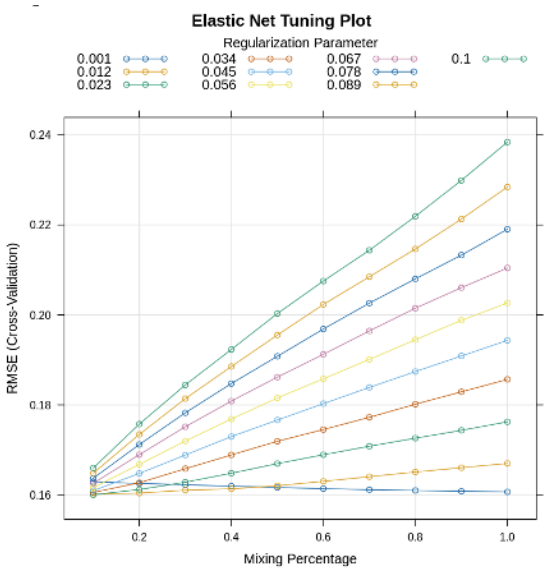
Lasso Regression Model

Tuning parameter 'alpha' was held constant at a value of 1 RMSE was used to select the optimal model using the smallest value. The final values used for the model were alpha = 1 and lambda = 0.001. RMSE: 0.1607747 R-squared: 0.8402121



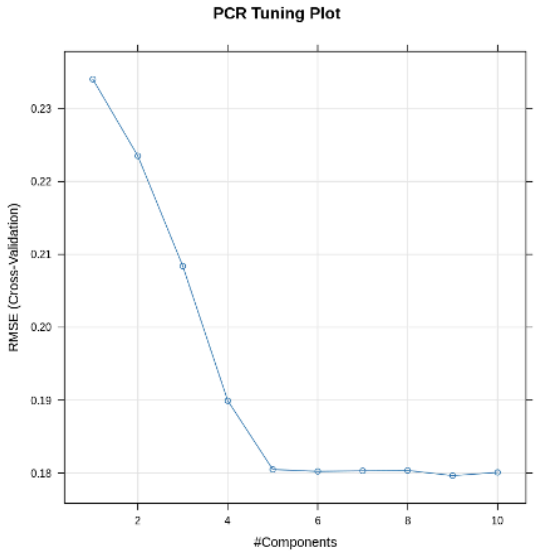
Elastic Net Regression Model (E-NET)

RMSE was used to select the optimal model using the smallest value. The final values used for the model were alpha = 0.1 and lambda = 0.023. RMSE: 0.1600687 R-squared: 0.8424046



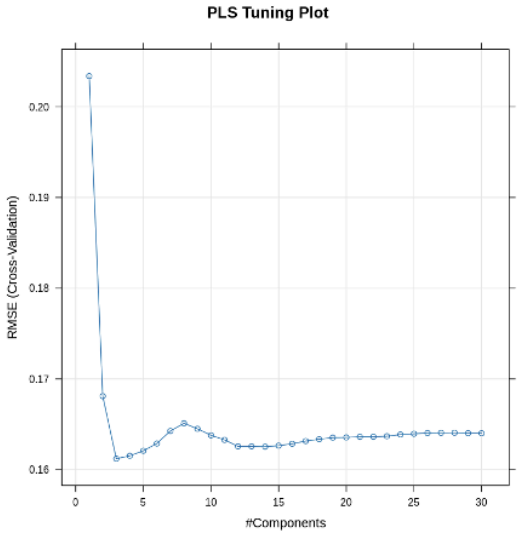
Principal Component Regression Model (PCR)

RMSE was used to select the optimal model using the smallest value. The final value used for the model was ncomp = 9. RMSE: 0.1793181 R-squared: 0.8007603



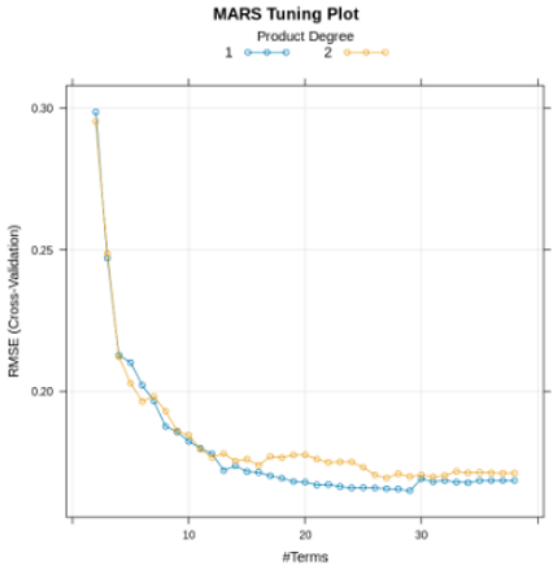
Partial Least Squares Regression Model (PLS)

RMSE was used to select the optimal model using the smallest value. The final value used for the model was ncomp = 3. RMSE: 0.1613994 R-squared: 0.8384435



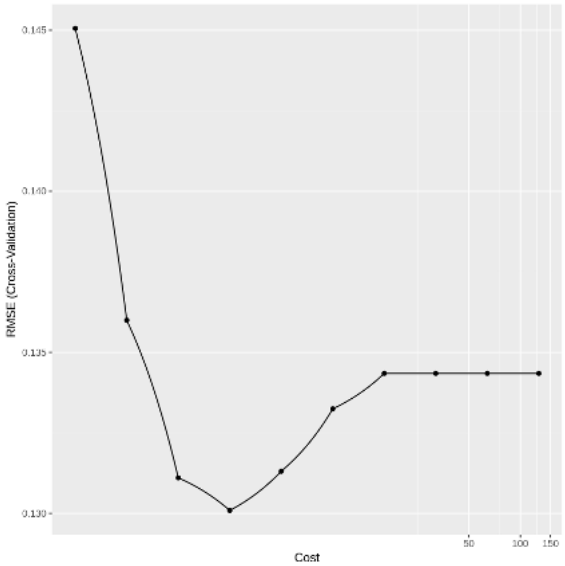
Multivariate Adaptive Regression Model (MARS)

RMSE was used to select the optimal model using the smallest value. The final values used for the model were nprune = 29 and degree = 1. Model Performance: RMSE: 0.1649153 R-squared: 0.8321994



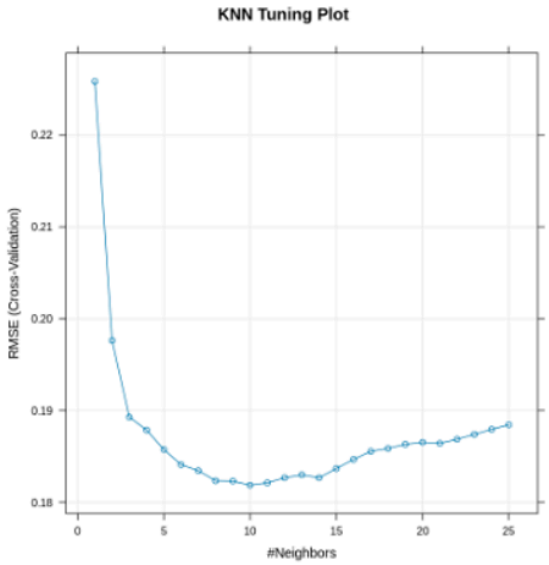
Support Vector Machines Model (SVM)

Tuning parameter 'sigma' was held constant at a value of 0.004969824 RMSE was used to select the optimal model using the smallest value. The final values used for the model were sigma = 0.004969824 and C = 2. RMSE: 0.1307137 R-squared: 0.8903392



K-Nearest Neighbors Model (KNN)

RMSE was used to select the optimal model using the smallest value. The final value used for the model was k = 10. Model Performance: RMSE: 0.1818747 R-squared: 0.7971668



Neural Network Model (NN)

RMSE was used to select the optimal model using the smallest value. The final values used for the model were size = 1 and decay = 0.5. Model Performance: RMSE: 0.1460579 R-squared: 0.8632528