Data Analysis Report

# Raw data summary

The underlying data has 506 observations with 13 variables. The raw data was split into training and testing data using 0.9 partition, which resulted in a training data size of 455 and a testing data size of 51. Model Trees was used as the for data analysis with the training dataset.

# Data analysis summary

The goodness-of-fit measured in adj R-squared is 0.8298259305995466. Based on the best model selected, the prediction error on the testing data is 3.425483198229281. The residual plot is attached in the following section

# The performance measure table

|  |  |  |
| --- | --- | --- |
| Performance measure index | Performance measure name | Value |
| 1 | goodness-of-fit | 0.8298259305995466 |
| 2 | training rmse | 2.9283799847792173 |
| 3 | testing rmse | 3.425483198229281 |

# The residual plot

