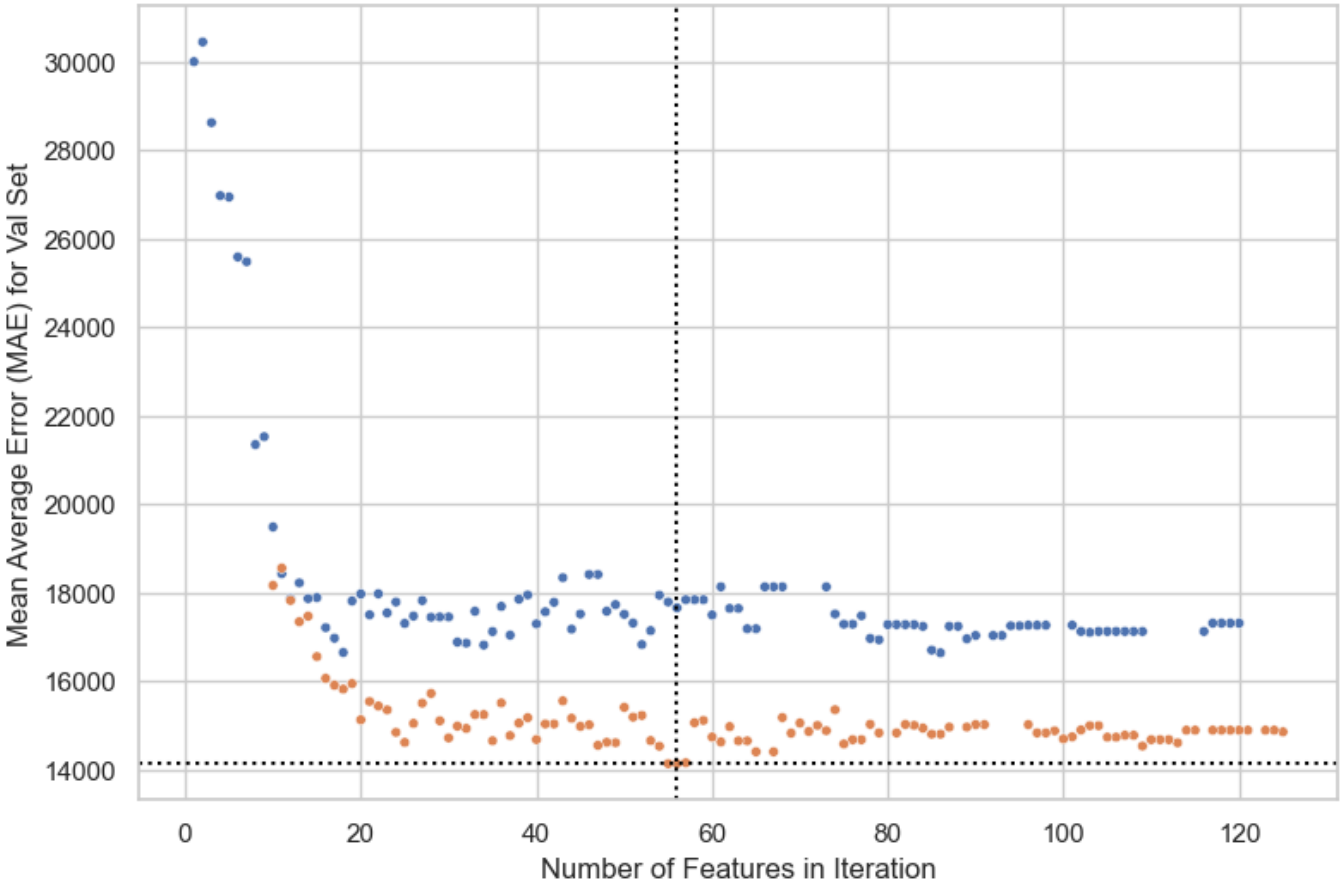
Feature Selection with Recursive Model Training

Recursive Training Results



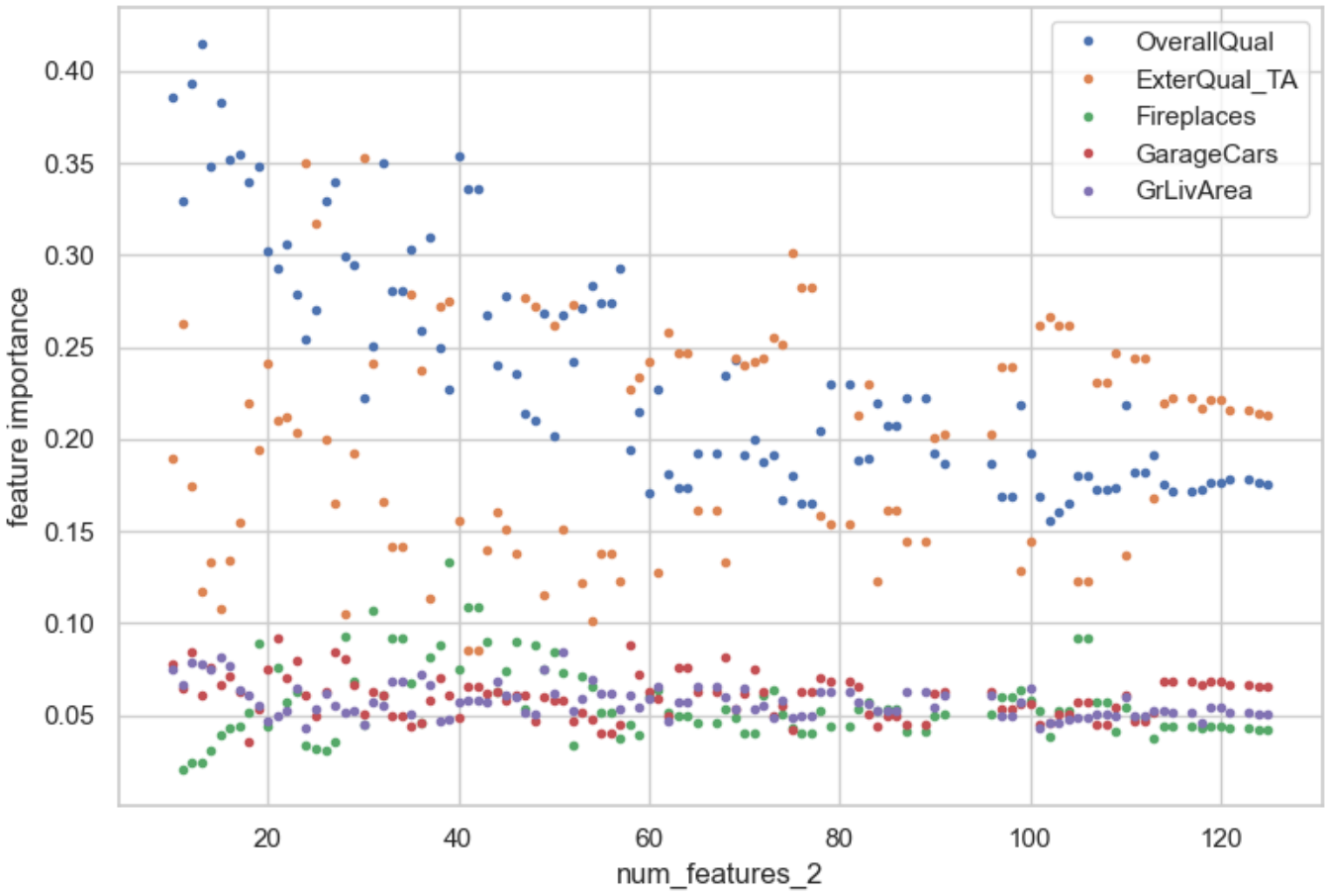
Note: The point from the first iteration with 287 features and an MAE of 14853 was removed from this plot.

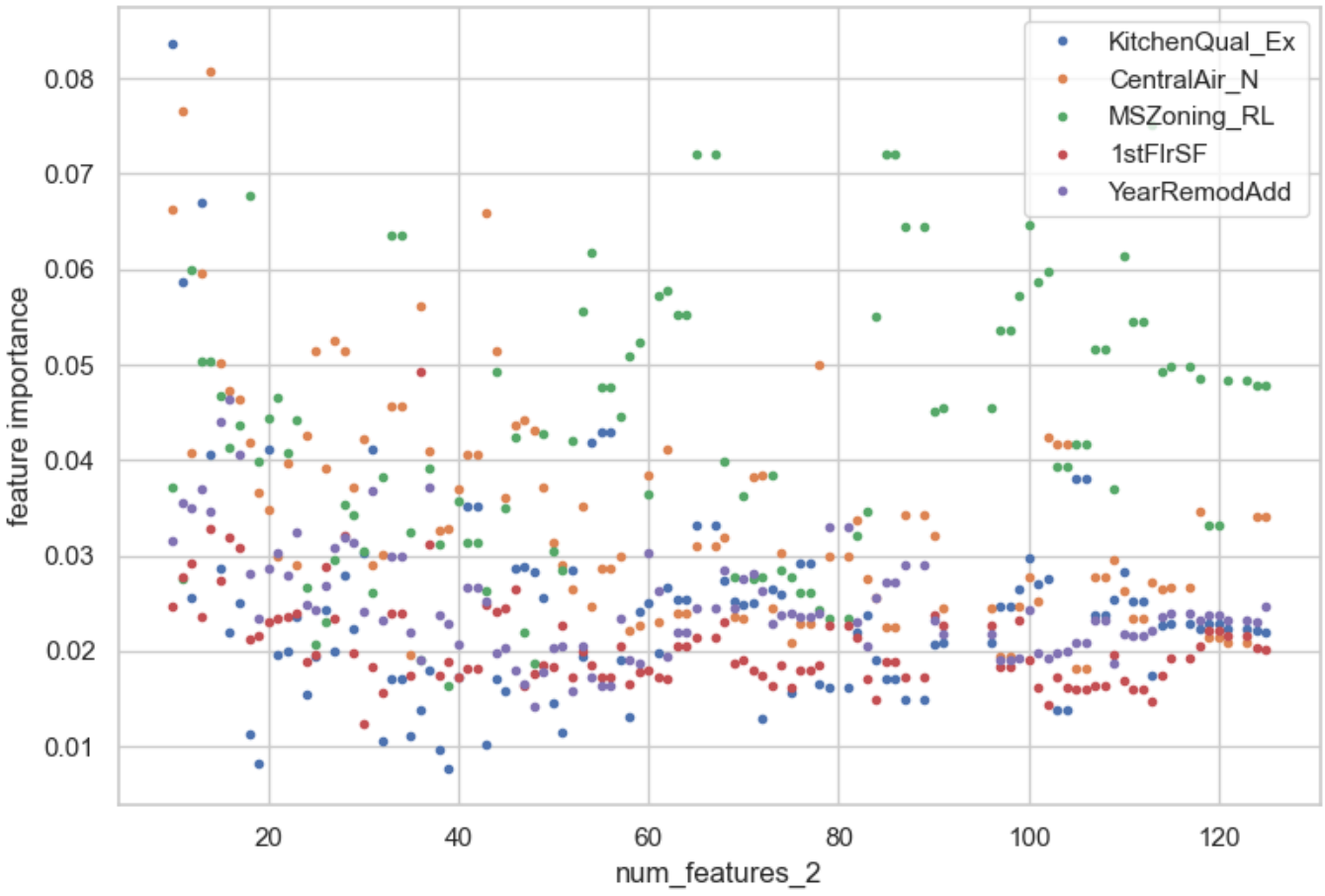
Normally, the way this recursive model training works is that it removes the feature with the lowest importance at each iteration. However, if there are multiple features that have exactly zero importance, then all of those zero importance features are removed at once.

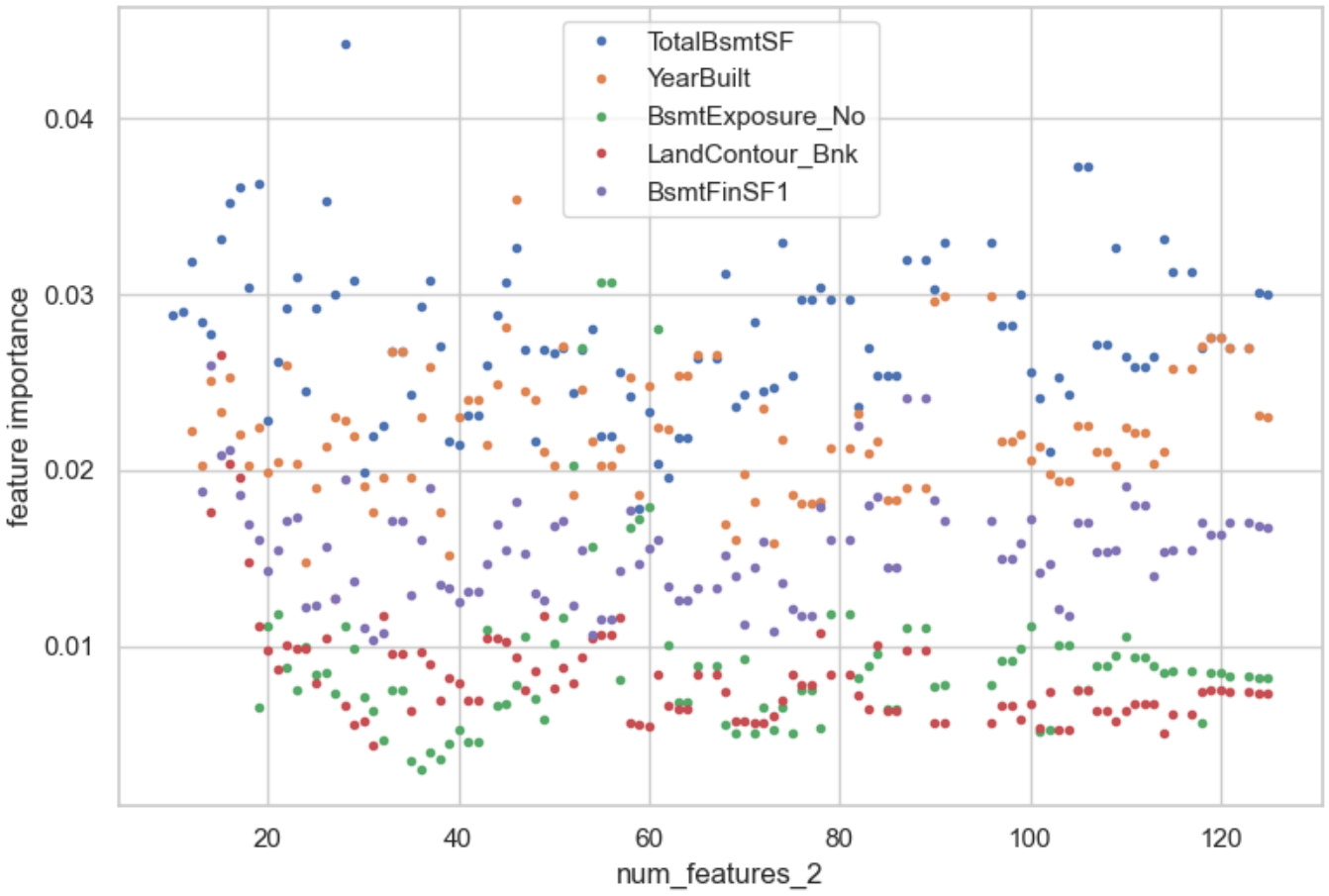
The above plot has our model metric on the y-axis, and the number of features for each model training iteration on the x-axis. In other words, each dot here represents an iteration of the recursive model training.

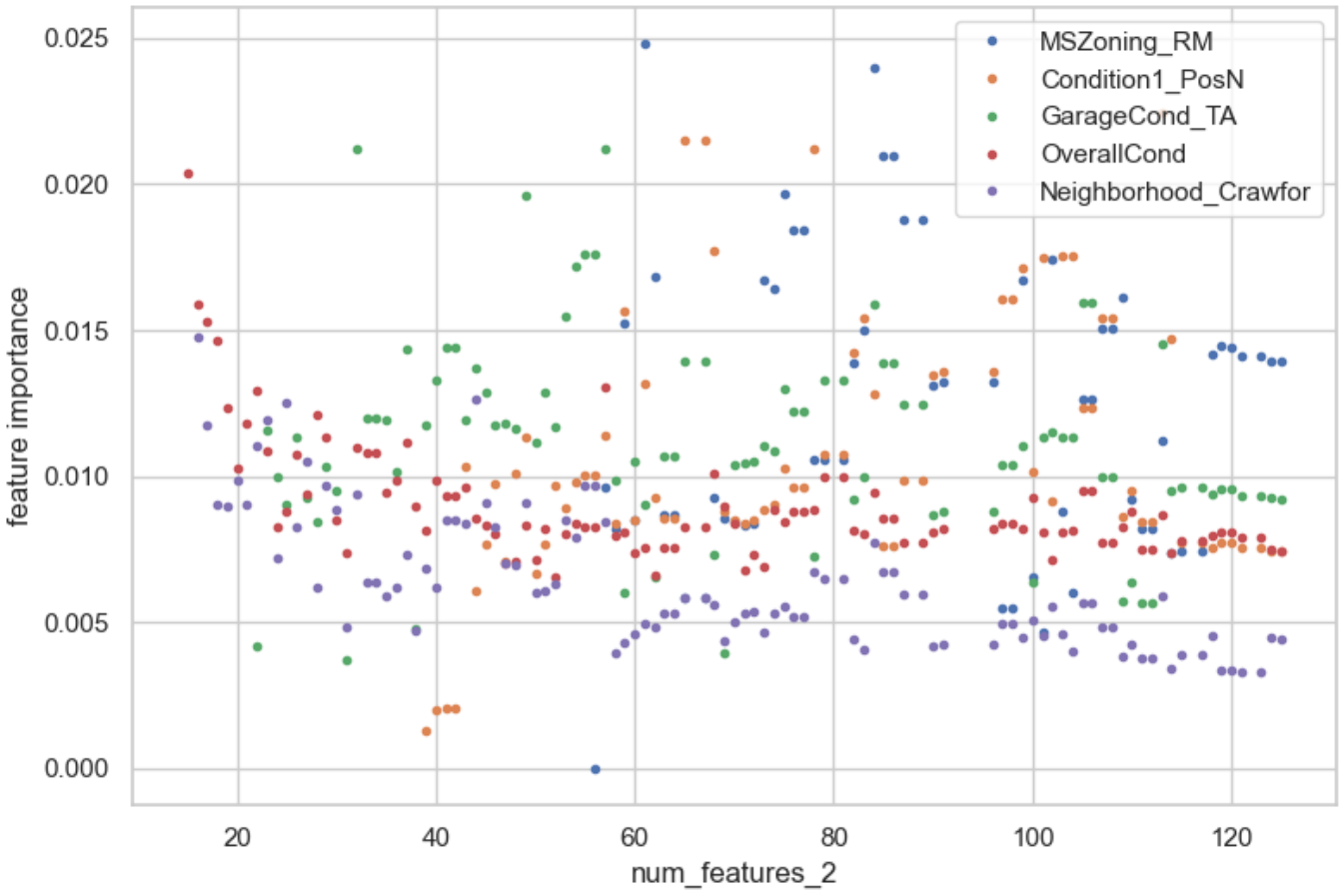
As the number of features is reduced, eventually the model will start to perform much more poorly. The vertical line is the location with the best value of the evaluation metric, which is an MAE of 14134), compared to the starting MAE of 14853.

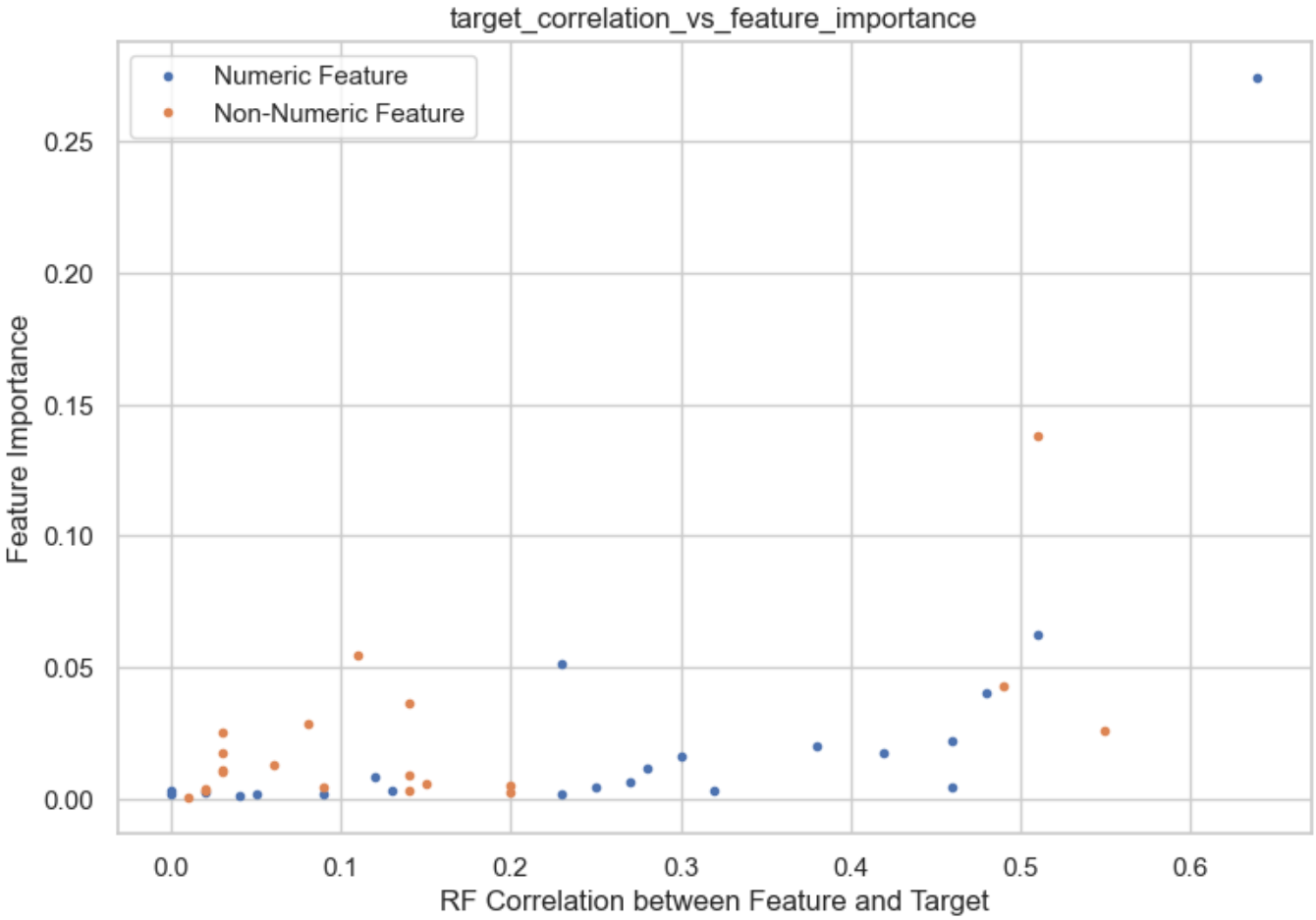
The model training started with 287 features (after one-hot encoding any categorical features), and achieved the best model training results with 56 features.


OverallQual
ExterQual_TA
Fireplaces
GarageCars
GrLivArea
feature importance
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
20
40
60
80
100
120
num_features_2


KitchenQual_Ex
CentralAir_N
MSZoning_RL
1stFlrSF
YearRemodAdd
feature importance
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
20
40
60
80
100
120
num_features_2


TotalBsmtSF
YearBuilt
BsmtExposure_No
LandContour_Bnk
BsmtFinSF1
feature importance
0.04
0.03
0.02
0.01
20
40
60
80
100
120
num_features_2


MSZoning_RM
Condition1_PosN
GarageCond_TA
OverallCond
Neighborhood_Crawfor
feature importance
0.025
0.020
0.015
0.010
0.005
0.000
20
40
60
80
100
120
num_features_2


target_correlation_vs_feature_importance
Numeric Feature
Non-Numeric Feature
Feature Importance
0.25
0.20
0.15
0.10
0.05
0.00
0.0
0.1
0.2
0.3
0.4
0.5
0.6
RF Correlation between Feature and Target