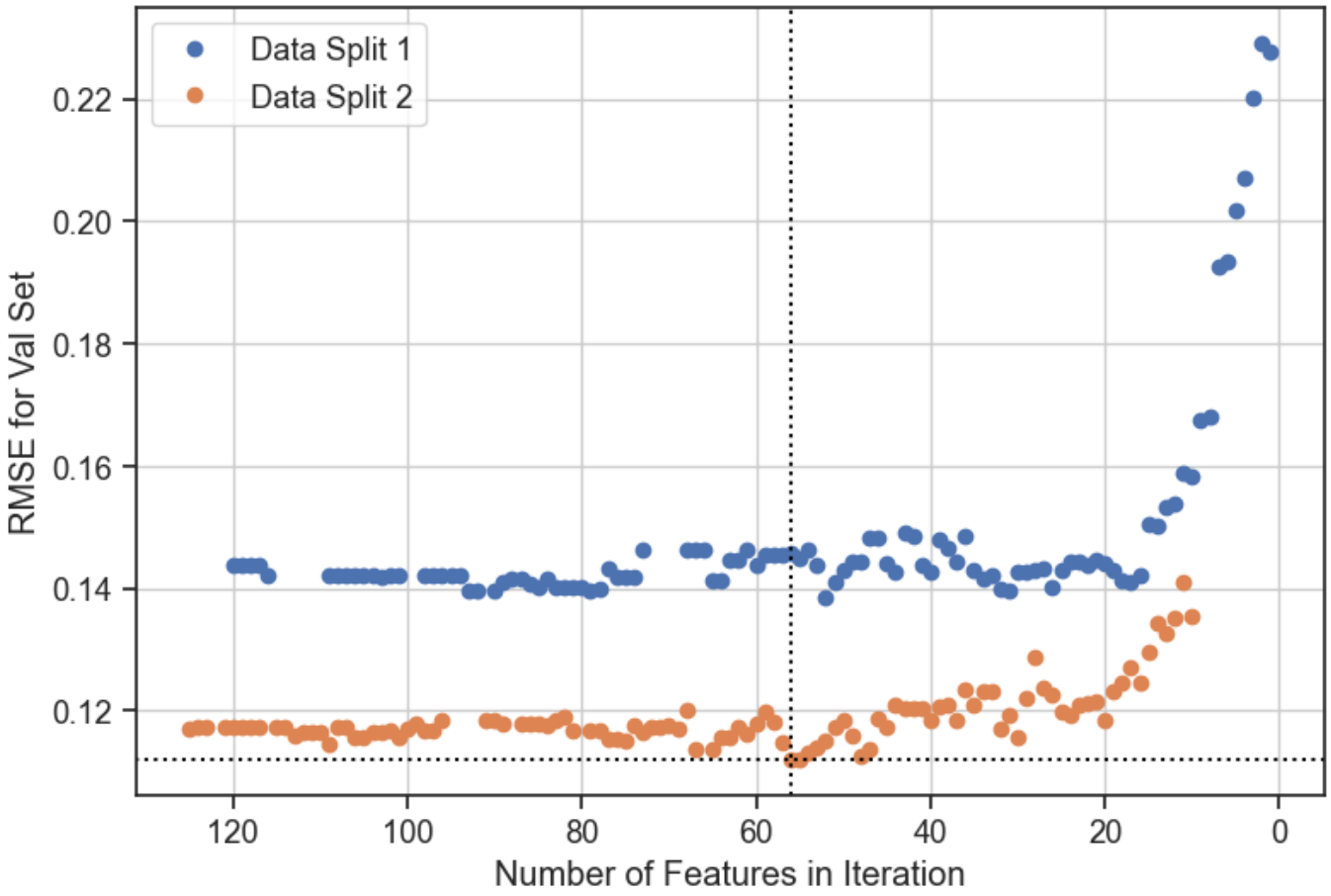
Feature Selection with Recursive Model Training

Recursive Training Results



Note: The point from the first iteration with 287 features and a RMSE of 0.11719 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.

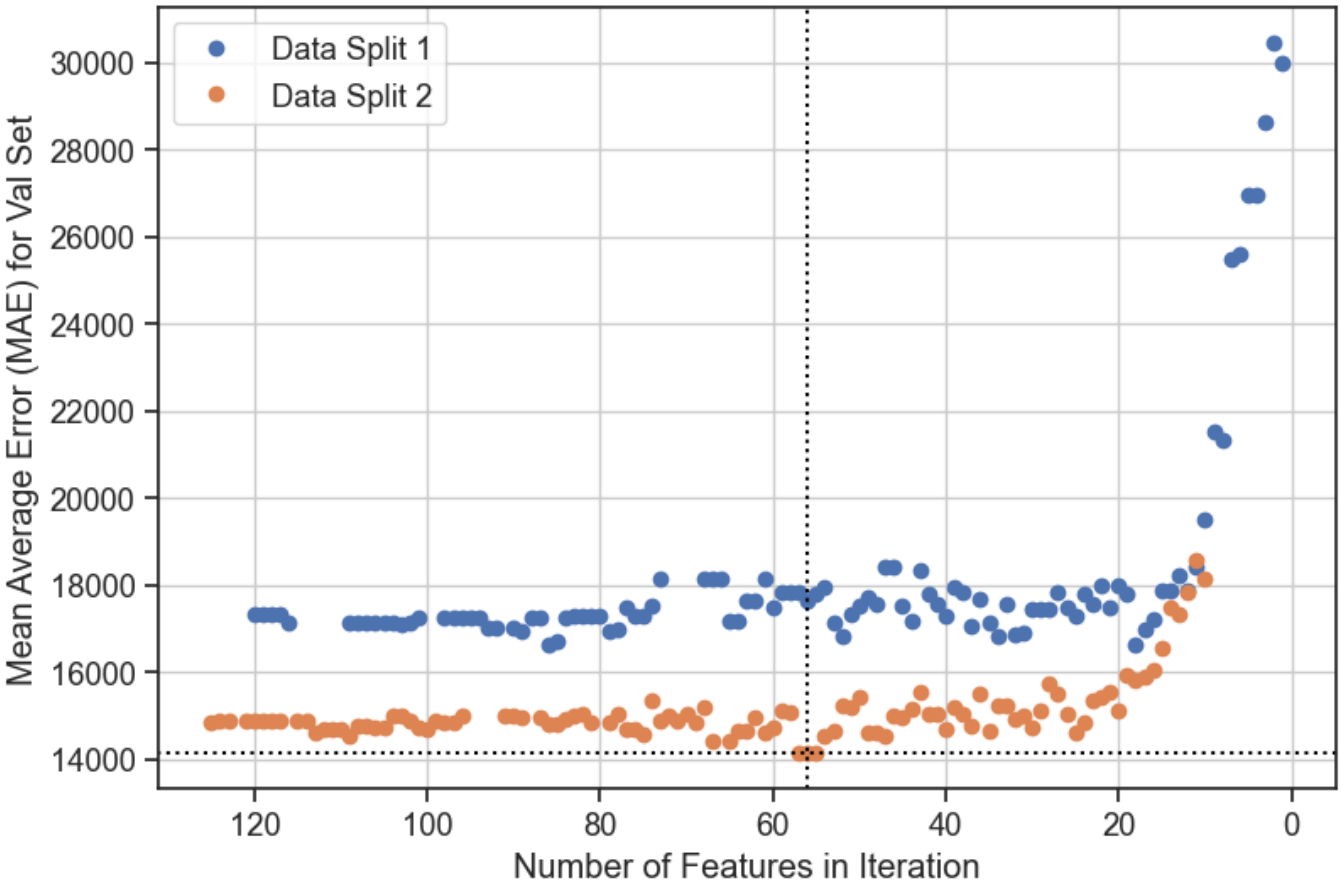
The above plot has our model metric on the y-axis, and the number of features for each model training iteration on the x-axis. In other words, each dot here represents an iteration of the recursive model training.

As the number of features is reduced, eventually the model will start to perform much more poorly.

This plot shows the primary metric that was used in model training, which is RMSE. The vertical line is the location with the best value of this metric, which is a RMSE of **0.11197**, compared to the starting RMSE of **0.11719**.

Note that lower values of RMSE indicate better model performance.

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



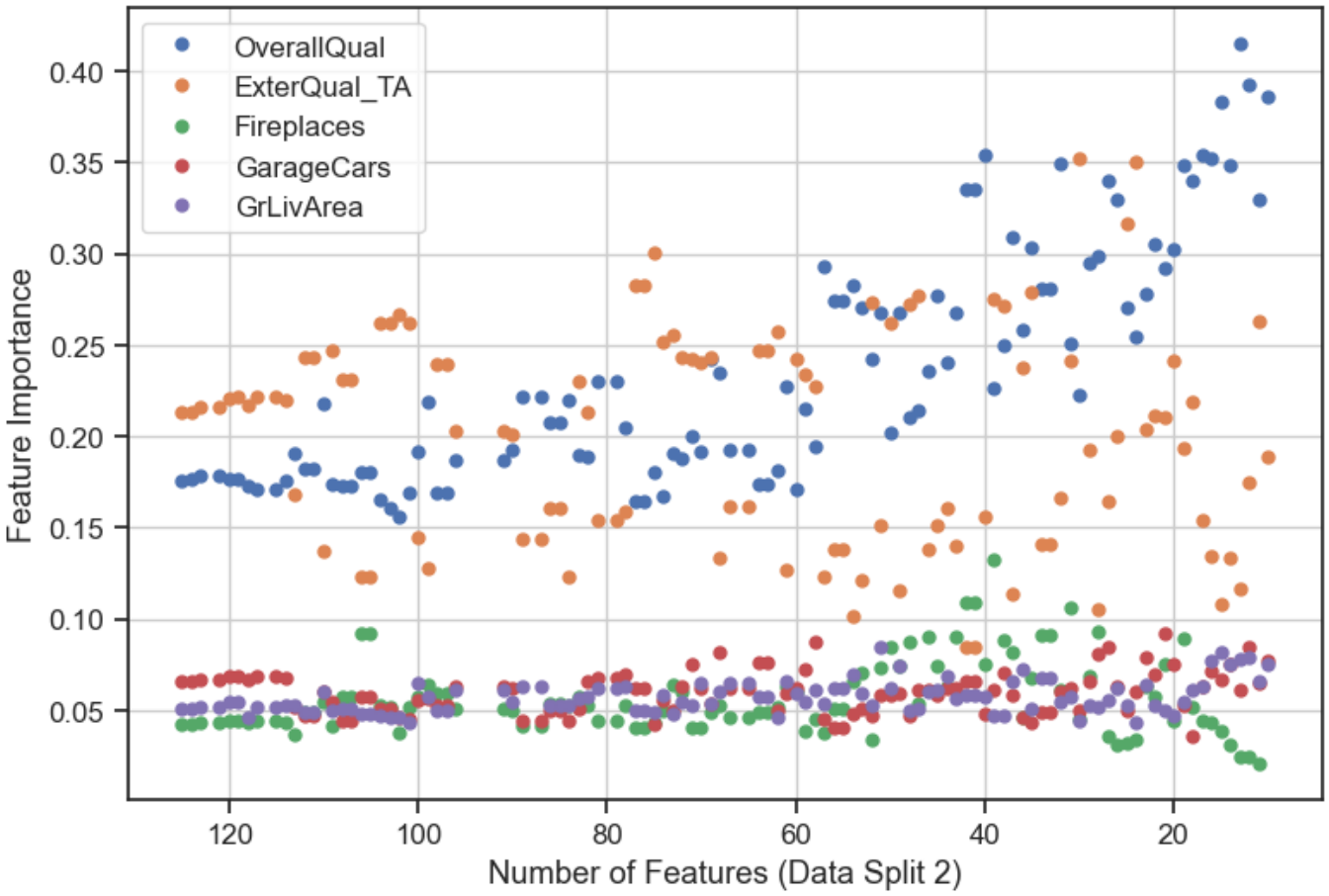
This plot is the same as the previous page, except with the secondary metric, MAE.

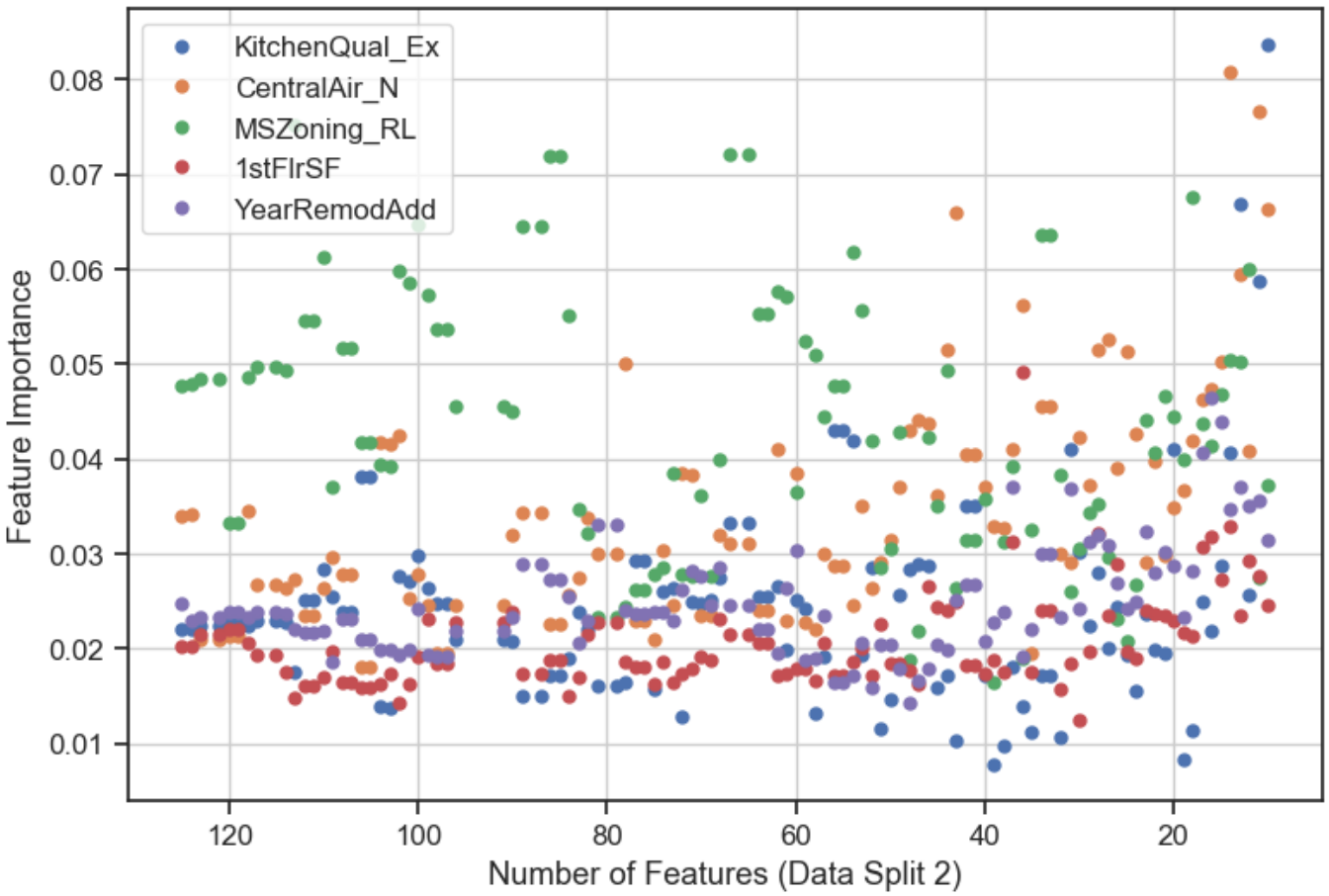
It is therefore possible that the 'best' iteration will be different in this plot.

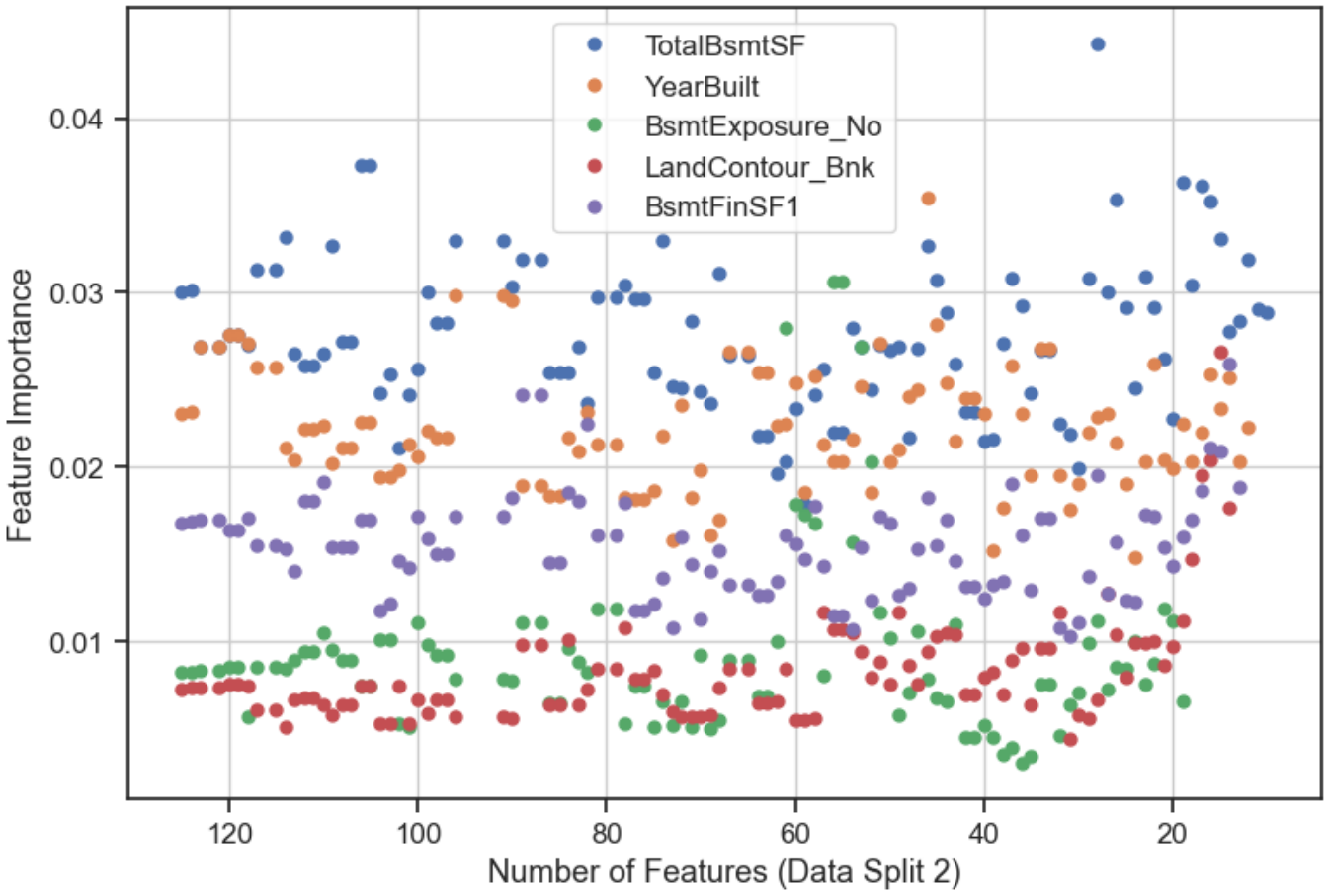
The best value of MAE is **14,134**, compared to the starting value of **14,853** (the average SalePrice is 180,920).

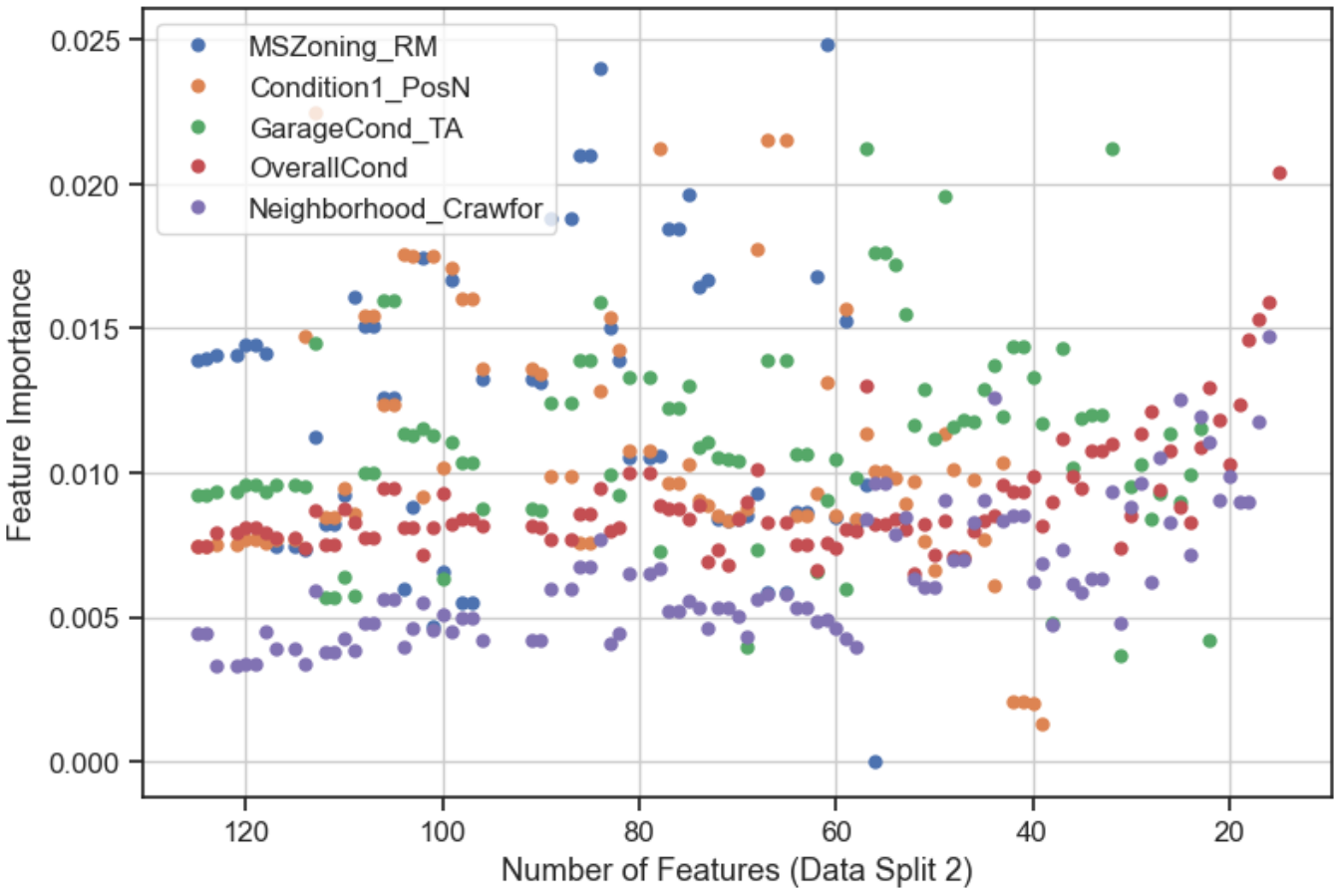
With this metric, the best model training result occurred with **56 features**.

Exploring Feature Importance during Iterative Training






TotalBsmtSF
YearBuilt
BsmtExposure_No
LandContour_Bnk
BsmtFinSF1
Feature Importance
0.04
0.03
0.02
0.01
120
100
80
60
40
20
Number of Features (Data Split 2)


MSZoning_RM
Condition1_PosN
GarageCond_TA
OverallCond
Neighborhood_Crawfor
Feature Importance
0.025
0.020
0.015
0.010
0.005
0.000
120
100
80
60
40
20
Number of Features (Data Split 2)

Exploring Feature Importance Compared to Individual Feature Correlations