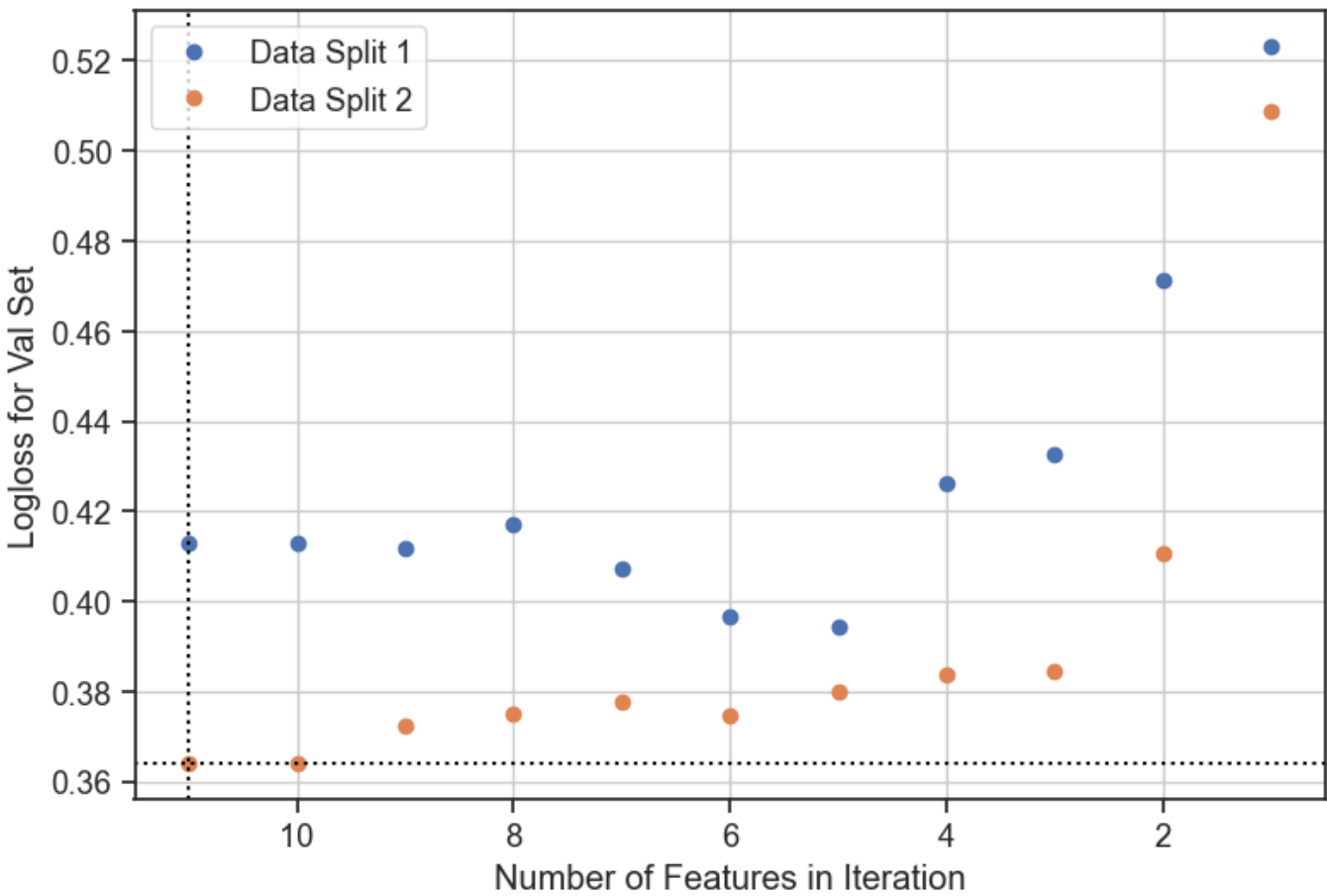# Feature Selection with Recursive Model Training

Recursive Training Results

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 logloss. The vertical line is the location with the best value of this metric, which is a logloss of **0.36394**, compared to the starting logloss of **0.36394**.

Note that lower values of logloss indicate better model performance.

The model training started with 11 features (after one-hot encoding any categorical features), and achieved the best model training results with **11 features**.
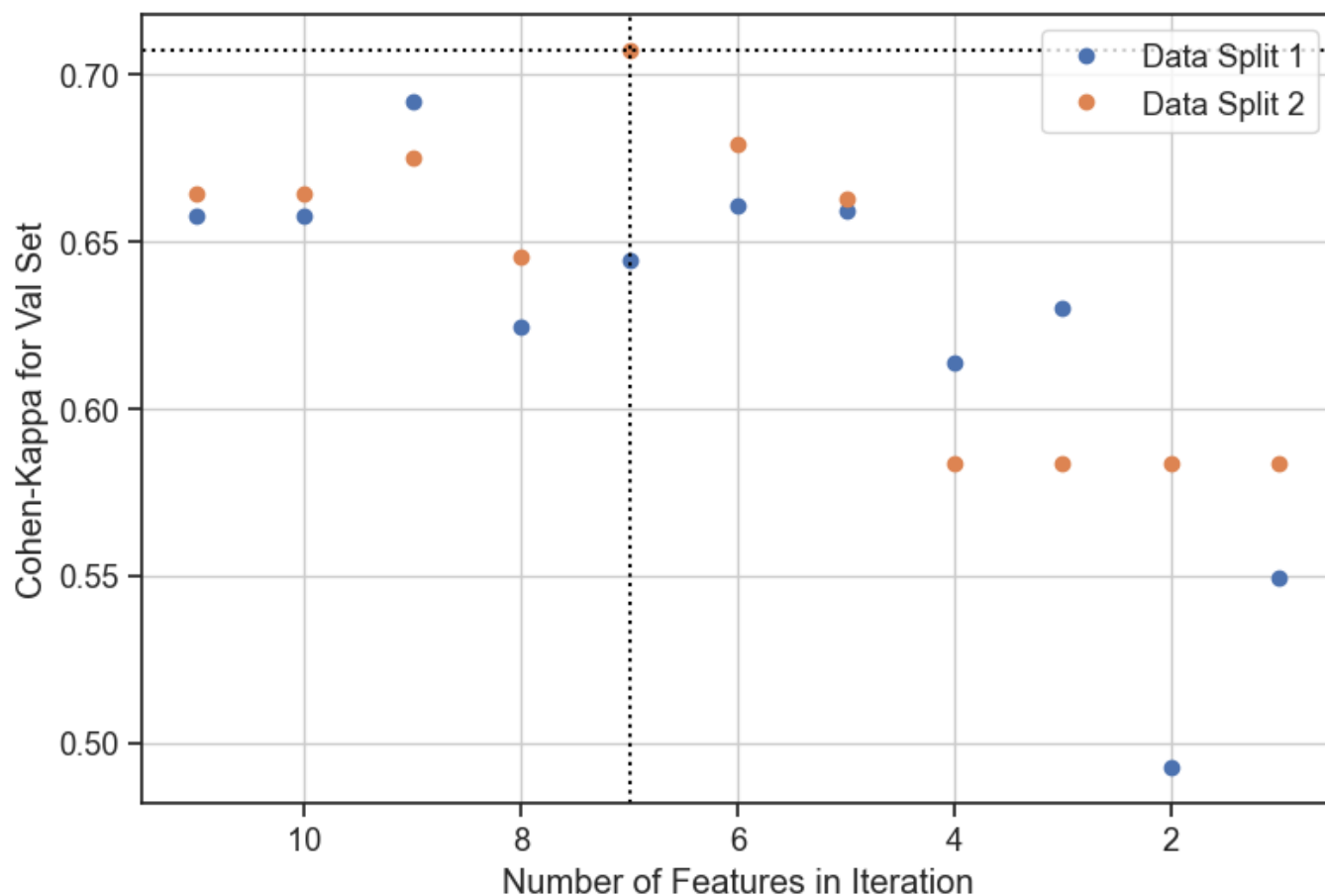
This plot is the same as the previous page, except with the secondary metric, CohKap.

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

The best value of CohKap is **0.70702**, compared to the starting value of **0.66424**.

Note that with the Cohen-Kappa score, the possible range is 0 to 1, with 0 meaning that the model has no predictive power and 1 is the best it could be.

With this metric, the best model training result occurred with **11 features**.

Exploring Feature Importance during Iterative Training
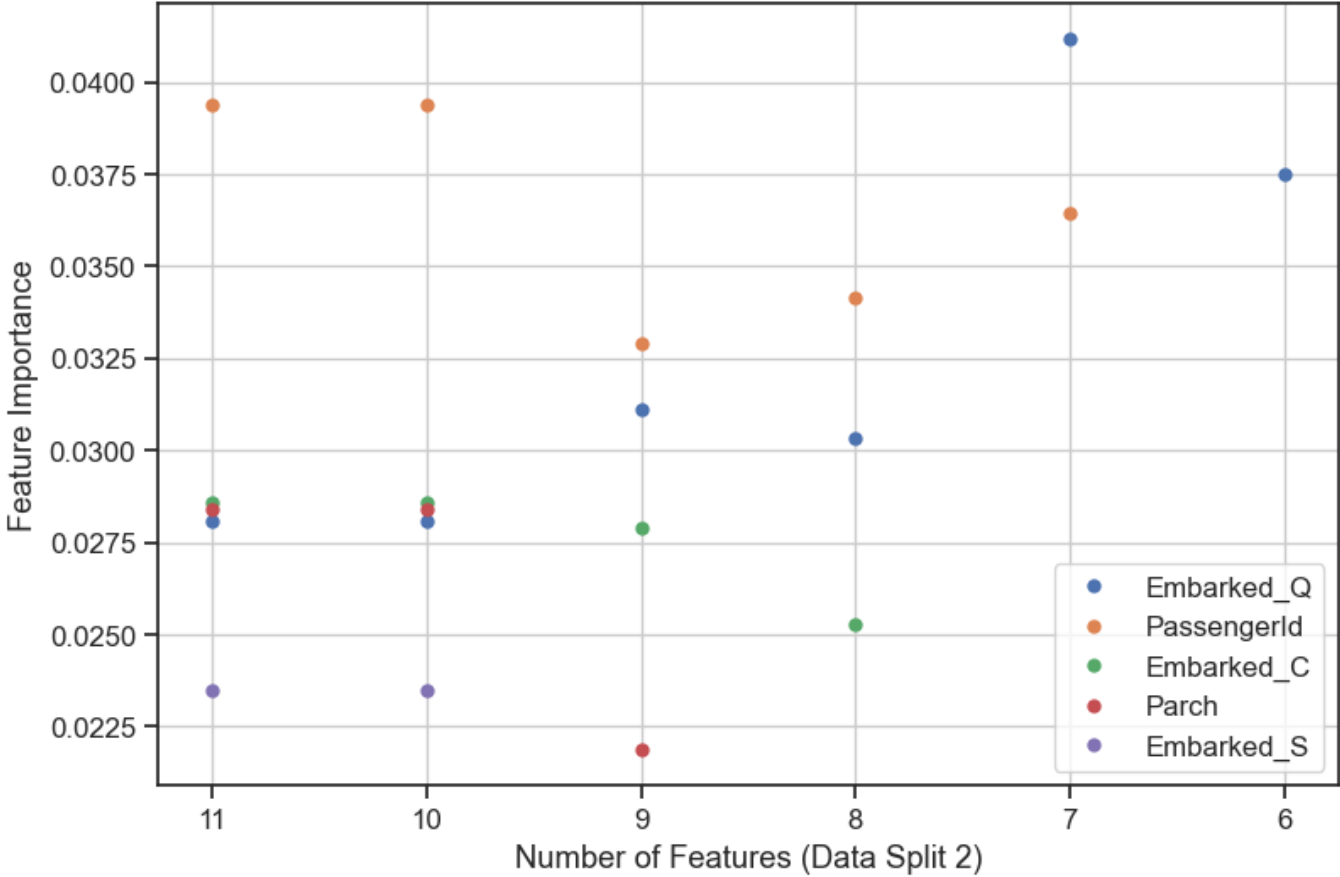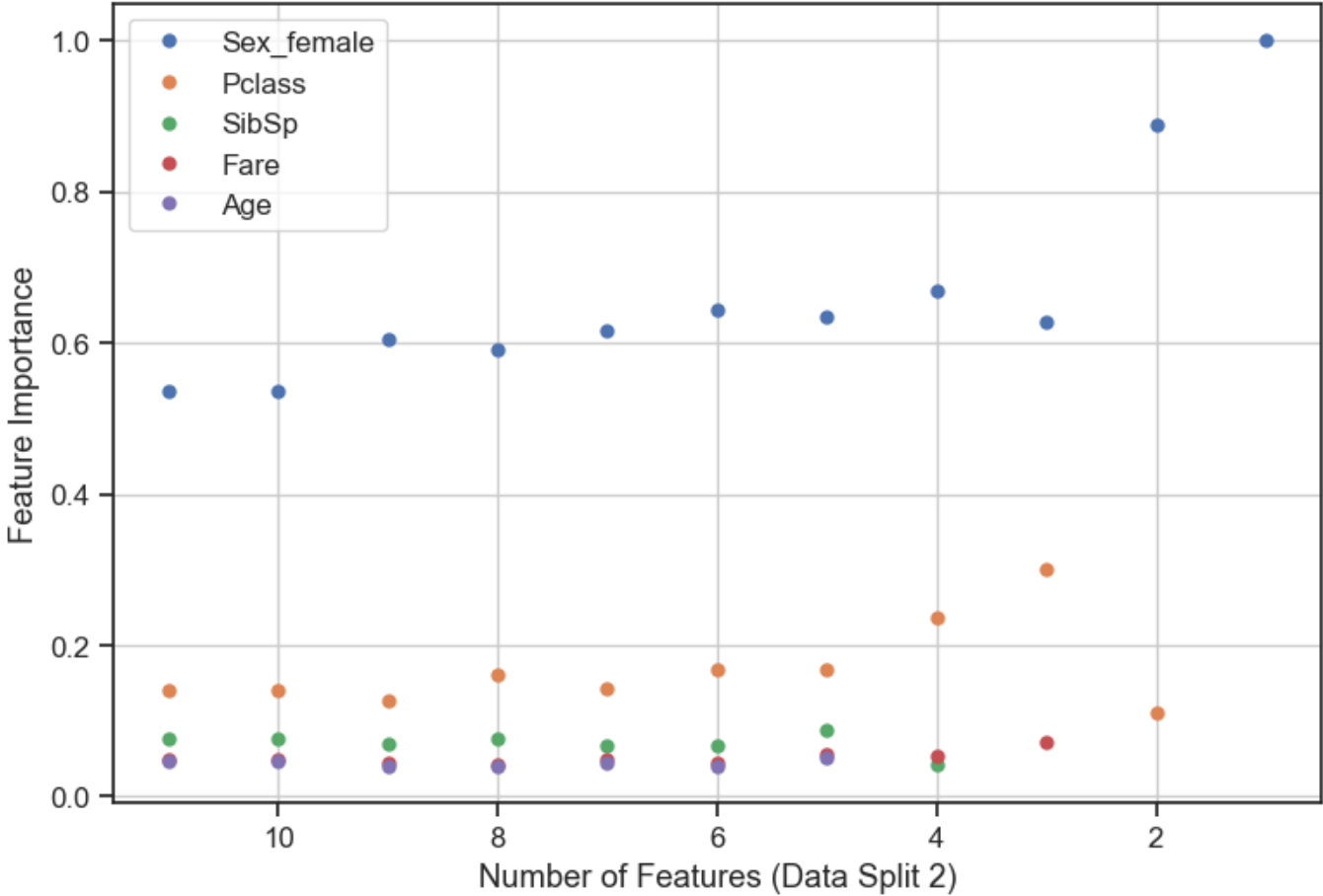
Exploring Feature Importance Compared to Individual Feature Correlations

Target Correlation vs Feature Importance