Dear Danielle,

The followings illustrate the analysis and prediction we conducted on customer brand preferences based on the completed survey to predict missing values in the incomplete survey.

Stochastic and Gradient Boosting model classifiers have been for our prediction. Both Models showed Salary and Age as the most significant contributors to the models. We used mtry =5 tunings for Random forest to optimize the model and automatic tuning for GBM. We decided to use GBM model due to greater accuracy and efficiency- please refer to the table below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | mtry | Accuracy | Kappa |
| GBM | 3 | 0.93 | 0.86 |
| RF | 5 | 0.92 | 0.83 |

When we applied GBM model to the incomplete survey for our prediction, the accuracy level dropped to 39% due to data corruption and no ground truth.

Based on our analysis, we can conclude 62% of the customers prefer Sony brand in comparison Acer - 38% and salary and Age are the most critical features in our prediction. Please see the chart below:

Armin