Customer Brand Preferences Report

Thomas Higginbotham

# Executive Summary

Danielle Sherman, CTO of Blackwell Electronics, had requested a prediction of customers’ brand preferences that were missing from incomplete surveys by conducting two classification methods using the R language (C5.0 or RandomForest).

Of the two methods, the RandomForest method performed slightly better, and was used to predict the results. The final results, out of 14898 total survey responses from the complete and incomplete surveys, suggest that 5623 customers prefer the Acer brand, while 9275 customers prefer the Sony brand.

# Methods and Findings

Two files were provided:

* CompleteResults.csv – containing 9898 survey results that were complete
* SurveyIncomplete.csv – containing 5000 survey results that were incomplete

Both files contained columns for salary, age, education level, type of car, first digit of zip code, credit and brand preference.

Using the CompleteResults.csv file, a training and testing data set were created using a 75%/25% split. With all classification methods, 10-fold cross-validation was used. Additionally, the Accuracy and Kappa values (representations of how well the methods performed with the data provided) were calculated for each possible method, along with the most important variables each model identified for prediction.

Both methods had similar Accuracy and Kappa scores, but the Random Forest classification method scored slightly better in both, so that was used. Salary was identified as the most important variable for both methods. Both methods also identified age as very important, and the C5.0 method also identified car as very important. The final predictions were done using the Random Forest classification method, based on Salary as it was 100 across the board.

# Conclusions

Of 5000 incomplete survey responses, 1879 were predicted to prefer the Acer brand, and 3121 were predicted to prefer the Sony brand.

Chart, treemap chart

Description automatically generated

Of the 14898 existing and predicted responses (the CompletedResults.csv and the above predicted results, combined), 5623 favored the Acer brand, and 9275 favored the Sony brand.

Chart, treemap chart

Description automatically generated