

CUSTOMER BRAND PREFERENCES REPORT

July 1, 2020

Danielle Sherman
Chief Technology Officer
Blackwell Electronics
www.blackwellelectronics.com

Re: Sony vs Acer computers – Customer Preferences

Dear Ms. Sherman,

Please find the summary of our findings below.

PRE-PROCESSING AND EXPLORATORY DATA ANALYSIS

The provided data was pre-processed. Parameters such as “elevel”, “car”, “zipcode”, and “brand” were converted into the “factor” datatype. Coefficients of correlation and statistical visualizations (scatterplots) detected the categorical relationships between the age, salary, and brand with other features lagging behind.



CUSTOMER BRAND PREFERENCES REPORT

PREDICTIVE MODELS

The classification models were trained on the data consisted of the completed surveys.

C5.0 (tree-based/rule-based classifier)

The model showed the following performance:

C5.0 with 10-fold cross validation with Automatic Tuning Grid (optimal model out of 8 models)

Model	Winnow	Trials	Accuracy	Kappa
-------	--------	--------	----------	-------

rules	FALSE	10	0.9251083	0.8395142
-------	-------	----	-----------	-----------

C5.0 with 10-fold cross validation with Random Search (optimal model out 3 models)

Model	Winnow	Trials	Accuracy	Kappa
-------	--------	--------	----------	-------

tree	TRUE	69	0.9220077	0.8342263
------	------	----	-----------	-----------

C5.0 with 10-fold cross validation with Manual Grid

Model	Winnow	Trials	Accuracy	Kappa
-------	--------	--------	----------	-------

rules	FALSE	10	0.9212053	0.8310463
-------	-------	----	-----------	-----------

Random Forest Classifier with 10-fold cross validation and manually tuned 5 different mtry values

Accuracy	Kappa
----------	-------

0.8956071	0.7772633
-----------	-----------

All models assigned 100% importance to both “age” and “salary” features.

PREDICTIONS

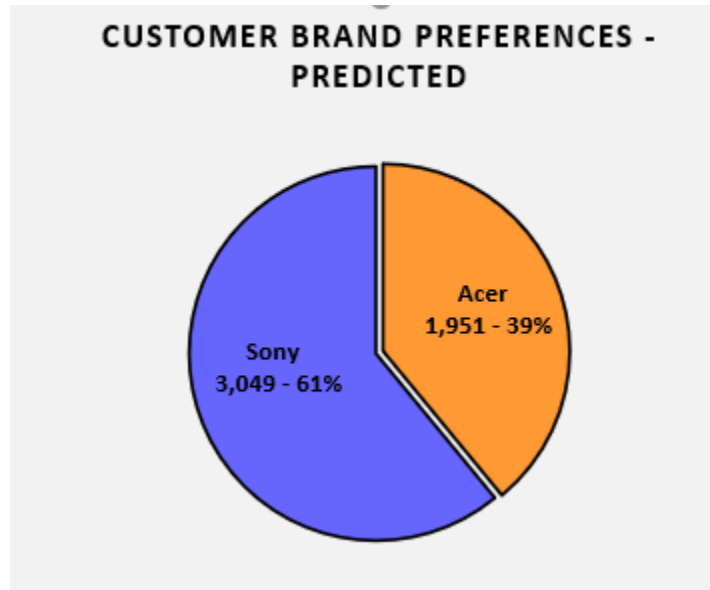
C5.0 with 10-fold cross validation with Random Search was selected as the optimal model for predictions due to the higher accuracy during testing.

The predictions were made based on the data consisted of incomplete surveys. The accuracy and the Kappa score (Accuracy - 0.40080000, Kappa - 0.01290381) significantly dropped due to the fact that values in certain parameters were missing.

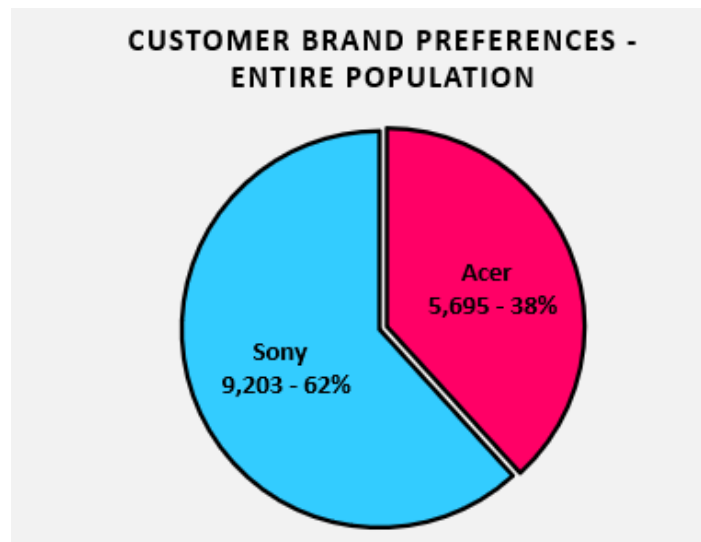
CUSTOMER BRAND PREFERENCES REPORT

CONCLUSION

The selected optimal model predicted that 1,951 customers who did not complete surveys would prefer Acer and 3,049 - Sony.



Overall, 5,695 customers would prefer Acer and 9,203 – Sony.



Let me know if you have further questions.