

Predicting Brand Loyalty in Grocery Shoppers

Rafael Rivera-Soto
rivera43@stanford.edu

Daniel Gardner
dangard@stanford.edu

Abstract—The abstract goes here.

I. INTRODUCTION

The Introduction goes here.

II. RELATED WORK

The related work goes here.

III. DATASET AND FEATURES

The Dataset and Features go here.

IV. METHODS

A. Logistic Regression

Our initial efforts were concentrated upon whether consumers tend to buy more branded or non-branded products. This is a binary classification task for which we implemented a Logistic Regression model. Logistic Regression squashes the output of the model in the range $y = \{0, 1\}$ using the sigmoid function (Eq. 1). The output values are then interpreted as the probabilities, thus any output greater than 0.5 is classified as belonging to the positive class and to the negative class otherwise (Eq. 2).

$$h_{\theta}(x) = g(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}} \quad (1)$$

$$\begin{aligned} P(y = 1|x; \theta) &= h_{\theta}(x) \\ P(y = 0|x; \theta) &= 1 - h_{\theta}(x) \end{aligned} \quad (2)$$

We based our initial predictions on products whose distributions between branded buyers and non-branded buyers was almost even. We labeled someone as being on the "branded-buyer" class if they're above the ratio cut-off value of 0.5, all other consumers were labeled as belonging to the "non-branded" class.

B. Support Vector Machines

C. Adaptive Boosting

V. RESULTS

The results go here.

VI. CONCLUSION

The conclusion goes here.

REFERENCES

- [1] H. Kopka and P. W. Daly, *A Guide to L^AT_EX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.