1 Introduction

Online shopping is prominent form of e-commerce in the world today that has a large market for consumers, as evident by the popularity of online shopping platforms such as Amazon. As such, there have been efforts to identify different types of online shoppers to improve an online shopping website's sales, or revenue [1].

A similar effort will occur given a dataset describing the properties and behaviours of online shopper visiting an unnamed online shopping site [2], that will be referred to as Nozama throughout the rest of the report. The goals of this report are to create an accurate predictor of which customer will buy something on Nozama and to determine which properties of users are associated with eventually buying items.

The analysis will first involve analyzing the datasets' properties, such as distribution and correlation, to help remove possibly redundant attributes and provide metrics for evaluating prediction models. Then the attributes of the dataset will be ranked and filtered based on importance and redundancy to improve clustering analysis and prediction model building. Once the dataset has been assessed wholistically, an analysis of clusters in the data to find insights regarding properties of eventual buyers and to help with prediction model building will occur. After that, a predictor will be constructed based on the findings in the data and assessed before finally making recommendations for Nozama's website.

References

[1] A. J. Rohm and V. Swaminathan, "A typology of online shoppers based on shopping motivations", *Journal of Business Research*, Marketing on the web - behavioral, strategy and practices and public policy, vol. 57, no. 7, pp. 748–757, Jul. 1, 2004. [Online]. Available at: https://www.sciencedirect.com/science/article/pii/S014829630200351X Last accessed: Nov. 15, 2021.

[2] C. O. Sakar, S. O. Polat, M. Katircioglu, and Y. Kastro, "Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks", *Neural Computing and Applications*, vol. 31, no. 10, pp. 6893–6908, Oct. 2019. [Online]. Available at: http://link.springer.com/10.1007/s00521-018-3523-0 Last accessed: Nov. 15, 2021.