XAI Applications in Automobile Financing Default

Kevin Sweet†
Data Science Graduate
University of Washington
Seattle WA USA
ksweet1@uw.edu

ABSTRACT

Consumer loan underwriting requires a high level of transparency in the output of predictive models. This constraint has limited the techniques possible for predicting the creditworthiness of an applicant. This work seeks to apply novel explainable AI frameworks to black box models that predict the probability of a debtor defaulting on their auto loan. These frameworks and model combinations are assessed for performance in both explainability and accuracy.

KEYWORDS

Consumer Finance, XAI

1. INTRODUCTION

This paper serves as a progress report and outline of my final report. The "Method & Approach" section contains the work that has been completed so far. The plan towards my final product is to next identify a suitable set of explainability frameworks and models to assess against the dataset. Then, using the typical train/test with cross validation methods, I will train a model and assess its accuracy in predicting the probability of default.

2. BACKGROUND & RELATED WORK

Predictive models, in particular logistic regression, have been used to model credit risk in the U.S. since before the 1980s .[2] The models used in practice for underwriting have been limited to only models inherently interpretable such as logistic regression or decision trees. Since the advent of explainable AI methods, there have been attempts to apply black-box methods with an explainable AI framework to consumer loans in Malta [3], but no known efforts have been made in the U.S.

3. METHOD & APPROACH

Loan-level data is not normally publicly available. The exception or this is when loans are bundled into securities and sold. This financial product is known as an asset-backed security which have public filing requirements from the SEC. Data was collected from SEC filings and preprocessed for analysis. This processed involved scripting a program that downloaded all of the relevant files. These files were filtered to only automobiles and merged into a single dataset for analysis. Due to the size of the data, this collection process took over 24 hours to complete.

4. EXPERIMENT DETAILS & RESULTS

5. CONCLUSION & DISCUSSION

REFERENCES

[1] Felton, R. (2021, October 27). *Many Americans are overpaying for their car loans*. Consumer Reports. Retrieved January 12, 2022, from https://www.consumerreports.org/car-financing/many-americans-overpay-for-car-loans-a8076436935/

[2] Wiginton, J. C. (1980). A Note on the Comparison of Logit and Discriminant Models of Consumer Credit Behavior. *The Journal of Financial and Quantitative Analysis*, *15*(3), 757–770. https://doi.org/10.2307/2330408

[3] Demajo, L.M., Vella, V., & Dingli, A. (2020). Explainable AI for Interpretable Credit Scoring. ArXiv, abs/2012.03749.Why is this important?