

# Automating Home Loan Approval

Group 4: Chuang Teng-Chieh (tengchieh.chuang@mail.utoronto.ca) & Ngezigihe Jean Paul (j.ngezigihe@mail.utoronto.ca) & Yu Keren (keren.yu@utoronto.ca)

## I. INTRODUCTION

The home loan approval process is a crucial step for many individuals and families looking to purchase a home. As a lender, the home loan approval process involves carefully assessing the borrower's financial situation to determine their creditworthiness and ability to repay the loan. Ultimately, the goal of the approval process is to mitigate the risk of default and ensure that the borrower can afford the loan.

## II. BUSINESS OBJECTIVES

The company wants to automate the loan eligibility process (real-time) based on customer detail provided while filling out the online application form. Here we are provided with a partial data set. Fig. 1 illustrates a typical home loan process where a loan officer has to manually do the underwriting.



Fig. 1. Typical Home loan process  
(<https://www.rate.com/resources/mortgage-loan-processing>)

The assumption is that we can automate this process and optimize this process through machine learning. Below are our objectives.

1) *Predict which customers will be eligible for a loan.*  
Predicting which customers' loans will be approved is helpful because Dream Bank can save time in the home loan process which can improve customer satisfaction. Additionally, we can reduce costs as we will need fewer employee labor hours.

## III. RESULTS

1) *Predict which customers will be eligible for a loan.*

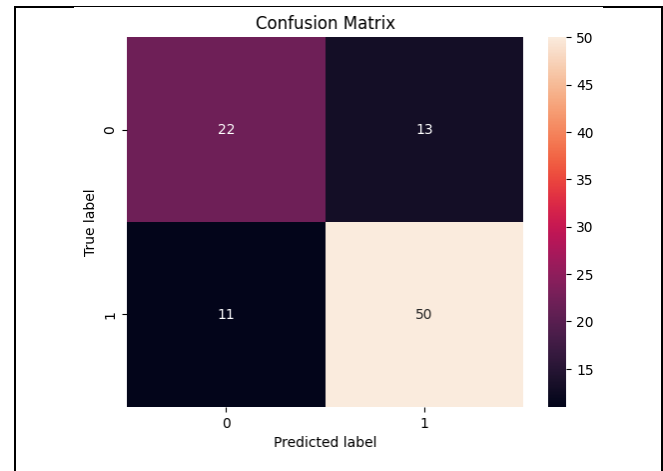
We trained two machine learning models; Logistic Classification and Decision tree.

### Decision tree

We estimated we could predict the correct loan approval 96% of the time. We obtained a 75% accuracy on the testing dataset. However, we want your model to be very accurate and not make many mistakes on the loan decision, therefore, we want high precision.

The model will predict the correct home loan approval 73% of the time (Metric: Precision. Initially, the model was

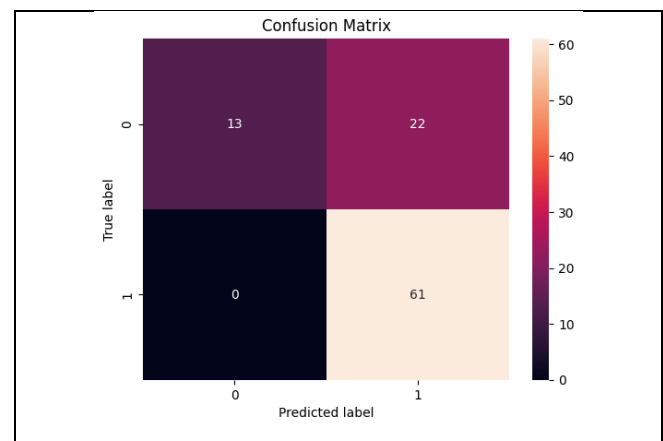
overfitting the training dataset. We have fixed that by balancing the class and it has made our model better. In Fig, 2, we see our model prediction results. In conclusion, we believe using decision three could assist in automating the home loan approval process.



### 2) Logistic Regression

We decide to use another model to compare with the decision tree. By using logistic regression as it is, we received a ~70% accuracy. We then normalized the data and were able to reach a 78% accuracy.

Looking at the precision, we reached 74% precision with the logistic regression.



## IV. CONCLUSION

Comparing the 2 models, we received very similar results. Both models produced decent predictions. We believe with a larger dataset, we could ultimately produce a trained model with better performance and potentially higher accuracy. This can lead to better home loan decisions and higher revenue.