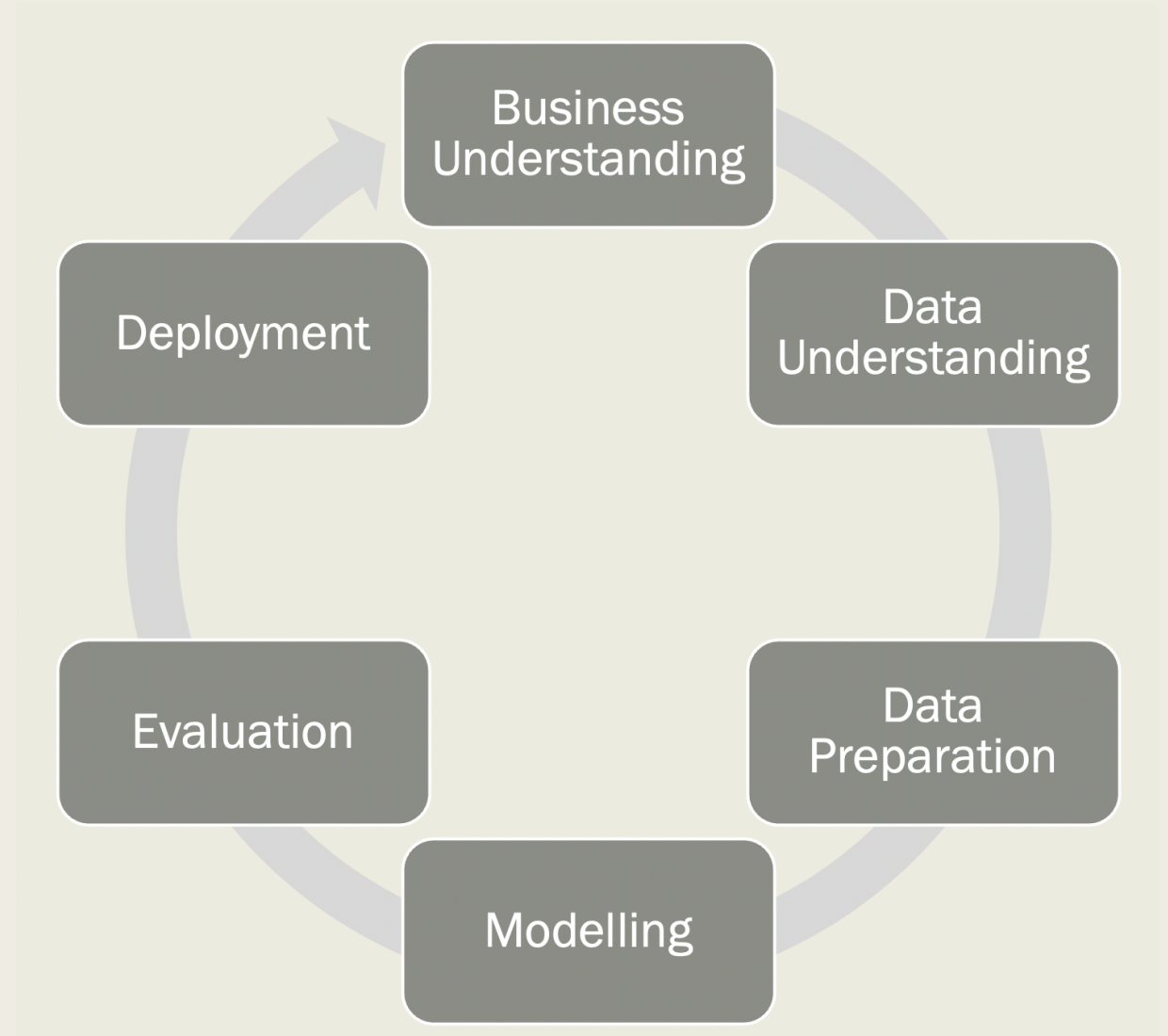


Home Loans: Exploration and Modeling

15/09/2022

Agenda: Data Science Lifecycle

- 1. Project Overview*
- 2. Data*
- 3. Analysis*
- 4. Modeling*
- 5. Model Evaluation*
- 6. Recommendations*



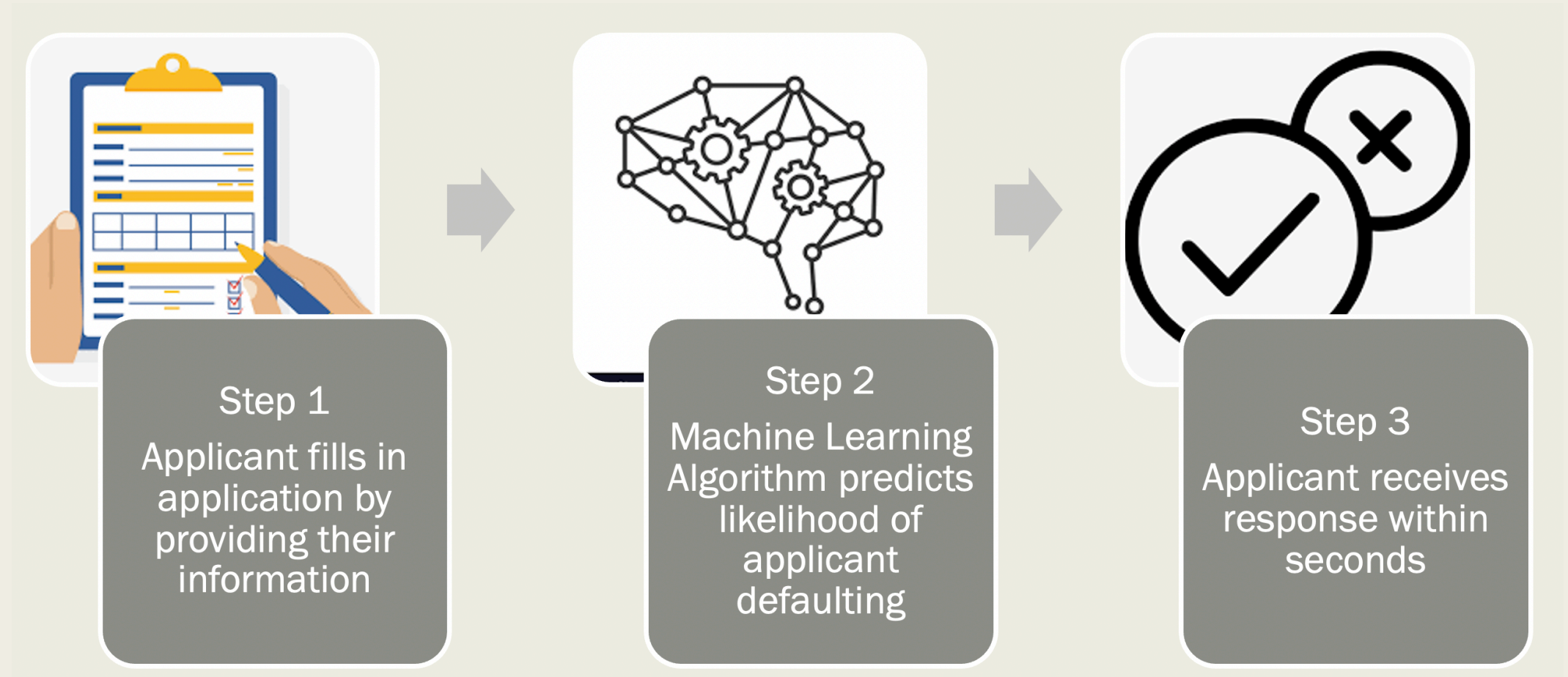
Project Overview

Business Problem: Home loan application processing is a manual process which takes around two or three days. The applicant must thus wait to receive their application decision.

Business Objective: To automate loan application processing, allowing applicants to obtain a response within a matter of seconds.

Method: A machine learning algorithm trained on historical data that will rapidly predict whether a loan applicant is likely to default or not.

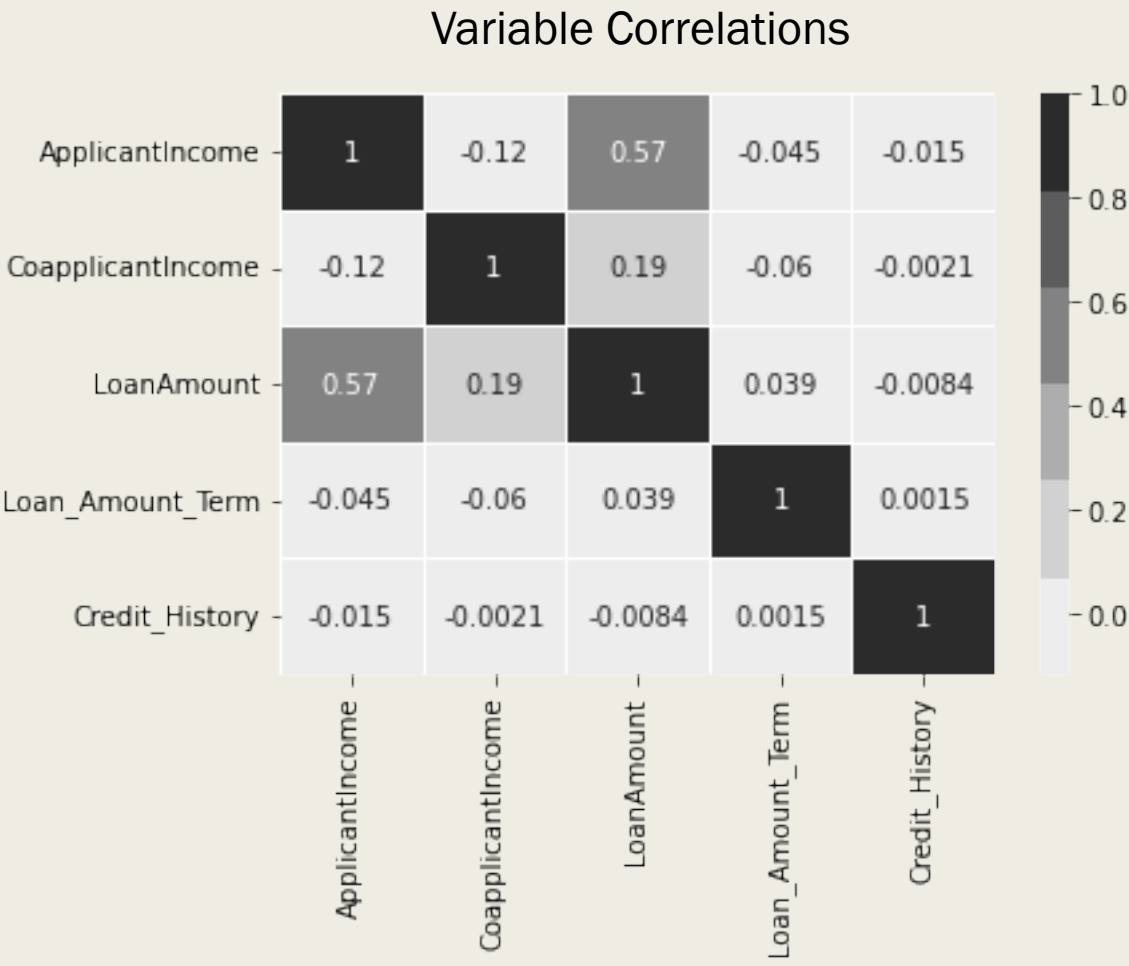
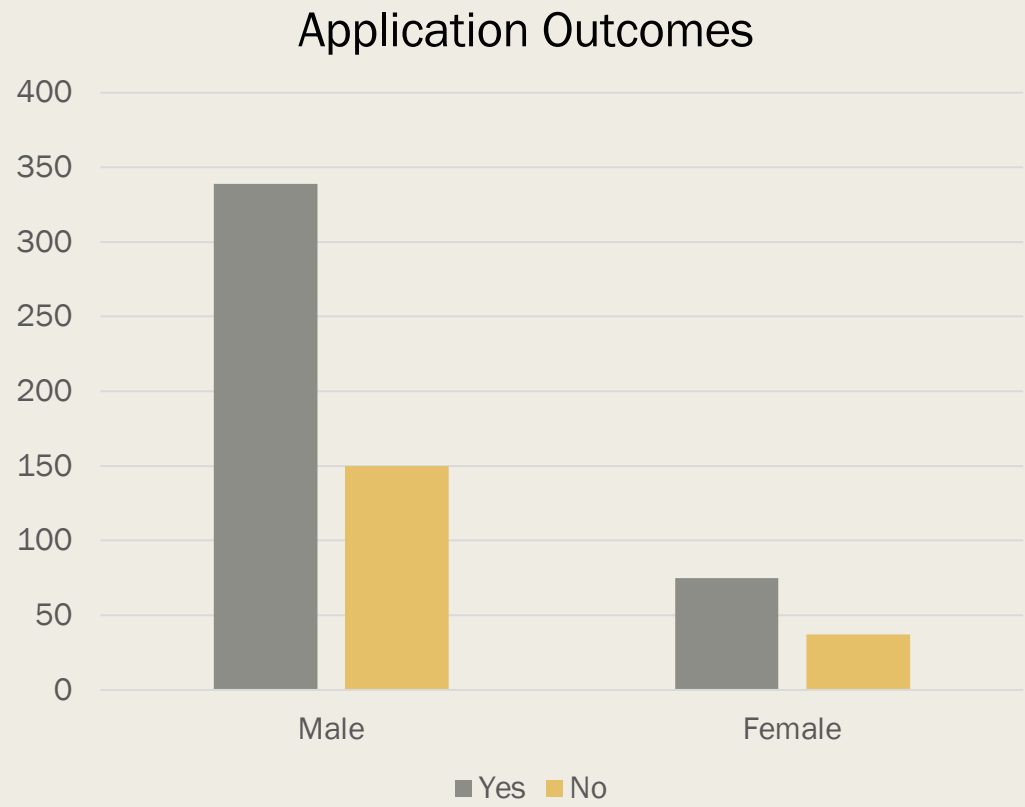
Process Overview / Solution



Training Data Description

- Number of entries: 614
- Number of variables: 13
 - 5 Numerical Variables
 - 8 Categorical Variables
- Target/Loan Status: Y (422 entries) and N (192 entries)

Data Analysis Overview



Modeling

- Two models were trained
- Bespoke model: required data preprocessing
- Automated model
- Both had similar accuracy

Model Evaluation

- Accuracy of models: approximately 78%
- Accuracy defined as percentage of correct predictions in the testing data set.

Recommendations

Bespoke model

- More understanding on algorithm used and its inputted data
- Less time required to train the model

Automated model:

- Commonly used as baseline model

Conclusion

- Bespoke machine learning model is better for our project.