ABC Bank Ltd.

Contents

## 1 EXECUTIVE SUMMARY

This document covers the model development process for **XGBoost\_2\_AutoML\_20210218\_195405** model. The model is a classification model that uses xgboost with input data consisting of 20000 observations and 70 features. The model achieves Auto of **75.84%** on validation dataset and **74.98%** on Out-of-Sample (OOS) test dataset.

## 2 MODEL PERFORMANCE SUMMARY

| Dataset | Size | Auto |
| --- | --- | --- |
| Validation | **?var:Validation.Size** | **?var:Validation.Auto** |
| OSS Test | **?var:OOS\_test.Size** | **?var:OOS\_test.Auto** |

## 3 DATASET

Following dataset were used for model training, tuning and OOS performance estimation:

| Dataset | Size | Features | Purpose |
| --- | --- | --- | --- |
| Train | **?var:Train.Size** | **?var:Train.Features** | **?var:Train.Purpose** |
| Validation | **?var:Validation.Size** | **?var:Validation.Features** | **?var:Validation.Purpose** |
| OSS Test | **?var:OOS\_test.Size** | **?var:OOS\_test.Features** | **?var:OOS\_test.Purpose** |

## 4 EDA

Following is a summary of input data. Refer Annexure-1 for detailed EDA.

|  |
| --- |
| Image |

## 5 Methodology Overview

**?var:Algorithm.Name** **?var:Algorithm.Overview**

| Dataset | Size | Auto |
| --- | --- | --- |
| Validation | 1920 | 75.84% |
| OSS Test | 1990 | 74.98% |

Following is a summary of steps performed to train the model: 

### 5.1 Data Preparation

**?var:Data\_Preparation**

### 5.2 Feature Transformation

**?var:Feature\_Transformation**

### 5.3 Model Tuning

**?var:Model\_Tunning**

### 5.4 Model Performance Evaluation

**?var:Model\_Perfor\_Eval**

### 5.5 Model Stability

**?var:Model\_Stability**

## 6 Model Details

**?var:Model\_Details.details**

### 6.1 Model Hyperparameters

Following is a summary of key model hyperparameters:

| Dataset | Size | Auto |
| --- | --- | --- |
| Validation | 1920 | 75.84% |
| OSS Test | 1990 | 74.98% |

‘+——–+————+——-+| Name | LastName | Age |+========+============+=======+| Name | LastName | Age |+——–+————+——-+| Juan | Lopez | 22 |+——–+————+——-+| Luisa | Perez | 24 |+——–+————+——-+| Ana | Sanchez | 23 |+——–+————+——-+’ ### Important Features Following is a list of important features for the model: | Dataset | Size | Auto | |————-|—–|——| | Validation | **?var:Validation.Size** | **?var:Validation.Auto** | | OSS Test | **?var:OOS\_test.Size** | **?var:OOS\_test.Auto** |

|  |
| --- |
| Partial Dependance Graph |

### 6.2 Model Performance

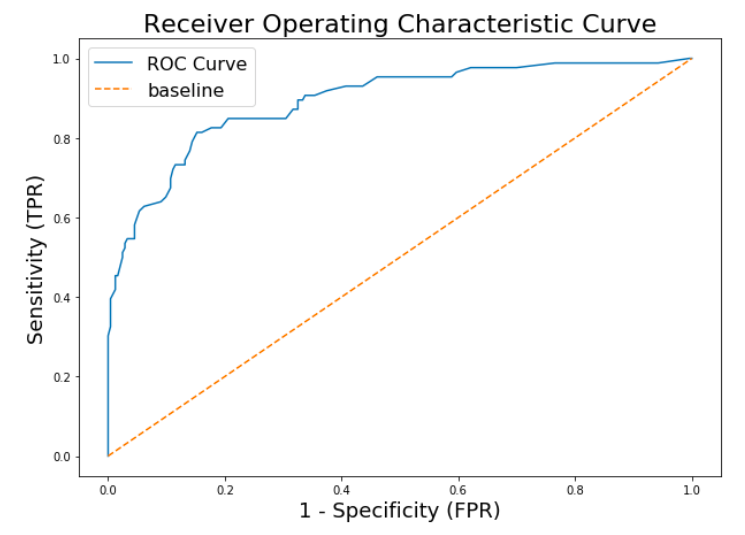
Following are the model performance statistics on validation and OOS test dataset: **Validation dataset**

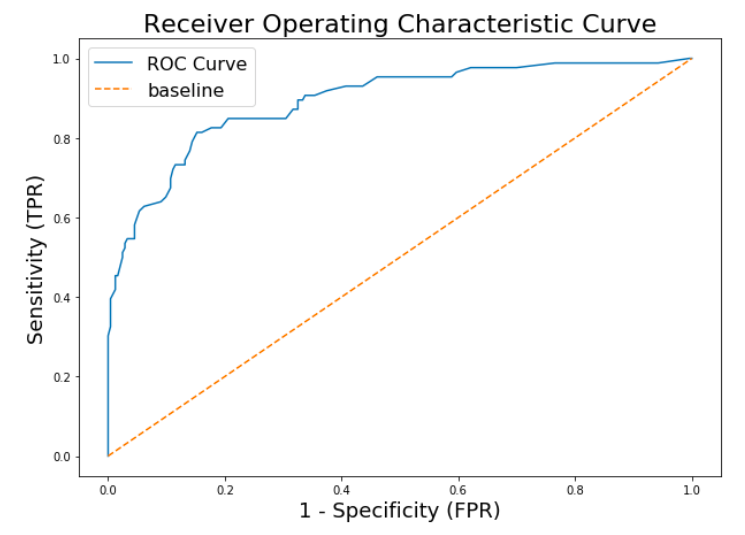
|  |
| --- |
| Model Performance on Validation dataset |

**Test dataset**

|  |
| --- |
| Model Performance on Test dataset |

### 6.3 Model Stability

Following are model stability statistics: **Train vs. Validation dataset** 

**Validation vs. Test dataset** 

### 6.4 Model Performance by Number of Features

Following is a summary of model performance, if it is replaced with a model with subset of important features. This may be used to identify if final model’s performance maybe matched with a simpler model with less number of features.