

# Configuration Manual

Research Project  
MSc Fintech 2021-22

Ashreet Sangotra  
Student ID: x20204523

School of Computing  
National College of Ireland

Supervisor: Mr Victor Del Rosal

**National College of Ireland**  
**MSc Project Submission Sheet**

**School of Computing**

**Student Name:** Ashreet Sangotra  
**Student ID:** x20204523  
**Programme:** MSc Fintech **Year:** 2021 - 22  
**Module:** Research Project  
**Lecturer:** Mr Victor Del Rosal  
**Submission Due Date:** 15.08.22  
**Project Title:** Exploring differences in Consumer Complaint Behaviour in Financial Products and modelling disputed responses using classification techniques  
**Word Count:** 545 **Page Count:** 5

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

**Signature:** .....

**Date:** 15th August 2022

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

|   |                          |
|---|--------------------------|
| Attach a completed copy of this sheet to each project (including multiple copies)   | <input type="checkbox"/> |
| <b>Attach a Moodle submission receipt of the online project submission,</b> to each project (including multiple copies).  | <input type="checkbox"/> |
| <b>You must ensure that you retain a HARD COPY of the project,</b> both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | <input type="checkbox"/> |

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

| <b>Office Use Only</b>           |  |
|----------------------------------|--|
| Signature:                       |  |
| Date:                            |  |
| Penalty Applied (if applicable): |  |

# Configuration Manual

**Student Name:** Ashreet Sangotra

**Student ID:** x20204523

## 1 Introduction

An overview of the research technique is provided in this document. This implies that this paper will have all of the information needed to repeat this study. This covers the Python version and packages used, the laptop used for the study, the code implemented for each of the necessary analyses, the memory of the laptop on which the analysis was performed, the plots produced by the analysis, the software processor, the speed utilization, and every entity or element used for the study. This helps to show how well the study was carried out and enables the study to be repeated if necessary.

## 2 Specification of Device Used

- **Model:** MacBook Air M1, 2020, 256 GB
- **Processor:** Apple M1
- **RAM:** 8 GB
- **Graphics:** Apple M1 GPU Built-In 7-core

## 3 Tools and Software Used

- Name of Programming Environment: Google Colab
- Name of programming language: Python
- Version of programming Environment: 3.7.13

## 4 Libraries Used

All libraries and settings mentioned below have been pre-installed into the code files and shouldn't need any further installation.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import os
import warnings
```

```
warnings.filterwarnings('ignore')
pd.set_option('display.float_format', lambda x: '%.5f' % x)
from tqdm import tqdm
from google.colab import files
import folium
import re

from sklearn.model_selection import train_test_split, cross_val_score, StratifiedKFold
from sklearn.linear_model import LogisticRegression
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.preprocessing import StandardScaler, RobustScaler, LabelBinarizer
from sklearn.metrics import recall_score, precision_score, f1_score, accuracy_score,
confusion_matrix, plot_confusion_matrix, roc_auc_score
import time
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
```

## 5 Steps to reproduce the study

- Go to the link - [bit.ly/codework22](https://bit.ly/codework22) for the dataset and the codework files.
- The codework won't work until the files are on your own google drive. So, download the entire folder named "Research Project Fintech" with all its content in your system. The folder will be zipped and downloaded.
- Go to the zipped folder "Research Project Fintech.zip" on your system, unzip it or extract its contents into a normal folder named "Research Project Fintech".
- Now upload the unzipped folder "Research Project Fintech" to the "My Drive" section of your Google Drive. Note: Please do not upload the folder into any subfolder on google drive. The file path defined in the codework will not work if done so.
- Open the folder "Research Project Fintech" that you just uploaded on "My Drive".
- You will find the following documents enclosed-
  1. complaints.csv (Project Dataset)
  2. EDA.ipynb (Code work file for EDA part of the project)
  3. Model Building (Main).ipynb (Code work file for main Model Building)
  4. Model Building (NLP).ipynb (Code work file for Model Building using NLP)
  5. zip\_code.csv (Used to define geographical location in the codework)

- We can start by first opening EDA.ipynb. the file will open in google colab.
- You can go to “Runtime” in the toolbar and Click “Run all”.
- Google Colab will ask to access your drive for files. Allow it by clicking on “Connect to Google Drive” and choosing your correct account.
- The notebook should execute without any interruption
- Similarly, run the Model Building (Main).ipynb and Model Building (NLP).ipynb files to check the execution of models used.
- Note: Some cells will take a few minutes to run, please give the code enough time to execute.