National College of Ireland

Configuration Manual

MSc Research Project
Artificial Intelligence

Likith Harish
Student ID: 22196269

School of Computing
National College of Ireland

Supervisor: Rejwanul Haque

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Likith Harish
Student ID:	22196269
Programme:	Artificial Intelligence
Year:	2024
Module:	MSc Research Project
Supervisor:	Rejwanul Haque
Submission Due Date:	02/09/2024
Project Title:	Configuration Manual
Word Count:	364
Page Count:	2

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 September 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

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

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

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Likith Harish
22196269

1 System Requirements

- Processor: AMD Ryzen 5 3550H or intel processor.
- RAM: 12 GB (9.88 GB usable)
- Operating System: Windows 11 Home Single Language, Version 22H2, OS Build 22621.3880
- Python Version: Python 3.11.4 (64-bit)

2 Software and Libraries

- Python 3.11.4 is required to run the experiments. Ensure that the correct version of Python is installed.
- IDE/Editor: You can use any Python-compatible IDE or text editor such as PyCharm, VS Code, or Jupyter Notebook.

3 Python Libraries

The following Python libraries need to be installed to replicate the experimental setup:

- pandas (for data manipulation and analysis)
- numpy (for numerical computations)
- scikit-learn (for machine learning models and evaluation metrics)
- matplotlib (for plotting and visualization)
- seaborn (for enhanced data visualizations)
- imbalanced-learn (for handling imbalanced datasets)
- joblib (for saving and loading models)
- xgboost (if you intend to explore Gradient Boosting algorithms)
- scipy (for additional scientific computing utilities)

4 Code Setup

- Place the Python scripts and data files in an organized directory structure.
- If using an IDE, ensure that the working directory is set correctly so that the scripts can access the data files and any saved models.

5 Results and Visualization

- The outcomes, encompassing model performance indicators and tailored loan proposals, ought to be automatically produced and stored by the script.
- Visualization libraries, including matplotlib and seaborn, will facilitate the creation of graphical representations such as confusion matrices, feature importance charts, and ROC curves. These visualizations should be incorporated into the code to ensure automatic generation and preservation of the plots as necessary.

6 Model Saving and Loading

Use joblib to save trained models for later use. Ensure that models are saved to a directory where they can be easily retrieved for predictions or further analysis.

7 Additional Notes

- The experimental configuration is tailored to operate on the designated hardware specifications. Should the experiment be conducted on an alternative system, it is imperative to verify that the system meets or surpasses the established requirements.
- The scripts are intended to be modular, facilitating straightforward modifications to the data processing workflow, model parameters, and evaluation metrics.

References