

Configuration Manual

MSc Research Project
MSc In Fintech

Jenyffer Noelia Mazat Ixcol
Student ID: 24112551

School of Computing
National College of Ireland

Supervisor: Victor del Rosal

National College of Ireland
MSc Project Submission Sheet
School of Computing



Student Name: Jenyffer Noelia Mazat Ixcol
Student ID: 24112551
Programme: MSc In Financial Technology **Year:** 2024-2025
Module: MSc Research Project
Lecturer: Victor Del Rosal
Submission Due Date: 11th of August 2025
Project Title: Design of an XAI-Enhanced Machine Learning Audit Model for Ethical, Technical, and Regulatory Compliance in the European Financial Sector.
Word Count: **Page Count:** 7 Pages excluding cover and ethics form.

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: Jenyffer Mazat
Date: 11th of August 2025

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

Jenyffer Noelia Mazat Ixcol
Student ID: 24112551

This document specifies the configuration, environment setup, and execution steps required to replicate the MSc thesis:

“Design of an XAI-Enhanced Machine Learning Audit Model for Ethical, Technical, and Regulatory Compliance in the European Financial Sector”

These configuration details ensure reproducibility and prevent undocumented assumptions. The FinalAuditModel.ipynb environment is intended for researchers, auditors, and regulators seeking to replicate results.

1. Software Components

Component	Version (Colab)	Purpose in Project
Python	3.11.x	Core programming language
pandas	2.1.x	Data manipulation and cleaning
numpy	1.26.x	Numerical operations
matplotlib	3.8.x	Data visualization
seaborn	0.13.x	Statistical visualizations
scikit-learn	1.3.x	Machine learning models & metrics
imbalanced-learn	0.12.x	Data balancing (SMOTE)
shap	0.43.x	Global & local explainability
lime	0.2.x	Local instance explainability
interpret	0.5.x	Explainable Boosting Machine
fairlearn	0.10.x	Fairness metrics & visualization
dill	0.3.x	Object serialization
dataframe_image	0.2.x	Export DataFrames as images
reportlab	4.0.x	PDF report generation
XlsxWriter	3.x	Excel export
weasyprint	60.x	HTML-to-PDF conversion
cairosvg	2.x	SVG conversion
tinycss2	1.x	CSS parsing
cssselect2	0.7.x	CSS selector parsing
joblib	1.3.x	Model saving/loading
IPython	8.x	Display & auxiliary functions
dateutil	2.x	Advanced date manipulation
datetime (builtin)	N/A	Date and time handling
json (builtin)	N/A	JSON file handling
os (builtin)	N/A	File path & OS operations
pickle (builtin)	N/A	Object serialization (native Python module)

2. Special Configurations

Runtime Settings:

- Platform: Google Colab (hosted Jupyter environment)
- Runtime type: Python 3.11
- Hardware accelerator: GPU (optional). Recommended for faster SHAP execution.
- Recommendation: Use Google Colab Pro for faster cross-validation and Explainable Boosting Machine (EBM) model training due to higher memory and dedicated GPU availability.

Persistent Storage:

- All datasets, models, audit logs, and reports stored in:
/content/drive/MyDrive/Thesis/ (*Google Drive mounted via Colab*)

python

Copy Edit

```
from google.colab import drive
drive.mount('/content/drive')
```

Reproducibility:

- Fixed random seed (random_state=42) applied to:
 - Data sampling
 - Model training
 - Cross-validation

3. Dataset Information

Primary Dataset:

- Name: Lending Club Loan Data (public version)
- Source: Kaggle – Lending Club Dataset
- Format: CSV
- Location: /content/drive/MyDrive/Thesis/full_dataset.csv

Processing Stages:

1. Raw Sampling – 200,000 random records.
2. Cleaning – Remove redundant features, define sensitive attributes.
3. Target Labeling – Regulatory-aligned risk categories.
4. Feature Encoding – Preprocessing aligned with transparency requirements.

Ethical Adjustments:

- Removal of personal identifiers.
- Replacement of geographical codes with anonymized equivalents.
- Full compliance with GDPR and EU AI Act.

4. Directory Structure

```
/content/drive/MyDrive/Thesis/
├── full_dataset.csv
├── sample_raw.csv
├── sample_encoded.csv
├── models/
│   ├── lr.joblib
│   ├── rf.joblib
│   ├── ebm.joblib
│   ├── X_train.pkl
│   ├── X_test.pkl
│   ├── y_train.pkl
│   └── y_test.pkl
├── plots/
│   ├── shap_summary_*.png
│   ├── ebm_*.png
│   └── fairness_*.png
├── reports/
│   ├── block*_*.csv
│   ├── fairness_comparison_rf_ebm.csv
│   └── master_audit_log.csv
├── Final_Audit_Report/
│   ├── final_audit_report.*
│   ├── final_audit_report_with_evidence.*
│   ├── exec_summary.*
│   └── Final_Audit_Report_Package.html
├── audit_logs/
│   ├── block1_log.json
│   └── ... block11_log.json
```

5. Execution Steps (Summary)

Step 1 – Environment Setup

- Open Google Colab (*see Section 2 – Special Configurations for detailed runtime and storage settings*).
- Set runtime to Python 3.11 (GPU optional; Colab Pro is recommended).
- Mount Google Drive. Install dependencies:

bash

Copy Edit

```
pip install pandas==2.1.* numpy==1.26.* scikit-learn==1.3.* \
imbalanced-learn==0.12.* shap==0.43.* lime==0.2.* interpret==0.5.* \
fairlearn==0.10.* dill dataframe_image reportlab xlswriter weasyprint \
cairosvg tinycss2 cssselect2
```

Step 2 – Data Preparation (Blocks 1–6)

- Load, sample, clean, label, encode dataset.
- Train baseline models (RF, LR, EBM).
- Outputs are stored in: /datasets, /models, /audit_logs.

Step 3 – Core Audit (Blocks 7–11)

- Technical Validation & Risk Audit.
- Explainability Analysis (SHAP, LIME, EBM).
- Fairness Audit.
- Compliance Mapping.
- Technical Audit Report generation.

6. Hardware Requirements

Minimum:

- CPU: 4-core.
- RAM: 8 GB.
- Storage: 2 GB.

Recommended:

- CPU: Quad-core or higher.
- RAM: 16 GB.
- GPU: NVIDIA Tesla T4 (*Colab Pro*).
- Storage: SSD, 5 GB.

7. Execution Notes

- Blocks must be executed sequentially (1–11).
- The directory structure must match exactly as specified.
- Audit logs must be preserved for compliance verification.

Appendix

Appendix A – Technical Execution Guide - AI Audit Framework

Purpose

This guide provides a **block-by-block execution process** to run the AI Audit Framework from data ingestion to final compliance report generation.

PHASE 1 — Model Simulation (Blocks 1–6)

- **Block 1 — Load & Sampling**
Load dataset, sample 200k rows, save sample_raw.csv, log stats.
Outputs: Thesis/sample_raw.csv, audit_logs/block1_log.json.
- **Block 2 — Cleaning / Minimisation**
Remove irrelevant/missing fields, save cleaning report, log.
- **Block 3 — Target Labeling**
Binarise loan_status → default/no default.
- **Block 4 — Encoding & Imputation**
Apply encoding, impute NaNs, create emp_group.
- **Block 5 — Split & Persistence**
Train/test split, save data artifacts, log.
- **Block 6 — Baseline Model Training**
Train LR, RF, EBM; save models, plots, metrics.

PHASE 2 — Audit Artefact (Blocks 7–11)

- **Block 7 — Technical Validation & Risk Audit**
CV unbalanced, leakage detection, CV with SMOTE, compare results.
- **Block 8 — Explainability (XAI)**
Global and local explanations (SHAP, LIME, EBM).
Map to EU AI Act articles 13–15.
- **Block 9 — Fairness Audit**
Calculate DI, SPD, EOD, Selection Rate; produce fairness plots.
- **Block 10 — Compliance Mapping + Evidence**
Compile master log, score traffic-light pillars, produce executive summary; *see Appendix B, Figures B.1 and B.2 for non-documentary and documentary traffic light scoring as displayed in the notebook.*
- **Block 11 — Packaging (Final Dossier)**
Generate final HTML/PDF report with KPIs, compliance heatmaps, gaps & mitigations; *see Appendix C, Figures C.1 and C.2 to view a section of the final audit report.*

Appendix B – Traffic Light Scoring

Figure B.1 Non-Documentary Audit Traffic-Light
(Visualization of audit scoring across technical validation and operational compliance pillars.)

	Pillar	Regulation	Expected_Blocks	Covered_Blocks	Missing_Blocks	Status	Status_Text	Score
0	Explainability	EU AI Act (Art 13, Art. 13, Article 13, Art 14, Art. 14, Article 14)	4,6,8	6,8	4	✓	Compliant	1.000000
1	Explainability	GDPR (Art 12, Art. 12, Article 12, Art 15, Art. 15, Article 15)	4,6,8	6	4,8	✓	Compliant	1.000000
2	Explainability	CBI Guidelines (Sec. 4.2, Section 4.2, Transparency)	4,6,8	4,6,8		⚠	Partial	0.500000
3	Explainability	EBA Guidelines (4.3.1, Model Explainability)	4,6,8	4,6,8		⚠	Partial	0.500000
4	Explainability	DORA (Art 15, Article 15)	4,6,8	4,6,8		⚠	Partial	0.500000
5	Fairness / Bias	EU AI Act (Art 10(2), Article 10, Art 14(4))	2,6,8,9	2,6,9	8	⚠	Partial	0.500000
6	Fairness / Bias	GDPR (Art 5(1)(a), Article 5(1)(a), Art 22, Article 22)	2,6,8,9	2,6,9	8	⚠	Partial	0.500000
7	Fairness / Bias	CBI Guidelines (Sec. 3.1, Ethics & Bias)	2,6,8,9	2,6,9	8	⚠	Partial	0.500000
8	Fairness / Bias	EBA Guidelines (5.2, Fair Treatment)	2,6,8,9	2,6,9	8	⚠	Partial	0.500000
9	Fairness / Bias	DORA (Art 6(2), Article 6(2))	2,6,8,9	2,6,9	8	⚠	Partial	0.500000
10	Traceability & Logging	EU AI Act (Art 10(3), Article 10)	1,2,3,4,5,6,7,8,9	5,6,7,8,9	1,2,3,4	⚠	Partial	0.500000

Figure B.2 Documentary Audit Traffic-Light
(Visualization of audit scoring across legal, ethical, and documentation-based compliance pillars.)

	Pillar	Regulation	Expected_Blocks	Covered_Blocks	Missing_Blocks	Status	Status_Text	Score	Status_Evidence	Status_Evidence_Text	Score_Evidence	Final_Status	Final_Status_Text	Final_Score
0	Explainability	EU AI Act (Art 13, Art. 13, Article 13, Art 14, Art. 14, Article 14)	4,6,8	6,8	4	✓	Compliant	1.000000	✗	Non-Compliant	0.000000	✗	Non-Compliant	0.000000
1	Explainability	GDPR (Art 12, Art. 12, Article 12, Art 15, Art. 15, Article 15)	4,6,8	6	4,8	✓	Compliant	1.000000	✗	Non-Compliant	0.000000	✗	Non-Compliant	0.000000
2	Explainability	CBI Guidelines (Sec. 4.2, Section 4.2, Transparency)	4,6,8	4,6,8	nan	⚠	Partial	0.500000	✗	Non-Compliant	0.000000	✗	Non-Compliant	0.000000
3	Explainability	EBA Guidelines (4.3.1, Model Explainability)	4,6,8	4,6,8	nan	⚠	Partial	0.500000	✗	Non-Compliant	0.000000	✗	Non-Compliant	0.000000
4	Explainability	DORA (Art 15, Article 15)	4,6,8	4,6,8	nan	⚠	Partial	0.500000	✗	Non-Compliant	0.000000	✗	Non-Compliant	0.000000
5	Fairness / Bias	EU AI Act (Art 10(2), Article 10, Art 14(4))	2,6,8,9	2,6,9	8	⚠	Partial	0.500000	⚠	Partial	0.500000	⚠	Partial	0.500000
6	Fairness / Bias	GDPR (Art 5(1)(a), Article 5(1)(a), Art 22, Article 22)	2,6,8,9	2,6,9	8	⚠	Partial	0.500000	⚠	Partial	0.500000	⚠	Partial	0.500000
7	Fairness / Bias	CBI Guidelines (Sec. 3.1, Ethics & Bias)	2,6,8,9	2,6,9	8	⚠	Partial	0.500000	⚠	Partial	0.500000	⚠	Partial	0.500000
8	Fairness / Bias	EBA Guidelines (5.2, Fair Treatment)	2,6,8,9	2,6,9	8	⚠	Partial	0.500000	⚠	Partial	0.500000	⚠	Partial	0.500000
9	Fairness / Bias	DORA (Art 6(2), Article 6(2))	2,6,8,9	2,6,9	8	⚠	Partial	0.500000	⚠	Partial	0.500000	⚠	Partial	0.500000
10	Traceability & Logging	EU AI Act (Art 10(3), Article 10)	1,2,3,4,5,6,7,8,9	5,6,7,8,9	1,2,3,4	⚠	Partial	0.500000	✓	Compliant	1.000000	⚠	Partial	1.000000

Appendix C – Final Audit Report

Figure C.1 Final Audit Report Cover
(Cover page of the generated audit report package.)

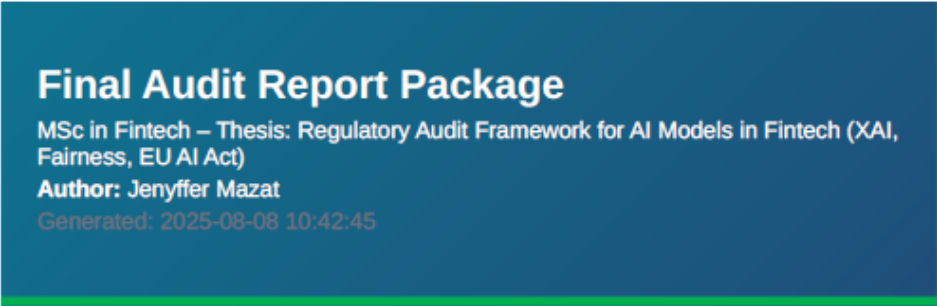


Table of Contents

- 1. Executive KPIs
- 2. Documentary Mapping (Block 10.2)
- 3. Evidence-Based Validation (Block 10.4)
- 4. Comparative Visuals
- 5. Top Gaps & Mitigation
- 6. Annex: Heatmaps & Raw Tables

1) Executive KPIs



Figure C.2 Mapping Features report
(Illustration of feature-level compliance mapping within the audit framework.)

2) Documentary Mapping (Block 10.2)

