National College of Ireland

Configuration Manual

MSc Research Project
Data Analytics

Prakruthi Barthur Prakash
Student ID: 23109742

School of Computing
National College of Ireland

Supervisor: John Kelly

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: ……Prakruthi Barthur Prakash ……

Student ID: ……X23109742……

Programme: ……Data Analytics…… **Year:** ……2023-2024.

Module: ……Research Project……

Lecturer: ……John Kelly……

Submission Due Date: ……12/08/2024……

Project Title: Utilizing Advanced Machine Learning Techniques for Predicting Fetal Health Risks

Word Count: ……690…… **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: ……**PRAKRUTHI BARTHUR PRAKASH**……

Date: ……12/08/2024……

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

Prakruthi Barthur Prakash
Student ID: X23109742

1. Introduction

This guide supports users in configuring the system according to individual operational needs with step by steps for basic setup and advanced personalization. This guide has been created to simplify the process of configuring it and make sure you obtain optimal performance/ functionality in any use case. Covering all important configuration facts, it offers straightforward explanations and concrete examples to help all users make the right decisions.

2. Dataset Description

Sl No.	Attribute Name	Data Type	Attribute Description
1.	Baseline Values	Float64	Heart Rate of Fetal
2.	Accelerations	Float64	Rate of accelerations per second
3.	Fetal Movement	Float64	Rate of fetal movements per second
4.	Uterine Contractions	Float64	Rate of uterine contractions per second
5.	Light Decelerations	Float64	Rate of Light Decelerations per second
6.	Severe Decelerations	Float64	Rate of Severe Decelerations per second
7.	Prolongued Decelerations	Float64	Rate of Prolongued Decelerations per second
8.	Abnormal Short-Term Variability	Float64	Percentile of time with abnormal short-term

			variability
9.	Mean Value of Short-Term Variability	Float64	Mean value of short-term variability
10.	Percentage of Time with Abnormal Long-Term Variability	Float64	Percentile of time with abnormal long-term variability
11.	Mean Value of Long-Term Variability	Float64	Mean value of long-term variability
12.	Histogram Width	Float64	Width of the histogram
13.	Histogram Min	Float64	Min Value of Histogram
14.	Histogram Max	Float64	Max Value of Histogram
15.	Histogram Number of Peaks	Float64	Rate of peaks in the exam histogram
16.	Histogram Number of Zeros	Float64	Rate of zeroes in the exam histogram
17.	Histogram Mean	Float64	Hist mean
18.	Histogram Mode	Float64	Hist mode
19.	Histogram Median	Float64	Hist median
20.	Histogram Variance	Float64	Hist variance
21.	Histogram Tendancy	Float64	Histogram trend
22.	Fetal Health	Float64	Fetal health: Normal, Suspect, Pathological

3. System Specification

3.1 Hardware Specification

Following are the hardware specifications of the system that was used to develop the project:
Processor: Apple M1 Chip
RAM: 16GB
Storage: 256GB
Graphics Card: 8-core GPU
Operating System: macOS Sonoma

3.2 Software Specification

Visual Studio is a comprehensive IDE used to write, edit, debug, and build code (Python) and the version of Visual Studio used here is 1.91.0 (Universal).

3.3 Python Packages/Libraries used

The following Python packages were installed using pip and used to implement the project

- Pandas
- Numpy
- Mathplotlib
- Seaborn
- Plotly
- Imblearn
- Scikit-learn
- Boruta

4. Implementation

1) Preparing of data
2) EDA (Exploratory Data Analysis)
3) Feature Selection using Boruta
4) Class balancing using SMOTE
5) Model Building: Decision Tree, Random Forest, KNN, GBM
6) Evaluation