



B.Sc. (Hons) in Software Development



Ollscoil  
Teicneolaíochta  
an Atlantaigh

Atlantic  
Technological  
University

Gesture UI

**By**  
**Rodrigo Almeida**  
**G00377123**

March 13, 2024

**Department of Computer Science & Applied Physics,  
School of Science & Computing,  
Atlantic Technological University (ATU), Galway.**

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Methodology</b>	<b>3</b>
2.1	Data Pre-processing . . . . .	3
2.2	Data Labelling . . . . .	4
2.3	Data Scaling . . . . .	4
2.4	Data Analysis and Visualisation . . . . .	5
2.5	Other Techniques . . . . .	5
<b>3</b>	<b>Experiments and Results</b>	<b>6</b>
3.1	Classifier Models . . . . .	6
3.2	Results . . . . .	6
<b>4</b>	<b>Conclusion</b>	<b>7</b>

# List of Figures

# List of Tables

2.1	Missing Values in the Dataset . . . . .	4
-----	---	---

# Chapter 1

## Introduction

# Chapter 2

## Methodology

### 2.1 Data Pre-processing

One of the biggest challenges in machine learning is the quality of the data. The quality of the data is crucial to the performance of the model. The data pre-processing phase is a critical step in the machine learning pipeline. It involves cleaning, transforming, and preparing the data for the model. The data pre-processing phase includes the following steps

#### **Imbalanced Dataset**

The dataset is imbalanced, which means that the number of instances for each class is not equal. The imbalanced dataset can lead to poor performance of the model.

Below are the number of instances for each class in the dataset:

- **Driving Style:**
  - EvenPaceStyle: 21,016 instances
  - AggressiveStyle: 2,759 instances
- **Road Surface Condition:**
  - SmoothCondition: 14,237 instances
  - UnevenCondition: 6,289 instances
  - FullOfHolesCondition: 3,249 instances
- **Traffic Condition:**
  - LowCongestionCondition: 17,764 instances

- HighCongestionCondition: 3,017 instances
- NormalCongestionCondition: 2,994 instances

For this project **Driving Style** class column was chosen due to its significant imbalance and potential for improvement. The **AggressiveStyle** class has 2,759 instances while the **EvenPaceStyle** class has 21,016 instances. The dataset is imbalanced, and the model may be biased towards the majority class.

## 2.2 Data Labelling

For the categorical variables (roadSurface, traffic, drivingStyle), needs to convert these into a format that can be used for machine learning models. Since drivingStyle is your target variable, let's focus on roadSurface and traffic.

First it was needed to handle the missing values in the dataset. The missing values were replaced with the mean of the column. Table below show the number of missing values in the dataset per column.

Table 2.1: Missing Values in the Dataset

Column	Missing Values
VehicleSpeedInstantaneous	9
EngineLoad	5
EngineCoolantTemperature	5
ManifoldAbsolutePressure	5
EngineRPM	5
MassAirFlow	5
IntakeAirTemperature	5
FuelConsumptionAverage	5

As this is a classification problem, the categorical variables need to be converted into a numerical format.

## 2.3 Data Scaling

Feature scaling is a method used to standardize the range of independent variables or features of the data. In data processing, it is also known as data normalization and is generally performed during the data pre-processing step. The **Standard Scaler** was used to scale the data. The Standard Scaler standardizes the features by removing the mean and scaling to unit variance. This results in a distribution

with a standard deviation of 1 and a mean of 0. The formula for standard scaling is given by:

$$z = \frac{x - \mu}{\sigma} \quad (2.1)$$

The scaler subtracts the mean of the feature and then divides by the standard deviation.

## 2.4 Data Analysis and Visualisation

## 2.5 Other Techniques



## Chapter 3

# Experiments and Results

### 3.1 Classifier Models

### 3.2 Results

# Chapter 4

## Conclusion

Briefly summarise your context and objectives. Remind the reader about the overall rationale and goals of the project. Highlight your findings from the System Evaluation chapter.