**CCT College Dublin**

**Assessment Cover Page**

|  |  |
| --- | --- |
| Module Title(s): | Programming for DA  Statistics for Data Analytics  Machine Learning for Data Analysis  Data Preparation & Visualisation |
| Assessment Title: | MSC\_DA\_CA2 |
| Lecturer(s) Names: | Sam Weiss  John O’Sullivan  Muhammad Iqbal  David McQuaid |
| Student Full Name: | Clelia Caetano |
| Student Number: | 2023060 |
| Assessment Due Date: | 26/05/2023 |
| Date of Submission: | 26/05/2023 |

**Declaration**

|  |
| --- |
| By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or another source. I declare it to be my work and that all material from third parties has been appropriately referenced. I further confirm that this work has yet to be submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution. |

Author: Clelia Caetano

E-mail: [2023060@student.cct.ie](mailto:2023060@student.cct.ie)

Student ID:2023060

Constructions:

Workplace Safety

CCT College Dublin

May (2023).

# Abstract

***Keywords:***

Table of Contents

[Abstract 3](#_Toc135823161)

[List of Tables 3](#_Toc135823162)

[Introduction 4](#_Toc135823163)

[Business Understanding 5](#_Toc135823164)

[Data Understanding 5](#_Toc135823165)

[Data Preparation 5](#_Toc135823166)

[Modelling 5](#_Toc135823167)

[Evaluation 5](#_Toc135823168)

[Deployment 5](#_Toc135823169)

[Conclusion 5](#_Toc135823170)

[1.1 - Statistics for Data Analytics 5](#_Toc135823171)

[1.1.1 – Dataset Overview 5](#_Toc135823172)

[References 5](#_Toc135823173)

# List of Tables

# Introduction

# Business Understanding

A picture containing text, line, screenshot, font

Description automatically generated

# Data Understanding

A picture containing text, line, screenshot, font

Description automatically generated

# Data Preparation

A picture containing text, line, screenshot, font

Description automatically generated

# Modelling

A picture containing text, line, screenshot, font

Description automatically generated

# Evaluation

A picture containing text, line, screenshot, font

Description automatically generated

-----Decision Tree Model Results-----

Accuracy on the training set: 1.0

Accuracy on the test set: 0.9775

Precision on the test set: 0.9772

Recall on the test set: 0.9772

F1-score on the test set: 0.9772

-----Random Forest Model Results-----

Accuracy on the training set: 1.0

Accuracy on the test set: 0.9888

Precision on the test set: 0.9881

Recall on the test set: 0.9896

F1-score on the test set: 0.9886

-----kNN Model Results-----

Accuracy on the training set: 0.9561

Accuracy on the test set: 0.9775

Precision on the test set: 0.977

Recall on the test set: 0.9792

F1-score on the test set: 0.9773

A picture containing text, screenshot, diagram, plot

Description automatically generated

# Deployment A picture containing text, line, screenshot, font Description automatically generated

# Conclusion

## 1.1 - Statistics for Data Analytics

## 1.1.1 – Dataset Overview

# References

**https://ec.europa.eu/eurostat/databrowser/view/HSW\_N2\_01\_\_custom\_6154887/default/table?lang=en**

<https://ec.europa.eu/eurostat/databrowser/view/HSW_N2_02__custom_6154914/default/table?lang=en>

<https://ec.europa.eu/eurostat/databrowser/view/SBS_SC_CON_R2__custom_6155496/default/table?lang=en>

# Appendix