A logo for college computing

Description automatically generated

**Assessment Cover Page**

|  |  |
| --- | --- |
| *Student Full Name:* | Kevin Scully |
| *Student Number* | 24275 |
| *Module Title* | Programming for AI |
| *Assessment Title* | CA2 |
| *Assessment Due Date* | 15th December 2024 |
| *Date of Submission* | 10th December 2024 |
|  |  |

**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 other source.

I declare it to be my own work and that all material from third parties has been appropriately referenced.

I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

Contents

[Introduction 1](#_Toc184749746)

[Objectives 1](#_Toc184749747)

[Problem Description 1](#_Toc184749748)

[Objectives 1](#_Toc184749749)

[Methodology 2](#_Toc184749750)

[A. Dataset 2](#_Toc184749751)

[B. Data Preparation 2](#_Toc184749752)

[i) Nil (0.00) values 2](#_Toc184749753)

[ii) Drop ID column 2](#_Toc184749754)

[iii) Outliers 2](#_Toc184749755)

[Architecture Diagram 3](#_Toc184749756)

[Results & Discussion 3](#_Toc184749757)

[References 5](#_Toc184749758)

[Appendices: 6](#_Toc184749759)

Introduction

Objectives

Problem Description

Develop a proposal by performing an initial analysis of the data (EDA) using python in your Jupyter notebook. Perform any preparation of the data, that you feel is necessary, using python in your Jupyter notebook. Explain your rationale behind your data preparation and how it will assist you. Discuss your findings and what relevance they might have on your planned Neural Network model. [0-20]

Data Preparation

Modelling

Evaluation

Deployment

Create and implement a scalable Dense Neural Network that will output a classification based on the 'Type of glass' feature (class attribute). Test this model and aim to enhance its performance by experimenting with various configurations of neurons, layers, loss functions, and activation functions. Additionally, apply hyperparameter optimization techniques to determine the optimal configuration. Finally, [0-35

Make a classification using your training as well as test data, using your final Neural Network configuration and discuss your findings and rationale for the chosen neural network configuration on the accuracy differential between the training and testing set.[0-15]

Objectives

Methodology

## Dataset

The ‘glass\_data’ set has been provided to be used for this assignment. This dataset provides the breakdown of nine different elements that are used to make a type of glass. There are 214 rows and 11 columns on the dataset. The columns are all numerical, with ‘type’ being categorical. The columns are ID, the nine elements (a column for each element) and the type of glass as a category.

## Data Preparation

The initial analysis of the data showed that there were no null values or duplicated rows. The categorical column, ‘type’, had the unique values of 1, 2, 3, 5, 6, 7 which was within the range expected.

### Nil (0.00) values

The description of the dataset showed that the columns ‘mg’, ‘k’, ‘ba’ and ‘fe’ had a minimum value of 0.00. There were no minus values within the dataset. It is assumed that the 0.00 values were accurate as it is possible that there was 0.00 of the elements in the formation of the glass type. As these were accurate values, it was decided for the model to only handle the values if they were extreme outliers. () The below graph shows the percentage of the 0.00 values in each column:

A graph of values per column

Description automatically generated

### Drop ID column

As the column ‘id’ had no impact on the glass type, this was dropped from the dataset. This would ensure that the column did not impact the model.

### Outliers

As the information within the dataset was accurate in the creation of the glass type, only extreme outliers were identified and handled. () The outliers were identified using the interquartile range (IQR) with 3 added to the lower and upper bounds. The range 3 was used instead of 1.5 as the outliers were deemed to be accurate so only extreme outliers were determined to have an impact on the data. The results are shown in the below graphs:

A group of images of a graph

Description automatically generated with medium confidence

configurations of neurons, layers, loss functions, and activation functions

Architecture Diagram

A dense neural network was built to

Results & Discussion

* 1200 (plus or minus 10%) words in a report on your topic that includes Introduction, problem description, objectives, methodology, architecture diagram, explanation about configurations of neurons, layers, loss functions, and activation functions, interpretation and discussion of your findings. In addition, also add the screenshots of your working code and its output [0-20].

# References

**There are no sources in the current document.**

***GitHub Link:*** https://github.com/kpscully116/Programmin-for-AI

# Appendices: