National College of Ireland

Configuration Manual

MSc Research Project
MSc Artificial Intelligence

Mugil Sivasamy Kalamani
Student ID: X22165096

School of Computing
National College of Ireland

Supervisor: Muslim Jameel Syed

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Mugil Sivasamy Kalamani
Student ID:	X22165096
Programme:	MSc Artificial Intelligence
Year:	2023 Jan - 2024 Jan
Module:	MSc Research Project
Supervisor:	Muslim Jameel Syed
Submission Due Date:	31/1/2023
Project Title:	Configuration Manual
Word Count:	480
Page Count:	3

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:	Mugil Sivasamy Kalamani
Date:	31st January 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

Mugil Sivasamy Kalamani
X22165096

1 System Requirements

Python: [Python 3.9 or higher]*Download Python* (n.d.)
Libraries: [TensorFlow, Keras, Numpy, Pandas, Matplotlib, SKlearn, NVIDIA CUDA, Python IDE(Pycharm or Notebook or Colab)]*NVIDIA CUDA Toolkit 12.1 downloads* (n.d.); *Download PyCharm: Python IDE for Professional Developers by JetBrains* (2021); *Install TensorFlow 2* (n.d.)
Hardware Requirements: [Min Requirement: i7 11750H Processor, RAM 16 GB DDR4 RAM, NVIDIA 3000 series or Higher Dedicated GPU with 6 GB VRAM]

2 Installation Guide

Python Installation: Ensure Python 3.9 or higher is installed on your system. Download and install from python.org if not already installed.[https://www.python.org/downloads/]
Set up a virtual environment: Set up a virtual environment to maintain the dependencies (like "conda" or "venv") . (Note: Only if needed)
Dependency Installation:
Install required libraries: pip install numpy pandas tensorflow keras matplotlib seaborn scikit-learn

3 Dataset Generation

The project uses synthetic datasets created through a Python script [Datagen.py].
The datasets involve various features derived from mathematical functions applied to linearly spaced 'x' values.
Features include linear transformations, trigonometric functions, polynomial features, and more.

The datasets are saved as CSV files for further use in the project.
The secondary dataset used is from the Google Air project, a street pollution lever data containing 5 million data points.*Google Project Air View Data - Dublin City (May 2021 - August 2022) - data.smartdublin.ie* (n.d.)

4 Neural Network Configuration and Training

Model Architecture:

Models are configured using TensorFlow and Keras. Configurable parameters include the number of layers, neurons per layer, and activation functions. For Model performance analysis, the Grid search technique is used for all possible combinations of Hyperparameters.
The Python script [NN_test.py] tests 3 different parameters and 5 values each. In total, there are 125 model combinations.
layers_grid = [1, 2, 3, 4, 5]
neurons_grid = [64, 128, 256, 512, 1024]
activation_grid = ['relu', 'tanh', 'sigmoid', 'softmax', 'linear']
For Simple line and curve fitting, The juypter notebook file [NN_neurons_layers.ipynb] configures all the parameters from no of neurons in each layer, no of layers and up-to learning rate.

Training Process: The models are trained on the generated synthetic datasets. Training parameters such as the number of epochs, batch size, and learning rate are configurable.

Evaluation:Models are evaluated using Mean Squared Error (MSE) as the metric. The performance of each model configuration is logged and analysed.

5 Data Visualization

Pair Plot and Heatmap: Utilising Seaborn and Matplotlib for visualising data distributions and correlations. Pair plots and heatmaps are generated to understand the relationships between different features in the dataset.

3D and 2D Plots for Model Performance: 3D scatter plots to visualise the model's performance across different configurations. 2D line and scatter plots to detail the MSE across various layers and neuron configurations.

6 Troubleshooting

Ensure all required libraries are correctly installed.
Verify the correct Python version is being used.
For TensorFlow and GPU usage issues, consult TensorFlow's official documentation.

References

Download PyCharm: Python IDE for Professional Developers by JetBrains (2021).
URL: *https://www.jetbrains.com/pycharm/download/?section=windows*

Download Python (n.d.).
URL: *https://www.python.org/downloads/*

Google Project Air View Data - Dublin City (May 2021 - August 2022) - data.smartdublin.ie (n.d.).
URL: *https://data.smartdublin.ie/dataset/google-airview-data-dublin-city*

Install TensorFlow 2 (n.d.).
URL: *https://www.tensorflow.org/install*

NVIDIA CUDA Toolkit 12.1 downloads (n.d.).
URL: *https://developer.nvidia.com/cuda-downloads*