National College of Ireland

Configuration Manual

MSc Research Project
Cloud Computing

Serena Santosh
Student ID: x23246642

School of Computing
National College of Ireland

Supervisor: Yasantha Samarawickrama

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Serena Santosh
Student ID:	x23246642
Programme:	Cloud Computing
Year:	2024
Module:	MSc Research Project
Supervisor:	Yasantha Samarawickrama
Submission Due Date:	12/12/2024
Project Title:	Configuration Manual
Word Count:	678
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:	Serena Santosh
Date:	10th December 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

Serena Santosh
x23246642

1 Setup Creation

This section describes the setup used to implement this research.

- IDE (Integrated Development Environment): Google Colab - Python 3 GCE (Google Compute Engine)
- Programming Language: Python 3
- Tools/Libraries: TensorFlow, Scikit-learn, Matplotlib, NumPy, Pandas
- System Specifications: As per Table 1

Table 1: System Specifications

Parameter	**Specification**
System Architecture	x64
Operating System	Windows 11
System RAM (GCE)	12.7
Disk (GCE)	107.7

- Code and Dataset: Utilise the code and optimised NAB (Numenta Anomaly Benchmark) dataset to reproduce the results.

2 Evaluation Parameters

This section describes the parameters used to evaluate the performance of the GRU-BERT (Gated Recurrent Unit-Bidirectional Encoder Representations from Transformers) model in detecting anomalies in cloud. The performance of GRU-BERT was analysed based on performance metrics such as cost, computational complexity, execution time, and energy consumption.

- Cost: Calculated using the value of the validation loss function (Binary Cross Entropy) that measures how well the prediction of the model matches with the actual target.
- Computational Complexity: Measures the efficiency of the algorithm based on the number of parameters used in the model, thus calculating the computational complexity of the model.

- Execution Time: Refers to the amount of time taken to train the model.
- Energy Consumption: Measured in terms of the power consumption based on CPU usage and execution time.

The performance of the GRU-BERT model was also analysed based on standard performance metrics which can be calculated by equations (1)(2)(3)(4) respectively for Accuracy, Precision, Recall and F1 score.

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Additionally, the performance of GRU-BERT was evaluated based on hyperparameters such as units, learning rate, dropout rate, and batch size.

- Units: Describes the number of GRU units, also known as hidden states.
- Learning Rate: Defines the size of the steps the model takes to update its weights.
- Dropout Rate: Determines the number of neurons to be dropped while training, to decrease overfitting.
- Batch Size: Total number of training samples that are processed before the weights are updated by the model.

3 Experiments Performed

In this section, the experiments performed using the evaluation parameters mentioned in section 2 are described. This involves the optimisation of hyperparameters in GRU-BERT, comparison of GRU-BERT and LSTM (Long Short-Term Memory) models based on performance metrics, standard performance metrics, and hyperparameters mentioned in section 2. Visualisation techniques were also utilised by making use of Matplotlib, for better analysis and interpretation of results.

- Optimisation of Hyperparameters in GRU-BERT: Bayesian Optimisation was performed using HyperOpt library in Python. This is an automated way of finding out the best hyperparameters that can be used to build and train the model. Initially, more than 20 hyperparameter combinations were searched. Each combination was evaluated on a validation set, thus optimising the validation loss. The best parameters were extracted and used to train the model for 20 epochs. The model was evaluated again based on the performance metrics.

- Comparison of GRU-BERT with LSTM: A comparison between GRU-BERT and LSTM was done based on the evaluation parameters mentioned in section 2, resulting in insightful conclusions.
  - Validation loss (BCE) of both the models was calculated to measure the cost.
  - The computational complexity of both the models was measured based on the number of parameters in the models.
  - The execution time was measured by making use of the time module in Python and calculating the time taken to train the models.
  - The energy consumption of GRU-BERT and LSTM was measured as a function of power and execution time, which was calculated by making use of the time and psutil modules in Python.
  - The metrics module in sklearn was made use of to calculate the standard performance metrics of both the models and compare them.
  - The values of the best hyperparameters in GRU-BERT and LSTM, calculated by using the Hyperopt library was made use of to compare and analyse the performance of the models.