Assessment 2: Neural Network Classification

Due: ***11:59 PM AEST Friday week 5***

Weight: ***40%***

Overview

In this assessment, you will produce a written report on building and evaluating a neural network model for predicting hospital readmission of diabetes patients. You will use Python and Jupyter Notebooks on a platform of your choice (e.g., Google Colab, local Jupyter, etc.).

Learning outcomes

1. Analyse real-world tasks using machine learning methods, particularly supervised learning for classification.
2. Develop and evaluate neural network models using Python.
3. Tune hyperparameters and apply regularisation techniques to improve model performance.
4. Programmatically interact with data and ML libraries using Python.
5. Communicate your methodology and findings effectively.

Format

You will need to submit the following:

- A PDF file clearly shows the assignment question, the associated answers, relevant Python outputs, analyses and discussions.
- Appendices include Jupyter Notebook/ Python code.
- The assignment **should not exceed 15-A4 pages.** Appendices do not form part of the page limit. • The task cover sheet

You have up to three attempts to submit your assessment, and only the last submission will be graded.

A word on plagiarism

Plagiarism is the act of using another's words, works or ideas from any source as one's own.
Plagiarism has no place in a University. Student work containing plagiarised material will be subject to formal university processes.

Background

Understanding the possibility of readmission of diabetes patients may provide hospitals and doctors insight into effectiveness of ongoing treatments, and potential changes in the treatments. The changes in treatments may potentially save patients' life.

You are asked to investigate if a neural network accurately classifies if a patient with diabetes is likely to be readmitted to the hospital given their current conditions.

Data

The data, **"diabetic_data.csv"**, contains 10 years (1999-2008) of clinical care at 130 U.S. hospitals and integrated delivery networks. Data description is provided in Table 1, *"Impact of HbA1c Measurement on Hospital Readmission Rates: Analysis of 70,000 Clinical Database Patient Records"* BioMed Research International, vol. 2014, Article ID 781670, https://www.hindawi.com/journals/bmri/2014/781670/

Assessment Tasks

1. Data preparation **(2 marks)**

a) Discuss a subset of relevant predictors used in a NN model
b) Apply and Discuss any appropriate cleaning or transformations
c) Apply and discuss the training and testing dataset.

2. Build, train and deploy a neural network **(14 marks)**

a) Propose a **baseline** structure of a multiple layer perception (MLP) neural network used for the classification. Justify your choice.
b) Report and discuss the performance of the proposed model.
c) Apply techniques such as dropout, early stop, batch normalisation to the baseline MLP (in (a)), and investigate their impacts on the performance of MLP.
d) Discuss limitations of the proposed model for the classification task.

3. Discuss the considerations of the Platform: **(4 marks)**

a) Describe the platform you used (e.g., Google Colab, JupyterLab).
b) Discuss the computational resources used (e.g., CPU/GPU, RAM).
c) Reflect on the cost (if any) and time efficiency of your chosen platform.

Marking Criteria and Rubric: MA5852 Assessment 2

Criteria	High Distinction	Pass	Fail
Prepare data for analysis	Demonstrate superior ability to clean and prepare data for ML and NN analysis. Highly developed awareness of the processes needed to prepare data for analysis	Demonstrate limited ability to clean and prepare data for ML and NN analysis. Limited awareness of the processes needed to prepare data.	Demonstrate poor or no ability to clean and prepare data for ML and NN analysis. No awareness of the processes needed to prepare data.
Build, train and deploy neural network	Demonstrate superior ability to justify and design a structure of MLP, elaborate choices of hyperparameters to train MLP. Demonstrate superior ability to logically arrange, present and communicate the information of analysis and comparison.	Demonstrate limited ability to design a NN, the training and testing dataset, elaborate choices of hyperparameters to train MLP. Demonstrate limited ability to logically arrange, present and communicate the information of analysis and comparison.	Demonstrate poor or no ability to design a NN, investigate and select hyperparameters for NN to analyse data. Poor to no ability to logically arrange, present and communicate the information of analysis and comparison.
Regularisation techniques	Provide detailed discussion regarding impacts of regularisation approaches and batch normalisation on the performance of the model.	Provide limited discussion regarding impacts of regularisation approaches and batch normalisation on the performance of the model.	Provide poor discussion regarding impacts of regularisation approaches and batch normalisation on the performance of the model.

	Provide sound suggestions to improve the performance of the model.	Provide some suggestions to improve the performance of the model.	Provide some suggestions to improve the performance of the model.
Implementing NN on a Platform	Demonstrate superior ability to build, train and deploy NN in any platform. Provide full evidence of successfully training, deploying and creating endpoints of all models in the selected platform.	Limited ability to build, train and deploy NN in any platform. Provide some evidence of successfully training, deploying and creating endpoints of all models in the selected platform.	Poor to no ability to build, train and deploy NN in any platform.