

**National Tsing Hua University**  
**11220IEEM 513600**  
**Deep Learning and Industrial Applications**  
**Homework 2**

Name:楊佳蓁

Student ID:109034048

**Due on 2024.03.21**

1. (20 pts) Select 2 hyper-parameters of the artificial neural network used in Lab 2, and set 3 different values for each. Perform experiments to compare the effects of varying these hyper-parameters on the loss and accuracy metrics across the training, validation, and test datasets. Present your findings with appropriate tables.

我選擇Epoch以及Learning rate作為要調整並測試的參數，以下表格為改變這兩項參數值之後，分別在training、validation以及 test datasets的表現（accuracy和loss）。

(1)Epoch：20/50/100三個level比較

Epoch		Training datasets	Validation datasets	Test datasets
20	accuracy	73.0159%	67.9012%	64.5161%
	loss	0.5215	0.6710	0.8076
50	accuracy	76.7196%	70.3704%	67.7419%
	loss	0.4360	0.5928	0.6009
100	accuracy	86.7725%	77.7778%	64.5161%
	loss	0.3273	0.5367	0.5376

(2)Learning rate：0.005/0.001/0.0001三個level比較

Learning rate		Training datasets	Validation datasets	Test datasets
0.005	accuracy	86.2434%	85.1852%	77.4194%
	loss	0.3632	0.3925	0.5673
0.001	accuracy	80.9524%	85.1852%	64.5161%
	loss	0.4218	0.5376	0.6043
0.0001	accuracy	75.1323%	75.3086%	61.2903%
	loss	0.5040	0.5842	0.6897

2. (20 pts) Based on your experiments in Question 1, analyze the outcomes. What differences do you observe with the changes in hyper-parameters? Discuss whether these adjustments contributed to improvements in model performance, you can use plots to support your points. (Approximately 100 words.)

(1) Epoch在我選擇的三個level中，當epoch越大，training、validation以及test set的accuracy皆會越來越大，且loss都會越來越小。然而，當epoch過大時，可能會讓運算負擔變大，且在結果收斂後造成多餘的資源浪費。

(2) Learning rate在我選擇的三個level中，當learning rate越大，training、validation以及test set的accuracy皆會越來越大，且loss都會越來越小。不過因為我選擇的三個level都在合理範圍內，若是learning rate太大或太小都可能會對model的表現造成負面影響。

3. (20 pts) In Lab 2, you may have noticed a discrepancy in accuracy between the training and test datasets. What do you think causes this occurrence? Discuss potential reasons for the gap in accuracy. (Approximately 100 words.)

在一個just right的model之下，training set的accuracy一般而言都會較test set高。因為model就是用training set訓練（讓training set的loss最小化）出來的結果，而test set是用來評估model一般化到new dataset的表現，用以反映model應用於真實資料的情況。

4. (20 pts) Discuss methodologies for selecting relevant features in a tabular dataset for machine learning models. Highlight the importance of feature selection and how it can impact model performance. You are encouraged to consult external resources to support your arguments. Please cite any sources you refer to. (Approximately 100 words, , excluding reference.)

一般而言，多數feature可能與想預測的目標沒有關係，常用來實現feature selection的方法例如filter、計算information gain等。Feature selection可以簡化機器學習模型，進而提高generalization的能力並降低overfitting的可能，此外還能降低機器學習模型訓練和測試的計算複雜度。

參考資料：<https://medium.com/@a.mirzaei69/how-to-use-deep-learning-for-feature-selection-python-keras-24a68bef1e33>

<https://www.analyticsvidhya.com/blog/2020/10/feature-selection-techniques-in-machine-learning/>

5. (20 pts) While artificial neural networks (ANNs) are versatile, they may not always be the most efficient choice for handling tabular data. Identify and describe an alternative deep learning model that is better suited for tabular datasets. Explain the rationale behind its design specifically for tabular data, including its key features and advantages. Ensure to reference any external sources you consult. (Approximately 150 words, , excluding reference.)

TabNet採用Sequential Attention的機器學習技術，在訓練模型的過程中利用計算

attention score來選擇feature，並經由動態調整保留真正重要而非多餘無關聯的feature，提升模型的效率。此外，TabNet以類決策樹的結構性方法處理表格資料，獲取決策樹可解釋性的優點，同時其可適用於大規模dataset。

參考資料：<https://www.ithome.com.tw/news/139718>

<https://openreview.net/pdf?id=BylRkAEKDH>