National Tsing Hua University

1130IEEM 513600

Deep Learning and Industrial Applications

Homework 2

Name: 黃允秀 Student ID: 112034701

Due on 2025.03.27

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.

Learning rate	Train Loss	Validation Loss	Test Loss	Train Accuracy	Validation Accuracy	Test Accuracy
0.001	0.427601	0.506727	0.538865	79.365079%	75.308642%	74.193548%
0.01	0.343976	0.360133	0.454755	84.656085%	87.654321%	77.419355%
0.1	0.692822	0.687212	0.694306	51.322751%	61.728395%	48.387097%

Epoch (learning rate = 0.001)	Train Loss	Validation Loss	Test Loss	Train Accuracy	Validation Accuracy	Test Accuracy
50	0.487630	0.549152	0.614950	75.66%	71.60%	64.52%
100	0.373737	0.501521	0.481538	84.66%	81.48%	77.42%
150	0.359785	0.417348	0.460555	84.66%	82.72%	74.19%

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.)

   - Learning rate：當學習率為0.001時，訓練準確度較高，但測試準確度較低，可能導致overfitting。學習率為0.1時，模型表現較差，特別是在測試集上，顯示過高的學習率會導致模型無法收斂。
   - Epoch：隨著epoch數量增加，訓練和驗證損失減少，訓練和驗證準確度提升，表示模型表現更穩定並學得更好。然而，測試集準確度的變化較小，顯示overfitting的可能性。

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.)

   訓練集與測試集的準確度差異，有可能是因為overfitting造成的結果，通常會出現在model過度學習訓練及資料的細節或是雜音，導致在測試集上的泛化能力下降。另外也可能兩者的數據在根本上有分布上的差異，會造成模型在測試集的表現較差。

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.)

- 過濾法（Filter）：基於統計指標評估特徵與目標的關係，如相關係數等，這種方法計算簡單，適合大型表格數據前處理。
- 包裝法（Wrapper）：使用目標機器學習算法評估特徵子集的性能，可達到較高的分類精度，但運算複雜度較高。
- 嵌入法（Embedded）：在模型訓練過程中自動進行特徵選擇，如Lasso回歸通過正則化實現變量選擇。

  Reference: Jie Cai, Jiawei Luo, Shulin Wang, Sheng Yang (2018). Feature selection in machine learning: A new perspective.

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 you to reference any external sources you consult. (Approximately 150 words, excluding reference.)

TabNet：由Google研究團隊設計，具有四大優勢：動態特徵選擇機制、序列化決策步驟、高可解釋性和稀疏學習特性。其注意力機制能在每個決策步驟選擇最相關特徵，提高效率同時減少過擬合。相比傳統神經網絡，TabNet在表格數據集上展現更佳性能、更少參數量和更佳的解釋能力。

Reference: Arik & Pfister (2021), TabNet: Attentive Interpretable Tabular Learning.