National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

Homework 2

Name:吳柏彥 Student ID:112034525

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.

Epochs/ lr	0.1	0.01	0.001
50	Epoch 50/50, Train loss:0.6891, Train acc:54.4974%, Val loss: 0.6866, Val acc: 54.3210% Test accuracy is 48.38709677419355%	Epoch 50/50, Train loss:0.3642, Train acc:86.7725%, Val loss: 0.4056, Val acc: 81.4815% Test accuracy is 64.51612903225806%	Epoch 50/50, Train loss:0.4578, Train acc:76.1905%, Val loss: 0.6240, Val acc: 66.6667% Test accuracy is 64.51612903225806 %
150	Epoch 150/150, Train loss:0.6908, Train acc:53.4392%, Val loss: 0.6845, Val acc: 56.7901% Test accuracy is 48.38709677419355%	Epoch 150/150, Train loss:0.2302, Train acc:89.4180%, Val loss: 0.5309, Val acc: 83.9506% Test accuracy is 80.64516129032258%	Epoch 150/150, Train loss:0.2870, Train acc:89.4180%, Val loss: 0.4590, Val acc: 77.7778% Test accuracy is 77.41935483870968%

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.)

從結果可以發現，lr(學習率)在三種水準下，會有不同的結果，在 lr=0.01 時對整個模型的準確率是較高的，而在 lr=0.1 的時候，會較低的準確率，因為學習的步數較大，所以才會導致準確率較低。而在 Epoch 的這個參數時，我們利用 epoch=50 跟 epoch=150 比較，可以發現 epoch=150 時跑出來的結果，都會比 epoch=50 時還要來的好，由此可知，epoch=50 時可能對於模型來說，跑的數量不夠多，需要再增加才可以達到最優化。

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.)

訓練和測試數據集之間的準確率存在差異。造成這種情況的原因可能有幾個：

1.過擬合（Overfitting）：模型在訓練數據上表現良好，但在測試數據上表現較差，暗示模型可能過度擬合訓練數據。

2.數據不平衡（Imbalanced Data）：訓練集和測試集中的類別分佈可能不一致，導致模型在測試集上表現不佳。

3.數據偏移（Data Shift）：訓練和測試數據集之間的分佈差異，可能導致模型在測試集上的泛化能力下降。

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.)

在表格資料集中，特徵選擇方法包括過濾法、包裝法和嵌入法。適當的特徵選擇有助於降低過度擬合風險，提高模型解釋性和效能。它可以減少計算成本，改善模型泛化能力。

1. 過濾法（Filter Methods）：
   1. 過濾法是在模型訓練之前對特徵進行篩選。它通常基於單變量統計測量，例如相關係數、卡方檢驗或信息增益，以確定特徵與目標之間的相關性。
   2. 這些方法快速且易於實現，但它們忽略了特徵之間的相互關係，可能無法捕捉到複雜的特徵關係。
2. 包裝法（Wrapper Methods）：
   1. 包裝法通過將特徵選擇視為搜索問題，根據特定的機器學習算法（如回歸或分類器）的性能，反覆選擇不同的特徵子集。
   2. 這些方法更費時且計算成本高，但通常可以獲得更好的性能，因為它們考慮了特徵之間的交互作用。
3. 嵌入法（Embedded Methods）：
   1. 嵌入法是將特徵選擇嵌入到機器學習算法的訓練過程中，通過正則化技術（如 Lasso 回歸）或樹模型（如隨機森林或梯度提升樹）來選擇最具信息量的特徵。
   2. 這些方法不僅考慮了特徵的相關性，還考慮了模型本身的複雜性和泛化能力。

[參考文獻: Brownlee, J. (2020). Feature Selection For Machine Learning in Python. Machine Learning Mastery.]

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.)

一個更適合處理表格數據的替代深度學習模型是 TabNet。TabNet 是一種基於注意力機制和遞歸門控單元（GRU）的深度神經網絡模型。它針對表格數據的特點進行了優化設計，通過選擇性地關注特徵和採用自注意力機制，能夠有效處理具有大量特徵和高度關聯性的表格數據。

1. 注意力機制：TabNet 利用自注意力機制選擇性地關注輸入特徵，從而自動學習重要特徵之間的關係。

2. 遞歸門控單元（GRU）：GRU 單元允許模型在處理序列型特徵時保持狀態信息，這對於時間序列或具有序列性質的表格數據非常有用。

3. 可解釋性：TabNet 能夠提供對模型決策的解釋，這對於某些應用場景（如金融和醫療）非常重要。

4. 高效性：TabNet 在訓練過程中具有高效性，並且在處理大型表格數據集時表現良好。

5. 抗過擬合：TabNet 通常能夠有效地處理高度結構化的表格數據，並且具有一定的抗過擬合能力。

[參考文獻: Arik, S. O., & Pfister, T. (2019). TabNet: Attentive Interpretable Tabular Learning. arXiv preprint arXiv:1908.07442.]