National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

Homework 2

Name: 華 Student ID: 112034822 **Due on 2024.03.21**

Q1

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.

epoch	learning rate	train loss	train acc	val loss	val acc	test acc	
100	0.001	0.3841	84.66%	0.7392	79.01%	64.51%	overfitting
80	0.001	0.3729	85.71%	0.6882	79.01%	74.19%	
50	0.001	0.4652	77.78%	0.6546	77.78%	61.29%	underfitting
200	0.001	0.2724	88.89%	0.6077	80.25%	64.51%	overfitting
80	0.01	0.3194	86.24%	0.6175	79.01%	74.19%	
80	0.0001	0.5108	74.60%	0.5982	77.78%	61.29%	underfitting
200	0.0001	0.457	79.37%	0.6118	81.48%	67.74%	

Q2

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.)

While learning rate equal 0.001, epoch over 100 will cause overfitting, and will underfitting under 50.

Set epoch as 80, let learning rate be 0.01 might not strongly influence test accuracy, but while let learning rate be 0.0001, model will be underfitting.

Epoch and learning will contribute to improvements in model performance with avoiding underfitting or overfitting but still can't make a correction on the bias of the model, maybe we should adjust the layer's size or preprocessing and select the features of the data.

Q3

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.)

I think is like the problem I have mentioned in Q2, which the model is overfitting. The model overfit with the training set and cause the inaccuracy with testing set. More precisely, model can predict the value in training set nicely, but will have a trouble while predicting test dataset.

Q4

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.)

Feature selection is crucial in machine learning to improve model performance by removing irrelevant or redundant data, enhancing computational efficiency, and reducing overfitting risk.

There have three main different ways to select features nowadays, that is filter methods, wrapper method and embedded methods.

Filter methods: assess feature relevance through statistical measures to select features.

Wrapper methods: evaluate subsets of features based on model performance.

Embedded methods: incorporate feature selection as part of the model training process, using techniques like LASSO and Ridge Regression.

(refer to net)

Q5

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.)

I will describe an alternative deep learning model TabNet, TabNet is designed for tabular data, surpass traditional ANNs by employing sparse attention in decision-making steps, enabling it to focus on relevant features for each prediction. This selective attention mechanism helps TabNet to understand complex inter-feature relationships and dependencies, which are common in tabular data but challenging for ANNs to capture due to their dense and generalized feature processing approach. By learning to choose which features to emphasize at each step, TabNet dynamically adapts to the data structure, improving interpretability and efficiency in handling diverse tabular datasets.

Reference: TabNet: Attentive Interpretable Tabular Learning (Sercan O. Arık, Tomas Pfister)