National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

Homework 2

Name: 林芮伃 Student ID: 112030509

Due on 2024.03.21

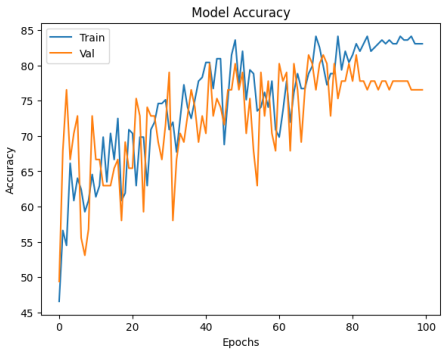
1. (20 pts) Select 2 hyper-parameters of the artificial neural network used in Lab 2, and set 3 different values for each. Perform experiments to compare the effects of varying these hyper-parameters on the loss and accuracy metrics across the training, validation, and test datasets. Present your findings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

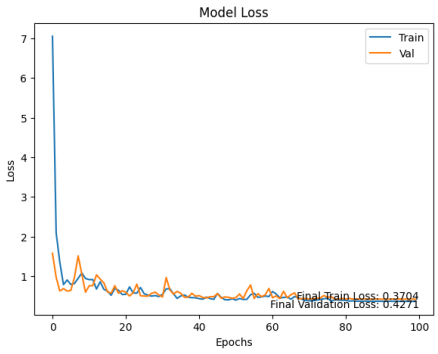
	Batch size	Epochs	Train loss	validation loss	Accuracy
0(original)	32	100	0.3704	0.4271	74.19%
Change the value of the epochs only					
1	32	10	0.6885	0.6986	58.06%
2	32	1000	0.9111	0.0854	80.66%
3	32	10000	11.453	0.0	80.65%
Change the value of the batch size only					
4	16	100	0.3523	0.3743	74.19%
5	64	100	0.4736	0.5428	64.52%
6	256	100	0.505	0.6677	61.29%

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.)

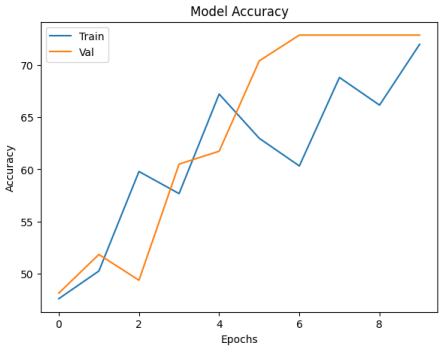
   Based on experiments, adjusting Epochs impacts model performance. Smaller values may cause underfitting, while larger values improve accuracy. However, excessively large Epochs may not enhance performance but increase training time, risking overfitting. Modifying Batch size influences convergence speed, but excessively large sizes may lead to overfitting without guaranteeing improved accuracy. Striking a balance between speed, performance, and overfitting concerns is crucial when adjusting these hyperparameters.

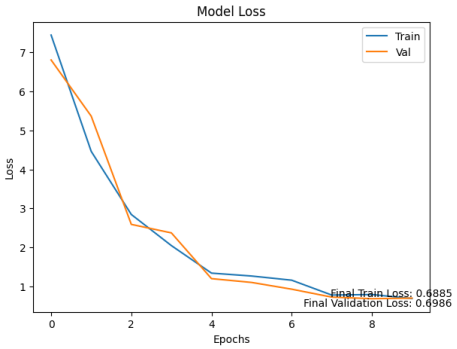
0(original)



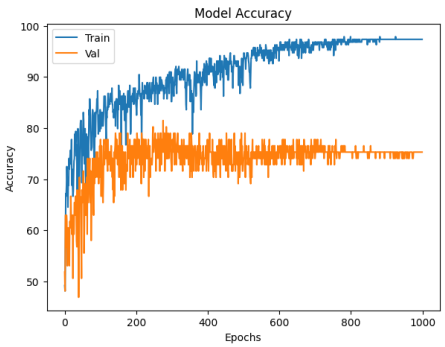


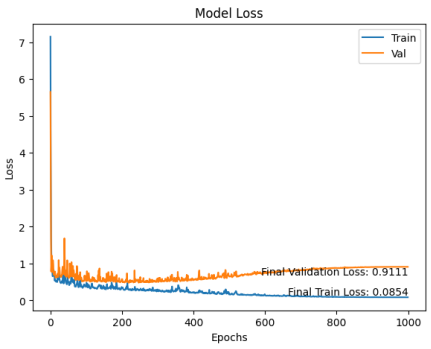
1



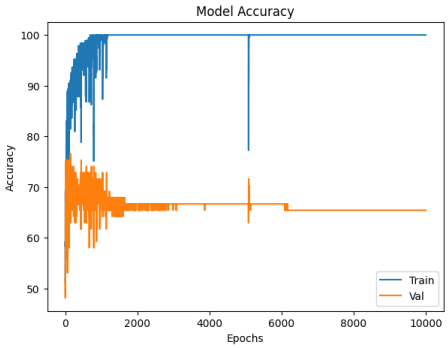


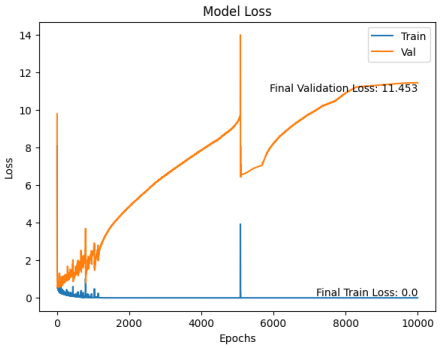
2



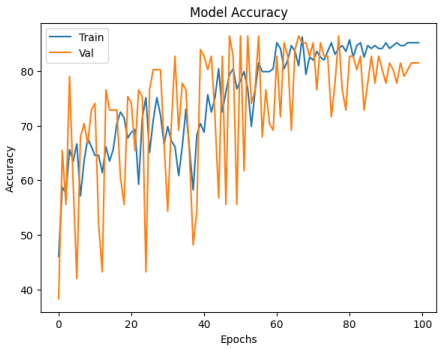


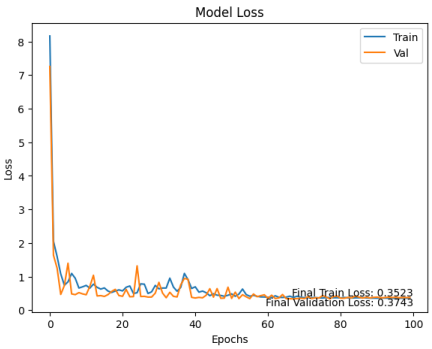
3



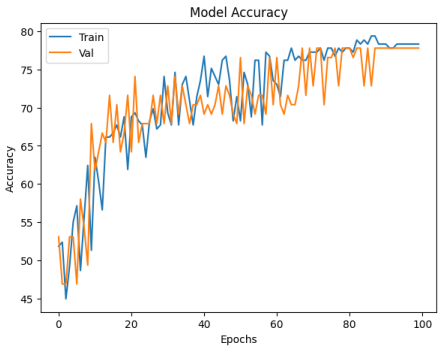


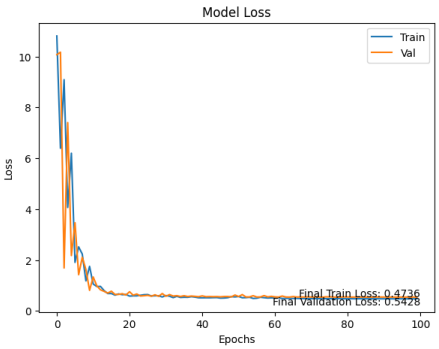
4



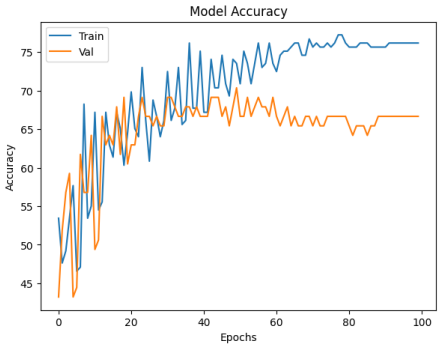


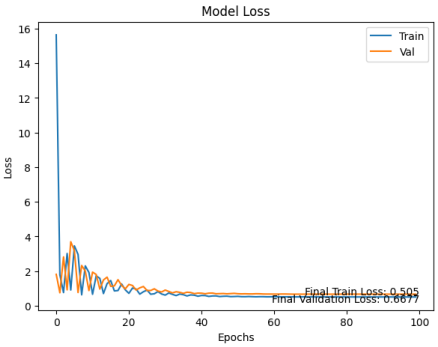
5





6





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.)

   In Lab 2, the disparity in accuracy between the training and testing datasets may be attributed to overfitting. Overfitting occurs when the model overly adapts to the training data, resulting in poorer performance on unseen testing data. This could be due to the model learning noise or patterns specific to the training set that do not generalize well to other datasets.

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 Methods:

   1) Correlation Analysis: Evaluate the contribution of features to a task by calculating the correlation between features and the target variable.
   2) Embedded Methods: Some machine learning models incorporate embedded feature selection, such as L1 regularization, which sets unimportant weights to zero.
   3) Tree Model Feature Importance: Utilize tree models (e.g., decision trees or random

forests) to compute the importance of features, providing a model-based analysis.

4) Recursive Feature Elimination (RFE): Iteratively train the model and eliminate the least impactful features on performance until the desired number of features is achieved.

Choosing relevant features helps reduce the complexity of the model, preventing overfitting, and enhancing generalization ability. Simultaneously, it lowers computational costs for training and prediction, particularly crucial for large-scale datasets. This approach also fosters interpretability, making it easier to understand and comprehend the contributions of features to predictions.

Reference: Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research, 3, 1157-1182.

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.)

While artificial neural networks (ANNs) are versatile, they may not always be the most efficient choice for handling tabular data. An alternative deep learning model better suited for tabular datasets is TabNet (Tabular Attention Network). Developed by the Google Cloud AI team, TabNet is specifically designed for structured data. Its key feature is the use of selective attention mechanisms, allowing it to adaptively focus on the most important features in the table while maintaining interpretability. TabNet excels in processing tabular data, demonstrating superior performance with lower computational costs. This makes it particularly suitable for handling large volumes of structured data in fields such as finance and healthcare.

Reference: https://arxiv.org/abs/1908.07442