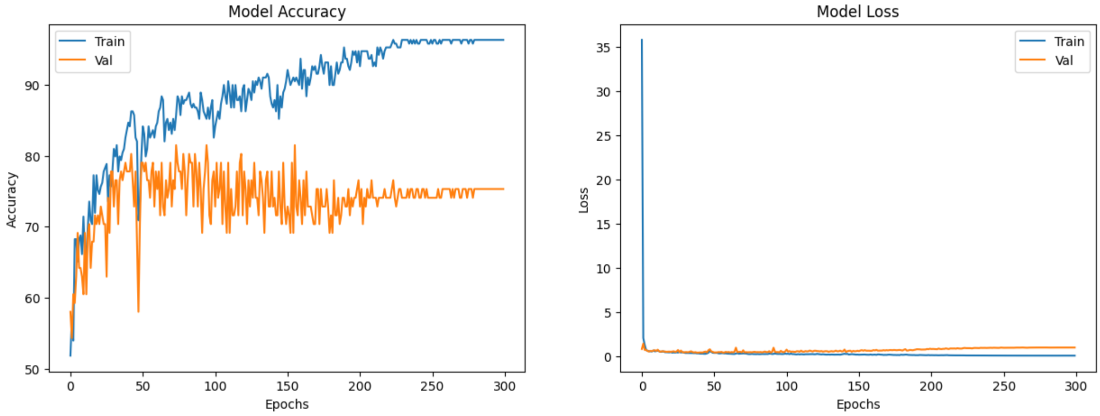
Deep Learning and Industrial Applications | Homework 2

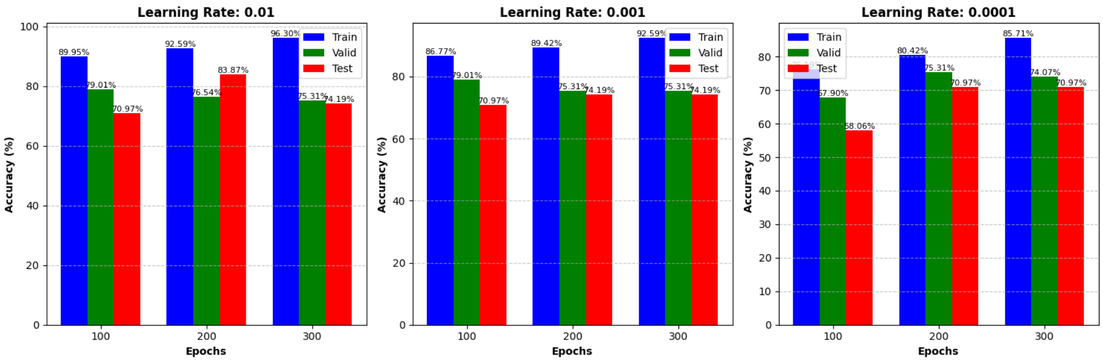
黃冠毓 113034512

1. Choose the learning rate and the epoch.

Learning rate	Epoch	Training accuracy	Training loss	Validation accuracy	Validation loss	Test accuracy	Test loss
0.01	100	89.95%	0.2729	79.01%	0.4941	70.97%	0.4597
0.01	200	92.59%	0.1944	76.54%	0.6515	83.87%	0.3955
0.01	300	96.30%	0.1173	75.31%	1.0342	74.19%	0.3976
0.001	100	86.77%	0.3484	79.01%	0.4613	70.97%	0.5350
0.001	200	89.42%	0.2699	75.31%	0.4820	74.19%	0.5347
0.001	300	92.59%	0.2196	75.31%	0.5658	74.19%	0.4991
0.0001	100	76.19%	0.4826	67.90%	0.5888	58.06%	0.6807
0.0001	200	80.42%	0.4239	75.31%	0.5690	70.97%	0.5492
0.0001	300	85.71%	0.3733	74.07%	0.5658	70.97%	0.5471

2. Based on the results of Question 1, the learning rate and the number of training epochs have a significant impact on the model's performance. A higher learning rate allows the model to train faster and achieve higher accuracy in fewer epochs, but the validation and test accuracies are lower, likely due to overfitting. Increasing the number of training epochs does not improve this issue and instead exacerbates the overfitting problem. A lower learning rate results in slower convergence, and as the number of epochs increases, overfitting also occurs. Therefore, choosing an appropriate learning rate and number of epochs is crucial for achieving stable performance and effectively improving accuracy on the validation and test datasets.
   The following image represents the model's performance with a **learning rate of 0.01** and trained for **300 epochs**.





3. The discrepancy in accuracy between the training and test datasets is likely caused by overfitting. During training, the model learns to fit the training data well, which leads to high training accuracy. However, it may not generalize effectively to unseen test data, resulting in lower test accuracy. Overfitting occurs when the model becomes too complex or is trained for too many epochs, memorizing the training data instead of learning general patterns. Additionally, differences in the distribution or noise between the training and test datasets can also contribute to this accuracy gap. Regularization techniques and early stopping can help mitigate overfitting.
4. Feature selection is a crucial step in machine learning that involves identifying the most relevant input variables to improve model performance. It can be done using methods such as filter (statistical tests like correlation or chi-square), wrapper (recursive feature elimination), and embedded (feature importance from algorithms like decision trees). Effective feature selection can reduce overfitting, improve accuracy, and decrease computational complexity. Irrelevant features can introduce noise, leading to poorer model generalization. Methods like L1 regularization (Lasso) can also be used to automatically penalize unimportant features, ensuring better model interpretability.

   Reference:

   ChatGPT and「Guyon, I., & Elisseeff, A. (2003). An Introduction to Feature Selection. Journal of Machine Learning Research, 3, 1157–1182」
5. One alternative deep learning model that is well-suited for tabular data is the TabNet model, introduced by Google Cloud AI researchers. Unlike traditional ANNs, which often struggle with high-dimensional tabular data due to their inability to interpret feature interactions effectively, TabNet leverages attention mechanisms to select relevant features dynamically, making it particularly effective for tabular data. It uses sparse attention to focus on important features and temporal dependencies, which reduces the risk of overfitting and improves generalization. Additionally, TabNet incorporates decision trees within its architecture, enhancing interpretability and offering insights into which features

contribute to predictions. This combination of deep learning with tree-like structures allows TabNet to outperform standard neural networks on tabular tasks like classification and regression. Its ability to handle sparse data efficiently and scale with larger datasets makes it a promising choice for practical applications in tabular datasets.

Reference:

ChatGPT and「Arik, S. Ö., & Pfister, T. (2021). TabNet: Attentive Interpretable Tabular Learning. Proceedings of the 37th International Conference on Machine Learning, 2021」