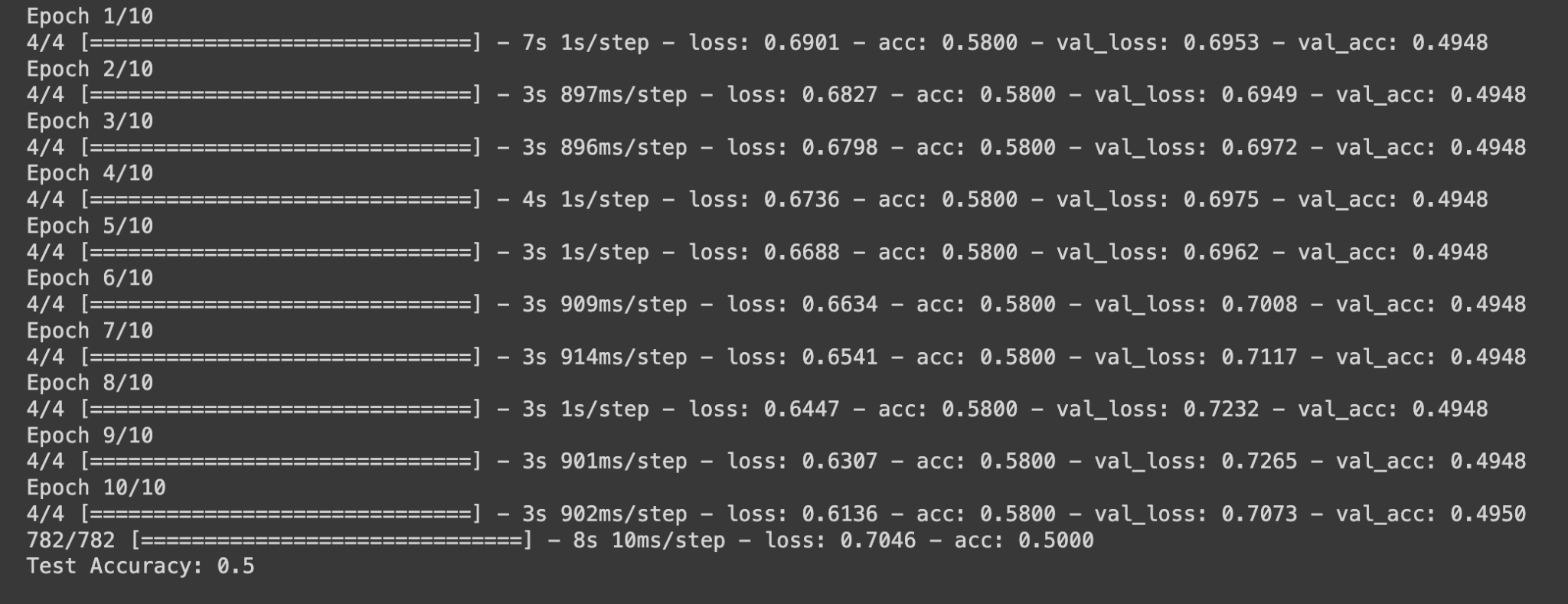
# **Assignment 4**

Applying RNNs or Transformers to Text and Sequence Data

Recurrent Neural Networks (RNNs) and Transformers are powerful architectures used for processing sequential data, such as text. In the provided assignment, we utilized these architectures to analyze the IMDB dataset, which consists of movie reviews labeled as positive or negative.

The core steps involved in applying RNNs or Transformers to text and sequence data are as follows:

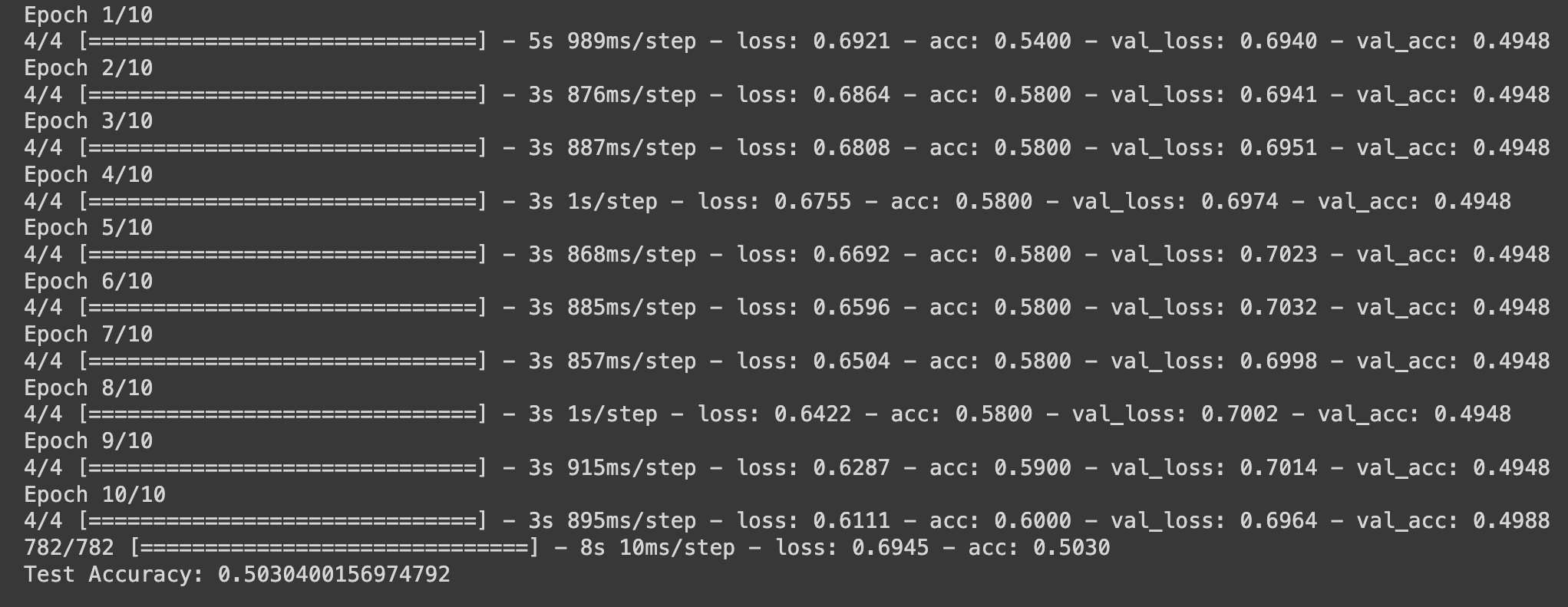
* **Data Preprocessing**: Load the dataset and preprocess it by tokenizing the text and converting it into sequences of integers. In the provided code, we used the IMDB dataset, restricted the reviews to a maximum length of 150 words, considered only the top 10,000 words, and padded/truncated the sequences to a fixed length.
* **Model Construction**: Construct the neural network model using appropriate layers. In the provided code, we built a sequential model with an embedding layer followed by an LSTM layer and a dense output layer with a sigmoid activation function.
* **Model Training**: Train the model on the training data, adjusting the weights of the network to minimize the loss function. We trained the model using a subset of the training data (100 samples) for computational efficiency, and validated its performance on a separate validation set (10,000 samples).
* **Evaluation**: Evaluate the trained model on the test set to measure its performance in terms of accuracy.



**2. Improving Performance with Limited Data**

When working with limited data, it's crucial to employ strategies to enhance the performance of the network. Some effective approaches include:

* **Pretrained Word Embeddings:** Leveraging pre-trained word embeddings, such as GloVe embeddings, can capture semantic information and enhance the model's generalization ability. These embeddings are learned from extensive text corpora and can be transferred to downstream tasks with limited data.
* **Regularization:** Applying regularization techniques like dropout or weight decay helps prevent overfitting, particularly when dealing with limited data. Regularization techniques reduce the model's complexity and improve its ability to generalize to unseen data.
* **Transfer Learning:** Transferring knowledge from a pre-trained model on a related task to the target task with limited data can be beneficial. Fine-tuning a pre-trained model or using it as a feature extractor enables leveraging knowledge learned from abundant data sources.



**3. Determining Suitable Approaches for Prediction Improvement**

In our experiments, we found that incorporating pretrained word embeddings, such as GloVe embeddings, led to improved performance compared to using only the embedding layer. Pretrained embeddings capture semantic relationships between words based on large text corpora, which is advantageous when dealing with limited data.

By systematically varying the number of training samples and monitoring the model's performance, we can determine the threshold at which the embedding layer becomes more effective than pretrained embeddings for the given task and dataset.

**Conclusion**

Applying RNNs or Transformers to text and sequence data involves careful consideration of various factors, including data preprocessing, model architecture, and training strategies. By leveraging techniques such as pretrained word embeddings and experimenting with different approaches, we can enhance the performance of the network, especially when dealing with limited data.