

Report

March 8, 2025

1 Homework 3 Report

1.1 Code Explanation

1. This solution implemented a GRU-based RNN for binary sentiment classification. The model uses a BPE tokenizer to pre-tokenize text reviews, which are then converted into sequences of token IDs. For each review, if the token sequence is longer than 100 tokens, a random contiguous sub-sequence is selected; if it's shorter, it's padded with zeros. An Embedding layer maps these token IDs into 100-dimensional vectors, and a 3-layer GRU (with a hidden size of 100) processes the sequence. The final hidden state is passed to a fully connected layer that outputs a single logit, which is converted to a binary sentiment prediction (positive/negative) via a threshold. For the design choices, I chose to have pre-tokenization to speed up the training to avoid repeating tokenization.
2. For external sources that were used was ChatGPT for any errors that I was not able to debug (making training faster).

1.2 Discussion

1. Compare the results of bag-of-words and RNN in terms of accuracy.
 - The test accuracy of the bag of words model was 0.8654 while the test accuracy of the RNN model was 0.8333. The bag of words model performed better than the RNN model. This is likely because the bag of words model is simpler and more interpretable, while the RNN model is more complex and required more time to train.