LSTM & Bi-LSTM

Programming Assignment 02 : Health Tweet Classification

(S19355)

This assignment focuses on classifying tweets as Personal Health Mentions (PHM) or Non-Personal Health Mentions (Non-PHM) using Recurrent Neural Network models. Two deep learning models were developed for this classification task,

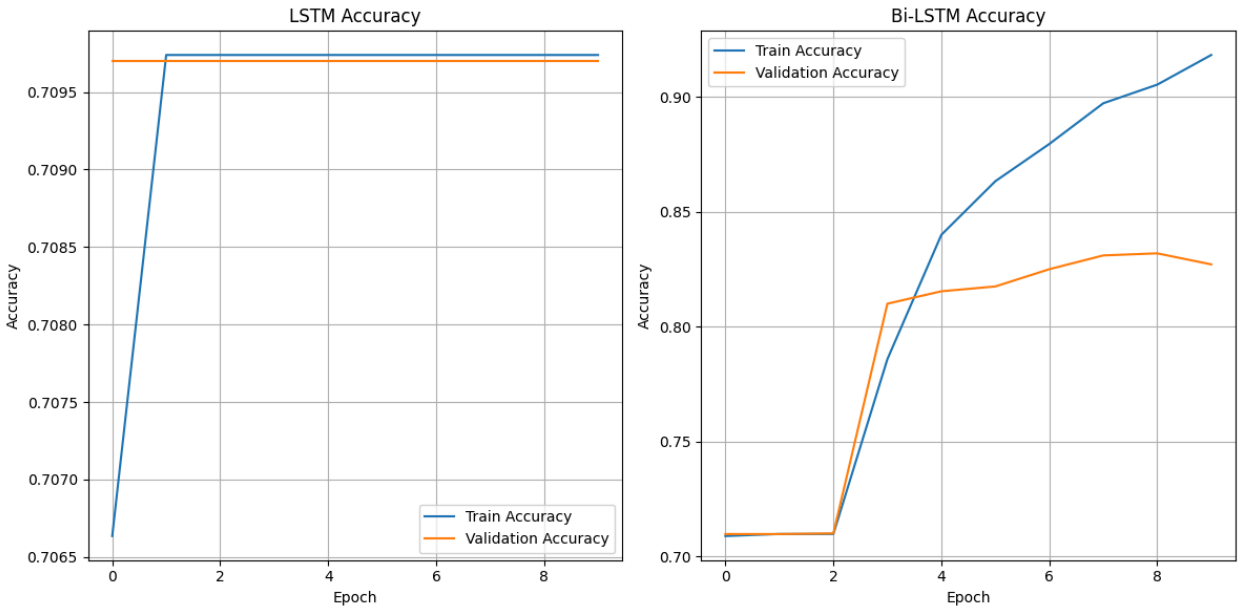
1. LSTM (Long Short-Term Memory)
2. Bi-LSTM (Bidirectional Long Short-Term Memory)

a. Performance comparison (e.g., accuracy and loss) table/plots of the two models.

1. Accuracy and Loss Table

Model	Train Accuracy	Train Loss	Validation Accuracy	Validation Loss
LSTM	70.97%	60.29%	70.97%	60.27%
Bi-LSTM	91.82%	23.85%	82.71%	45.11%

2. Accuracy Graphs



b. Discussion on the performance of the two models.

- Initially, both the LSTM model (similar to the IMDB review classification implementation) and the standard Bi-LSTM model achieved testing accuracies in the range of 60% to 70%. This baseline performance was reasonable but insufficient for robust tweet classification in a real-world health-related context.

- To enhance the performance, hyperparameter tuning was applied to both models:

  For the LSTM model:

  - The maxlen parameter for padding sequences was increased to 250, allowing the model to capture longer tweet contexts.
  - The dropout rate was increased to 0.3 to prevent overfitting.
  - The batch size was decreased to 16, allowing more frequent weight updates.
  - The number of epochs was increased to 10, giving the model more time to learn.

  For the Bi-LSTM model:

  - The second LSTM layer was changed from Bidirectional(LSTM(64)) to a smaller Bidirectional(LSTM(32)) for better regularization and efficiency.
  - The dropout rate was raised to 0.5, which helped control overfitting.
  - The batch size was decreased to 128, balancing training stability and speed.

- As a result of these tuning efforts:

  - The Bi-LSTM model achieved a training accuracy of 91.82% and a validation accuracy of 82.71%, with lower loss values indicating better generalization.

  - The LSTM model improved to 70.97% accuracy, but it still lagged behind Bi-LSTM in terms of both performance and robustness.

- Overall, the Bi-LSTM architecture proved more effective, likely due to its ability to capture both forward and backward dependencies in the tweet sequences. This bidirectional context is especially beneficial in understanding the nuanced and informal language commonly used in tweets discussing health. Therefore, the Bi-LSTM model is the recommended choice for this task.