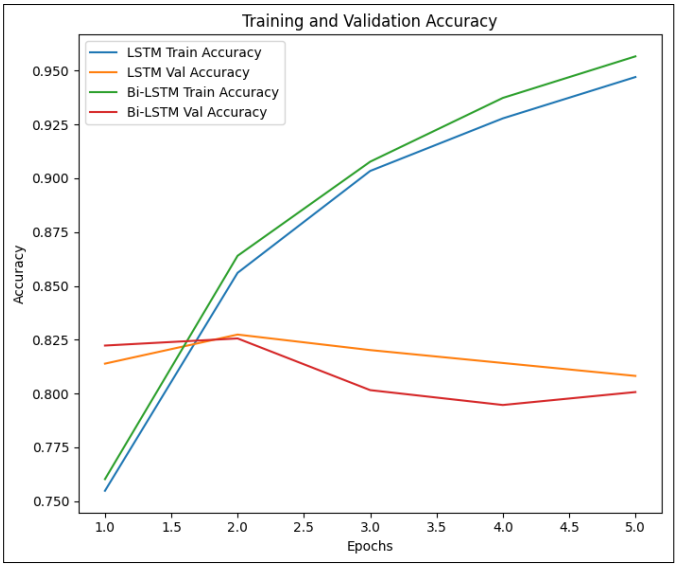
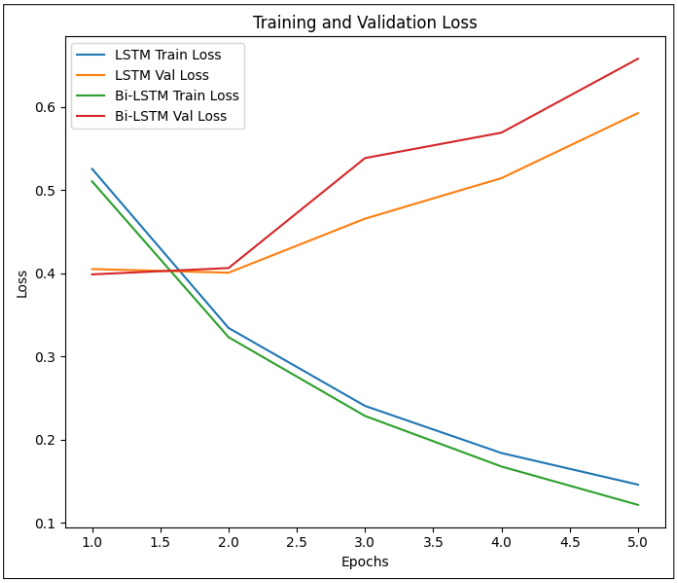
Performance Comparison of LSTM and Bi-LSTM Models for Tweet Health Mention Classification

Performance comparison

Model	Correct Predictions	Wrong Predictions	Accuracy
LSTM	2692	639	80.82%
Bi-LSTM	2667	664	80.07%





Discussion on the performance of the two models

Accuracy and Loss

- Training Accuracy: Both models show steadily increasing training accuracy, with Bi-LSTM slightly ahead of LSTM throughout training.
- Validation Accuracy: Both models reach their peak validation accuracy early (around epoch 2), after which validation accuracy plateaus or slightly declines. LSTM maintains a marginally higher validation accuracy than Bi-LSTM.
- Training Loss: Both models demonstrate steadily decreasing training loss, indicating effective learning on the training set.
- Validation Loss: Validation loss for both models increases after epoch 2, suggesting the onset of overfitting.

Performance

- LSTM Model:
    - Correct Predictions: 2692
    - Wrong Predictions: 639
    - Test Accuracy: 80.82%
- Bi-LSTM Model:
    - Correct Predictions: 2667
    - Wrong Predictions: 664
    - Test Accuracy: 80.07%

The LSTM model slightly outperforms the Bi-LSTM model on the test set in this experiment, both in terms of correct predictions and overall accuracy.

Model Behavior and Interpretation

- LSTM learns from past to future. It usually trains faster and reaches stable performance (equilibrium) quickly.

- Bi-LSTM learns from both past and future context at each step, which can help capture more information from the data and sometimes improve prediction accuracy, especially for text.
- In practice, Bi-LSTM may require more training and smaller batch sizes to tune its parameters well, and it can take longer to reach equilibrium compared to LSTM.
- In my results, both models performed similarly, with LSTM slightly ahead in accuracy. This suggests that for this dataset, learning from both directions did not yield a significant advantage, possibly due to the nature or size of the data.