

# **Math 179 Final Project**

**Rui Zhang**

**Apr 29, 2024**

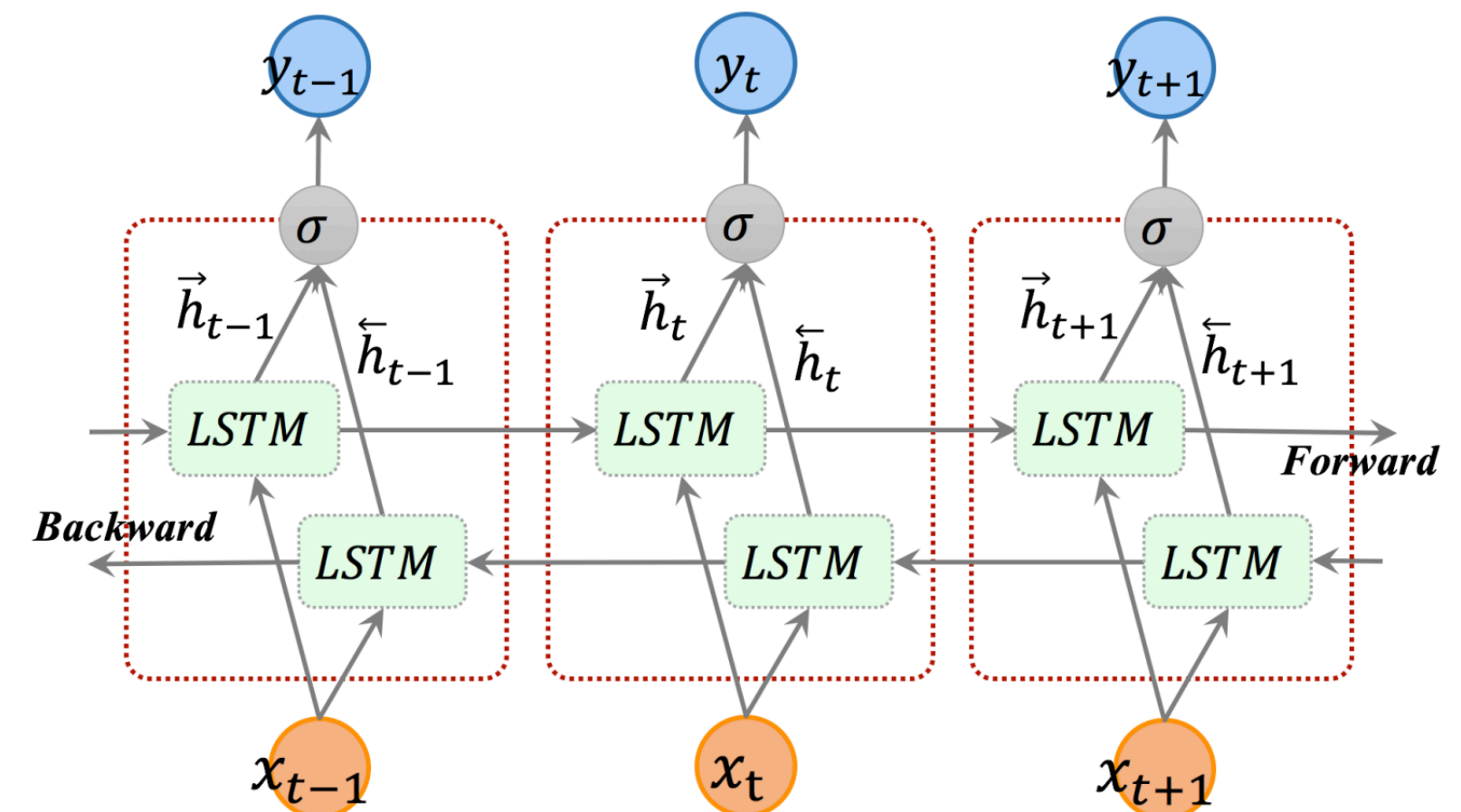
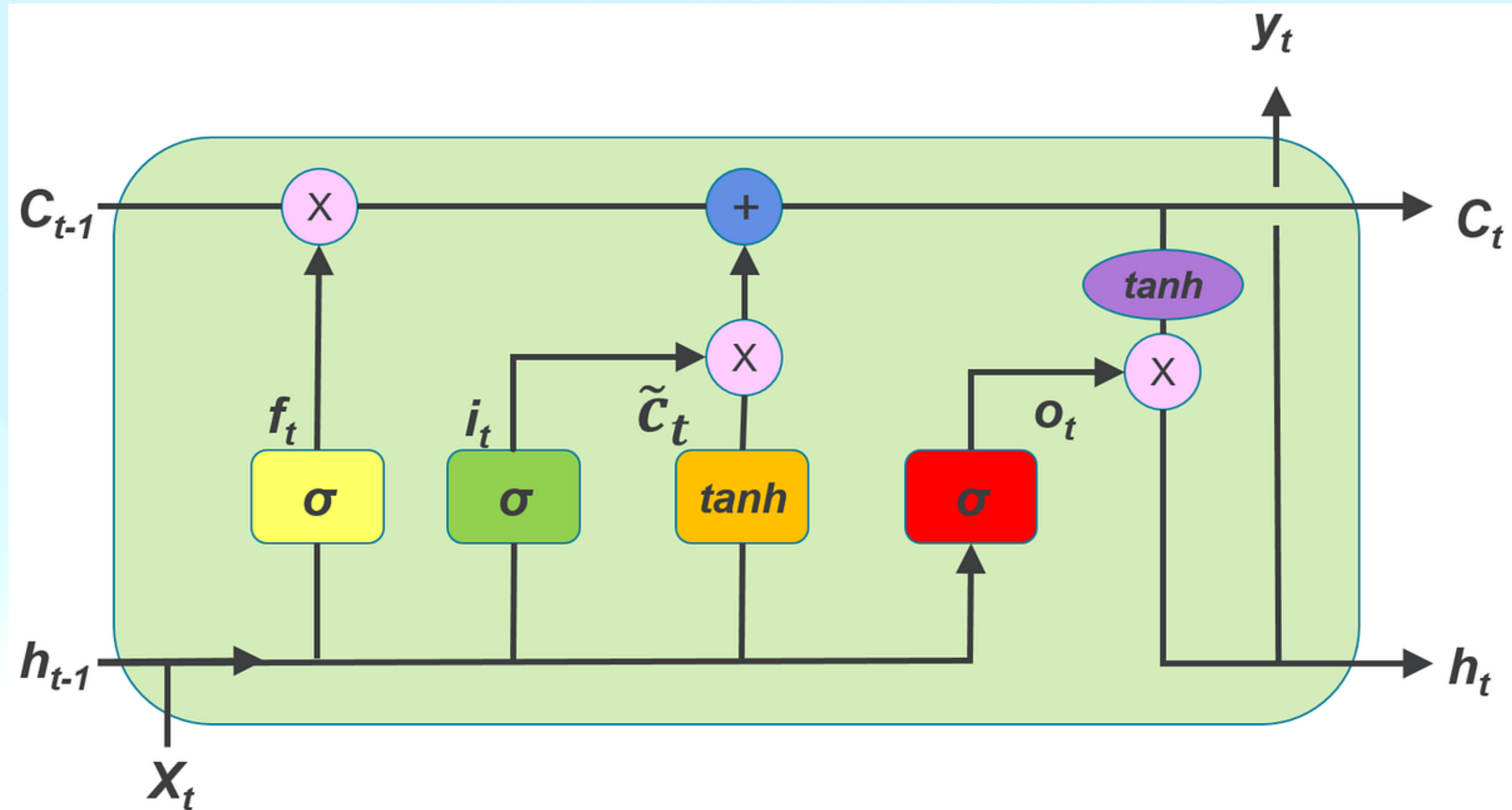
# Outline

- Model
- Exploration
- Data
- Experiment & Results
- Conclusion



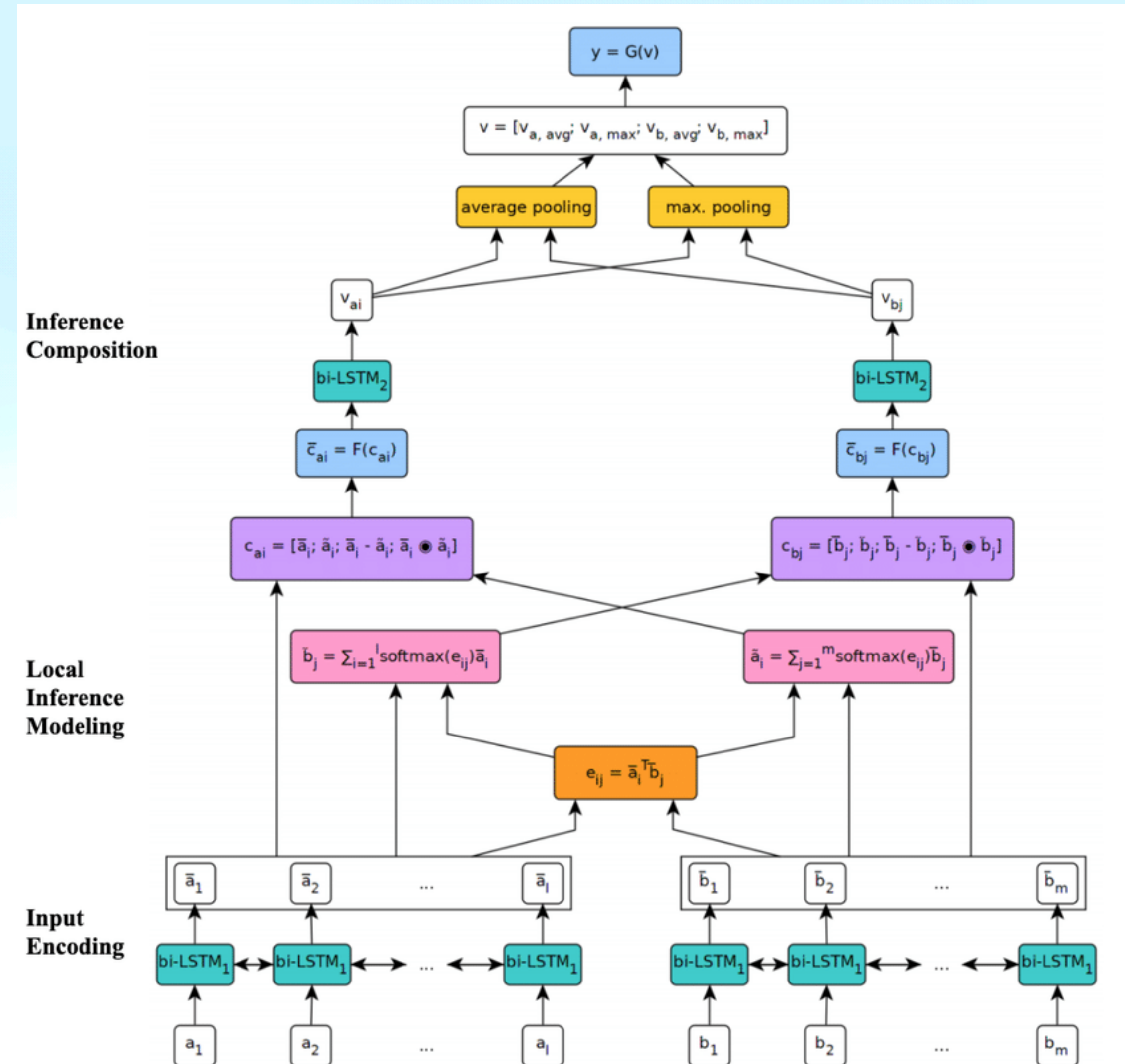
# LSTM

- Long Short-Term Memory
- A type of RNN
- Forget Gate, Input Gate, and Output Gate
- BiLSTM



# ESIM

- Enhanced Sequential Inference Model
- Input Encoding with BiLSTM
- Local inference with attention
- Inference Composition with BiLSTM, average and max pooling



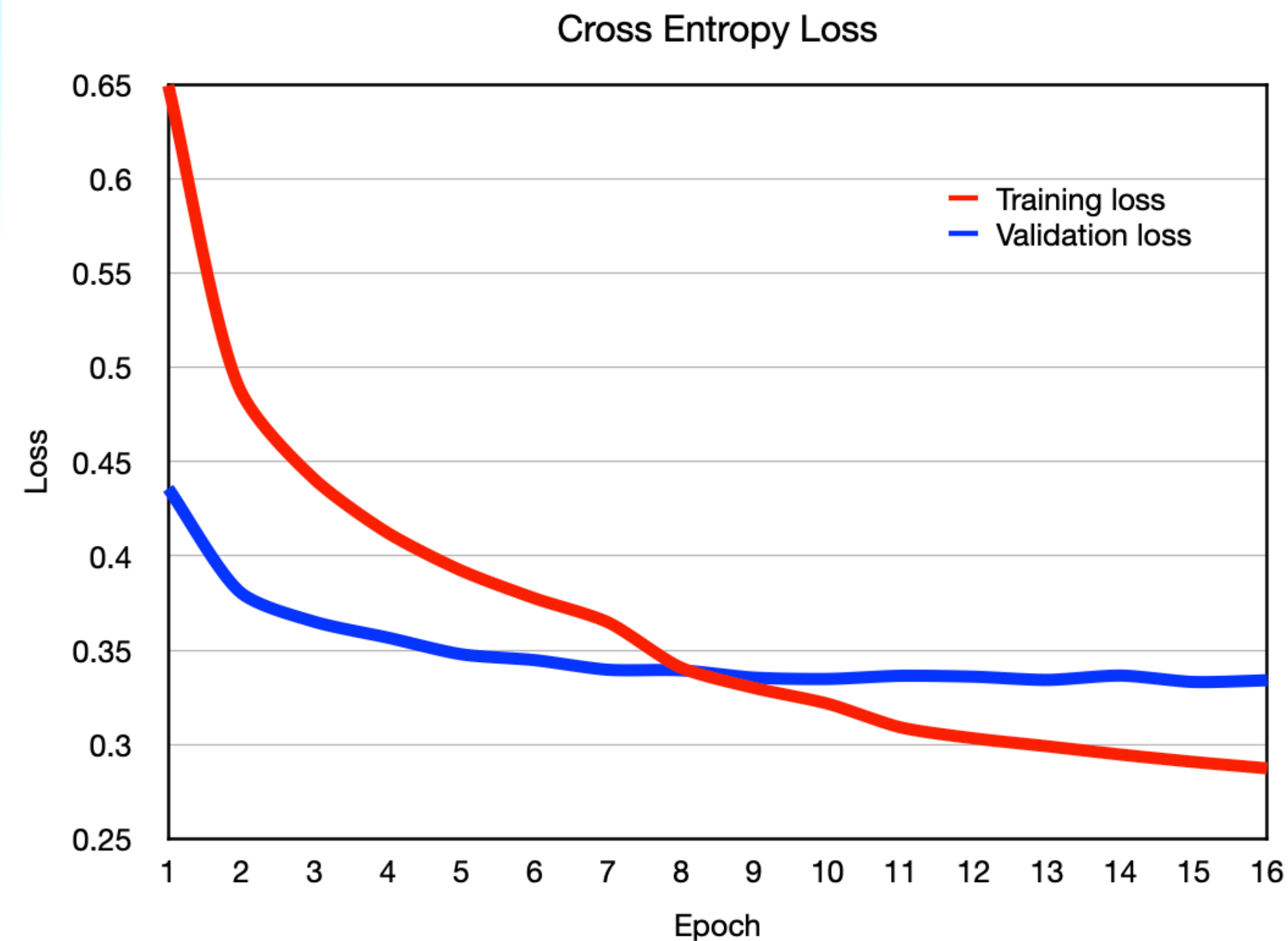
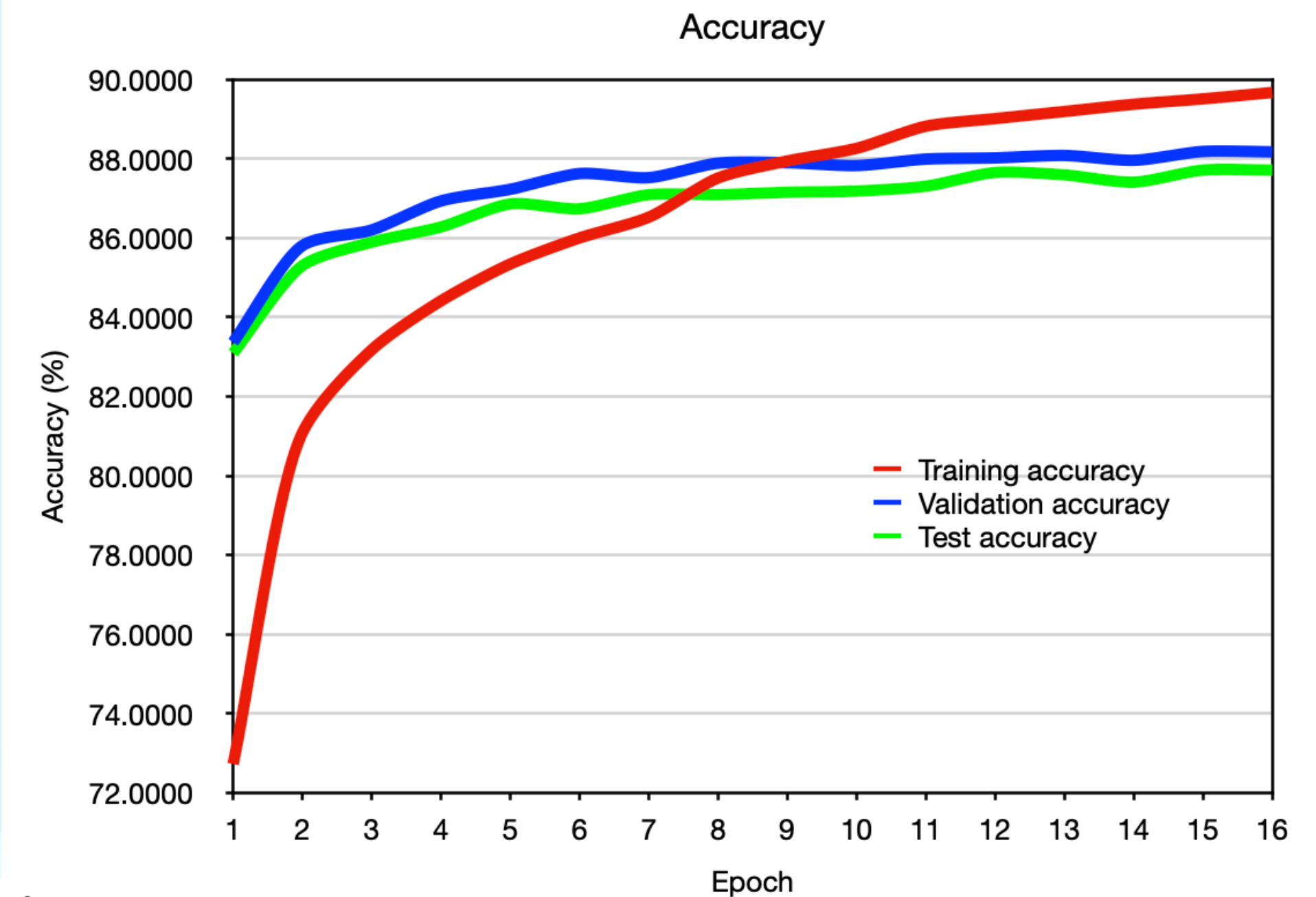


# Exploration: SNLI Dataset

- Stanford Natural Language Inference dataset
- 570k pairs of premise and hypothesis
- Benchmark for NLI

# Result

- My outcome after 16 epochs
  - training accuracy 89.7%,
  - validation accuracy 88.2%
  - test accuracy 87.7%
- Coet outcome
  - training accuracy 93.2%,
  - validation accuracy 88.4%
  - test accuracy 88.0%



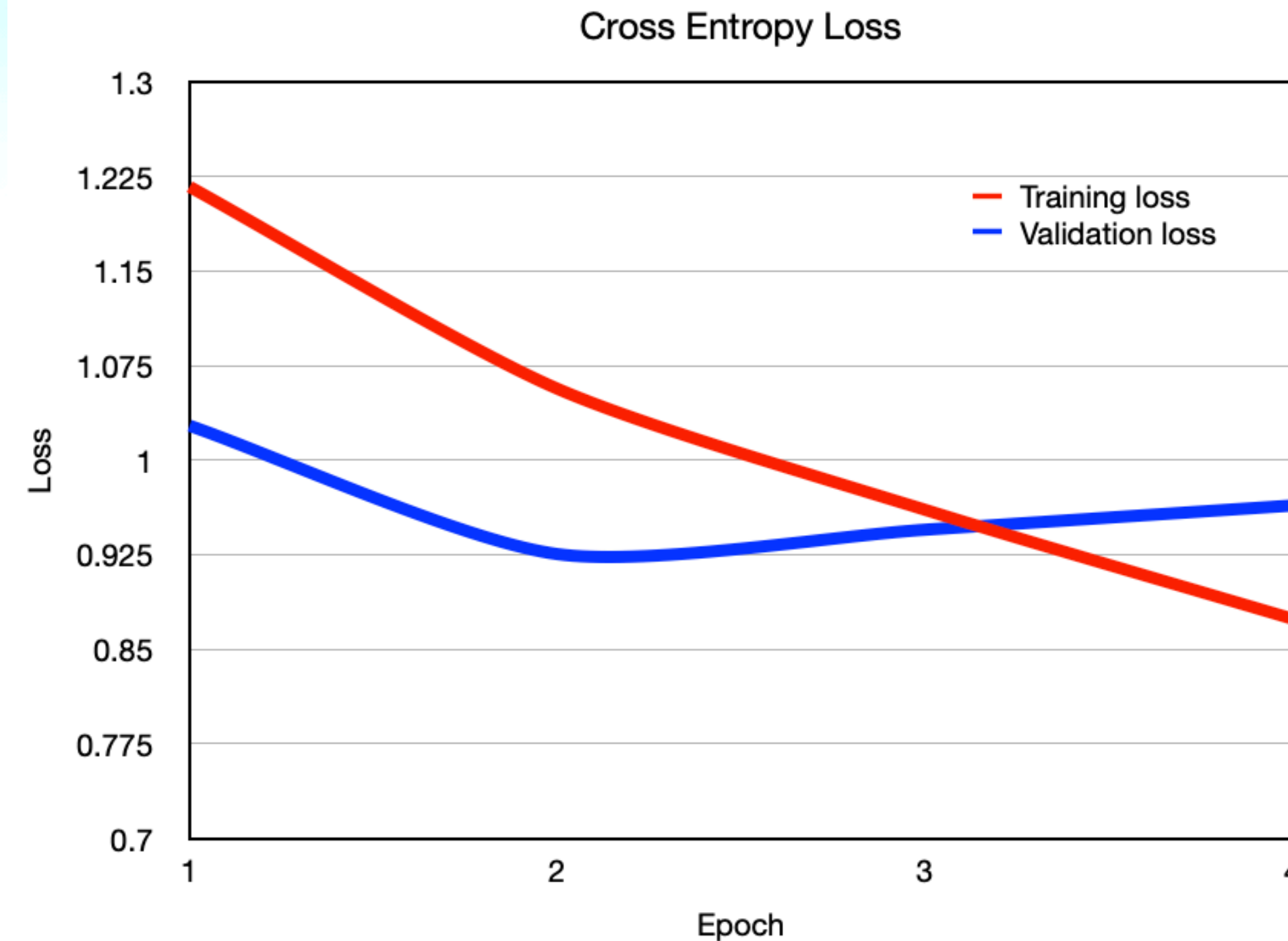
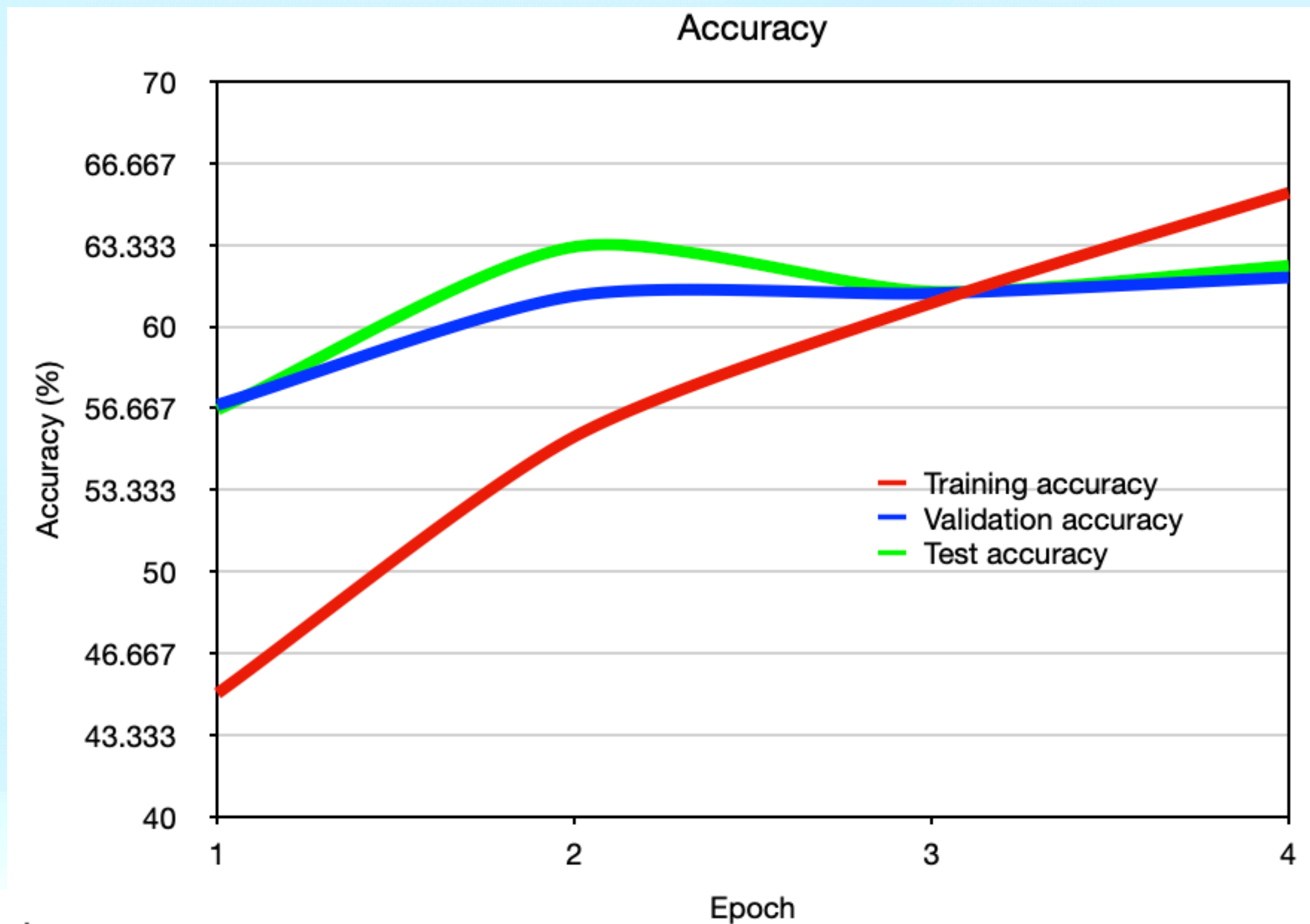


# SciNLI Dataset

- Published in 2022
- Capturing the formality in scientific text
- 107k pairs of premise and hypothesis
- Accuracy 61.32% in BiLSTM model

# Results: trial 1

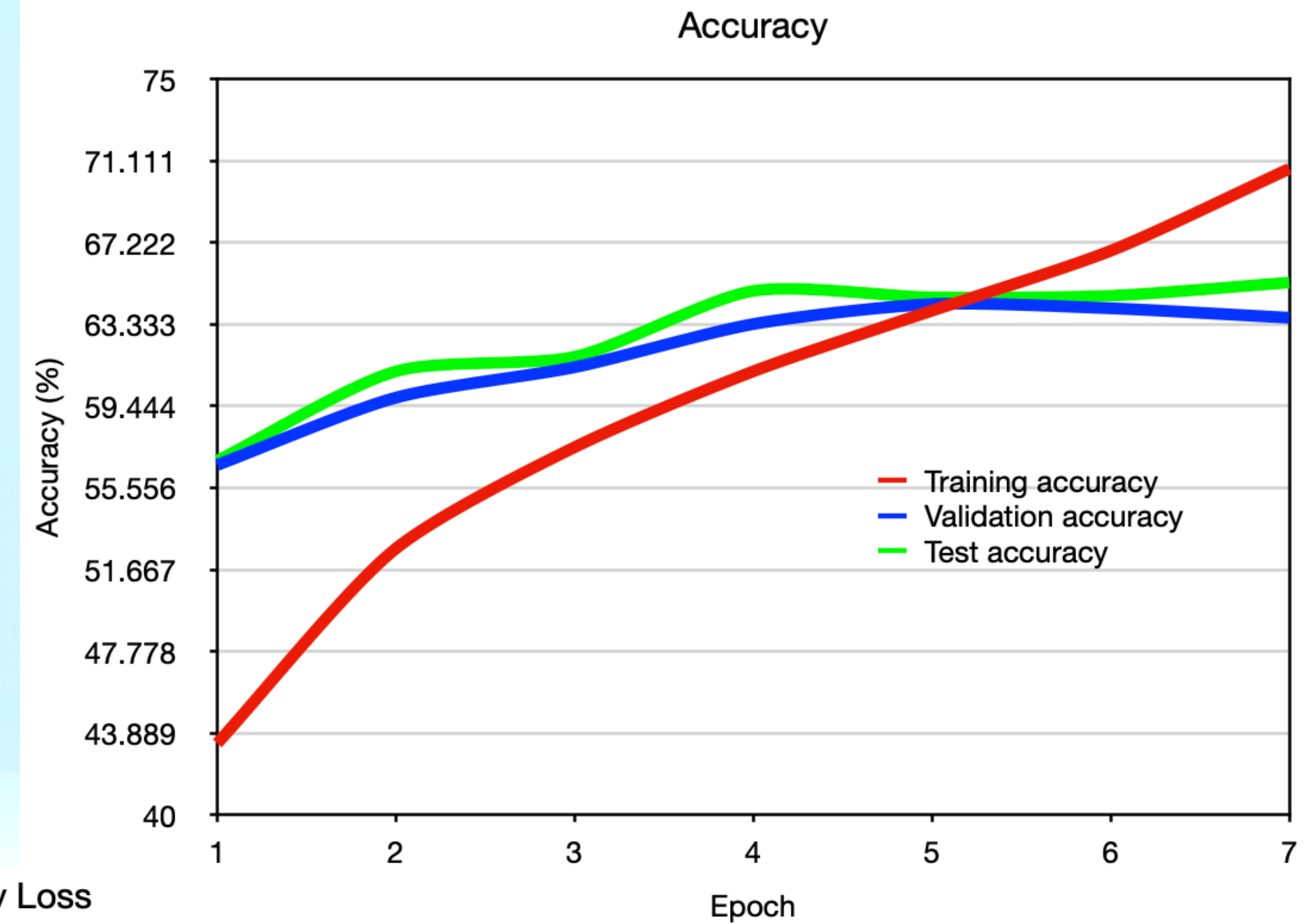
- Hidden size: 300
- Learning rate: 0.001
- Max gradient norm: 10.0
- Overfitting in 2 epochs
- Accuracy ~63%





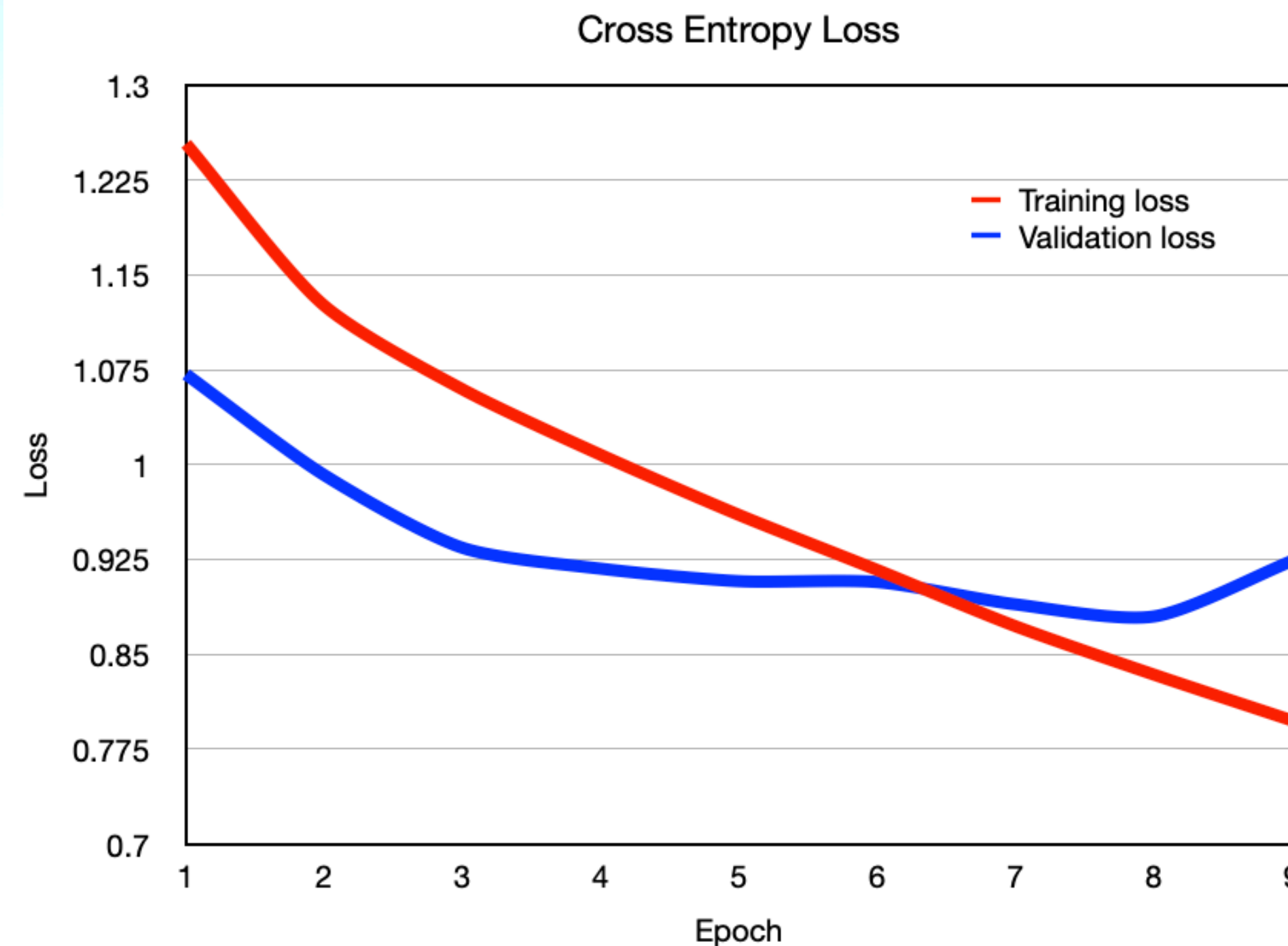
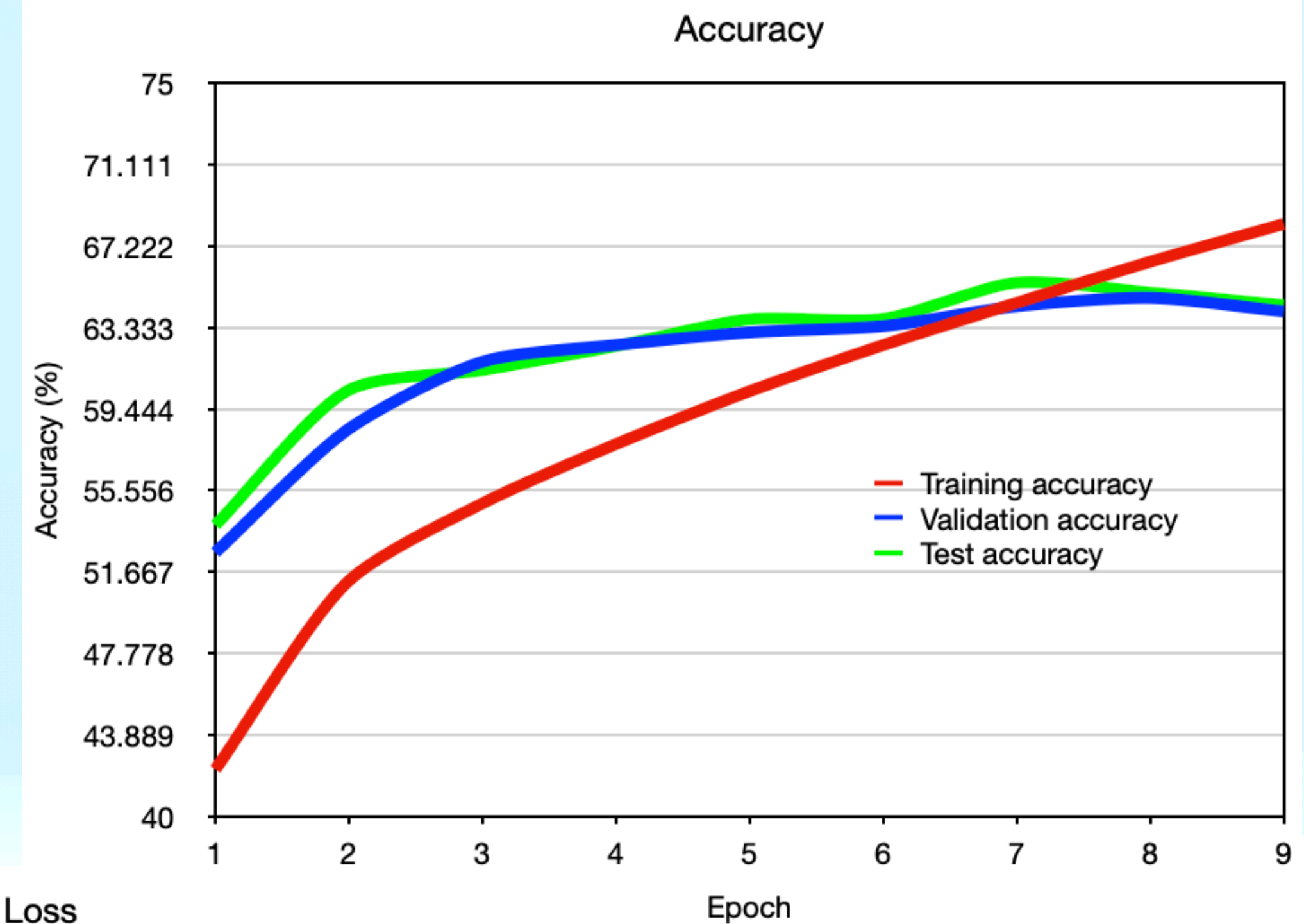
# Results: trial 2

- Hidden size: 300
- Learning rate: 0.0004
- Max gradient norm: 10.0
- Overfitting in 5 epochs
- Accuracy ~64.7%



# Results: trial 3

- Hidden size: 500
- Learning rate: 0.0002
- Max gradient norm: 5.0
- Overfitting in 8 epochs
- Accuracy ~65%





# Conclusion

- ESIM is stronger than BiLSTM on SNLI and SciNLI datasets
- But ESIM is not good at SciNLI dataset
- Maybe learn about other algorithms to improve accuracy on SciNLI dataset