

Report 1:

Shilo Avital 206487407 Tamar Michelson 323805861

1. Dataset

There are 10,000 sequences, 5,000 positive examples, and 5,000 negative examples. Of these, the model was trained on 8,000 sequences and tested on 2,000 sequences.

2. Model Architecture

The model used is an LSTM-based acceptor with an embedding layer, an LSTMCell and a multi-layer perceptron with one hidden layer (16 units, ReLU activation) and a final output layer for binary classification.

3. Training Details

- Embedding dimension: 16
- Hidden dimension: 128
- Optimizer: Adam
- Loss: Binary Cross-Entropy with Logits
- Epochs: 10
- Batch size: 64
- Learning rate: 0.001

4. Performance

The model was evaluated on a separate test set. The final test accuracy was approximately reported per epoch. Training loss and Test accuracy are below. The results were good and the model learned excellently to separate the right and wrong sequences. Already in the first EPOCH the model converged to good results. It took a minute on a PC.

5. Observations and Fixes

Some key aspects in model development included proper masking of padded elements in sequences to prevent state updates for padding. Initial trials showed high training accuracy but required adjustments for generalization such as tuning learning rate and embedding dimension.

