

CSE 4392 Special Topic: Natural Language Processing

Homework 9 - Spring 2025

April 3, 2025

Question 1 - 40%

Provide a step-by-step proof of the following formula for RNNs. Any omissions or gaps in the reasoning may result in mark deductions.

$$\frac{\partial L}{\partial \mathbf{W}} = -\frac{1}{n} \sum_{t=1}^n \sum_{k=1}^t \frac{\partial L_t}{\partial \mathbf{h}_t} \left(\prod_{j=k+1}^t \frac{\partial \mathbf{h}_j}{\partial \mathbf{h}_{j-1}} \right) \frac{\partial \mathbf{h}_k}{\partial \mathbf{W}},$$

Then discuss intuitively how you expect the formula to change for bidirectional RNNs.

Question 2 - 60%

Compare LSTMs, GRUs and RNNs with PyTorch for fake news detection by employing the following Fake News Dataset. Implement the whole NLP pipeline as we did before for spam classification with RNNs and submit a notebook or report with your results.

You may be penalized if your results (in terms of F1 score) significantly underperform compared to the ordinary so you may need to do some hyperparameter tuning. Let the validation set be 20% of the provided training set and include at least two visuals (one for the data and another for training and validation losses). At least one table comparing the three architectures should be presented.

Preferably, submit a PDF notebook using this tool. Or submit a Python file and a report documenting your work.