

ECE661 Quiz 3

Name: _____ UniqueID: _____ Score: _____

This quiz is closed-book. By signing your name above, you agree to follow Duke Community Standard. For True/False questions, no justification is needed.

1. (2pt) (T/F) The higher the perplexity that the language model achieves on the test set, the better it is.
False. Should be lower (From lecture 9 p14)
2. (2pt) (T/F) LSTM is designed to address the ‘gradient vanishing’ problem.
True
3. (2pt) (T/F) One-hot word vectors encode similarity between words.
False. One-hot word vectors are orthogonal. We cannot explore the similarity between words.
4. (2pt) (T/F) LSTM can train all the input words in parallel.
False. LSTM is hard to parallelize.
5. (1pt) (T/F) Including an attention mechanism in RNNs can provide more interpretability.
True
6. (1pt) To what degree do you think you are familiar with this course’s content? 1-5 (open answer) Feel free to leave some comments. :)