

CSSE 315 – Natural Language Processing
Rose-Hulman Institute of Technology

Quiz 03 Solutions

Name (Print): _____ Date: _____

1. Provide at least 2 characteristics of Large Language Models:
 - a.) Large training data
 - b.) Able to perform diverse tasks
 - c.) DL (transformer architecture)
2. Before transformers, what was the popular approach to model word sequences
 - a.) CNN
 - b.) **RNN**
 - c.) LLM
 - d.) CPU
3. Your input to the Deep Learning model is a vector (sequence of words) and your output is the probability of whether the text is positive or negative. Select the correct architecture frame for input-output mapping
 - a.) one-to-many
 - b.) many-to-many
 - c.) **many-to-one**
 - d.) LSTM
4. Which model currently allows for the largest context window?
 - a.) GPT4
 - b.) BERT
 - c.) Claude
 - d.) **Gemini**
5. What was the name of the paper published by Google researchers that proposed a transformational neural network called Transformer?
Attention All You Need (2017)
6. Look at the sentence **The animal didn't cross the street because it was too tired** from the article "Illustrated Transformer" [Reading for Tuesday]. Explain in your own words the process allowing the model to refer **it** to **animal**. You can use the back of the page to write your answer.

Self-attention