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

Worksheet 05

Name (Print):_____   Date:_____

1. Warm-up. What does each abbreviation stand for?

   (a) GPT:

   (b) BERT:

2. What is a problem with Bag-of-Word approach?

3. In RNN, the current output not only depends on _____, but also on the _____ , which takes into account _____

4. Provide Examples for each RNN architecture type:

Term	Example
One-to-Many	
Many-to-One	
Many-to-Many indirect	
Many-to-Many direct	

5. When there is a big gap between the critical information and the output, it is called _____

6. A traditional RNN has a _____ hidden state that is passed through time, which can make it difficult for the network to learn long-term dependencies. LSTM solved this problem by using a _____ that can hold information for an extended period

7. Describe Self-attention: