## Reading Assignment from Melanie Mitchell's "Artificial Intelligence: A Guide for Thinking Humans" Assignment by Carrie Corcoran

## Chapter 11: Words, and the Company they Keep

1. What is the phrase used to describe the field of AI that pertains to "getting computers to deal with human language." Natural Language Processing

3. TRUE/FALSE - Understanding natural language is among AI's most difficult challenges, largely because language is inherently ambiguous, is deeply dependent on context, and assumes a great deal of background knowledge common to the communicating parties. True

5. TRUE/FALSE - In the 1990's, rule-based NLP approaches were overshadowed by more successful statistical approaches, in which massive data sets were employed to train machine-learning algorithms, which today are informed by deep learning. True

7. TRUE/FALSE - Current speech recognition systems are so successful because they understand the meaning of the speech they are transcribing. False

9. What advance in NLP does MM believe will be required in order to deal with the outstanding issues of noise, unfamiliar accents, unknown words, and the fact that ambiguity and context sensitivity of language, that collectively impinge upon interpreting speech? She believes they will have to actually understand the meaning of the speech.

11. TRUE/FALSE - An AI system that could accurately classify a sentence (or longer passage) as to its sentiment-positive, negative, or some other dimension of opinion-would be solid gold to companies that want to analyze customers' comments about their products, find new potential customers, automate product recommendations ("people who liked X also like Y"), or selectively target their online advertisements. True

13. What kind of neural network tends to be used to process sentences (or melodies, or stock market trends), in such a way that a sequence of inputs of arbitrary lengths are mapped to a fixed size structure in order to capture, in some sense, a comprehensive representation of the inputs.

Recurrent neural network

15. In abstract terms, describe the process which governs recurrent neural network execution for "encoding" networks.

A recurrent neural network examines each word in sequence in a series of time steps.

17. Neural networks require their inputs to be numbers. So when processing a sequence of words, a mechanism is needed to map words to numbers. Name, and describe, the simplest such mechanism that is presented in the text.

This is known as one-hot encoding. Each word in the machine's vocabulary corresponds to a numeric value.

19. TRUE/FALSE - In linguistics, the idea captured by John Firth's poetic phrase is more formally known as distributional semantics, which expresses the following theory: "The degree of semantic similarity between two linguistic expressions A and B is a function of the similarity of the linguistic contexts in which A and B can appear." True

21. TRUE/FALSE - Once all the words in the vocabulary are properly placed in semantic space, the meaning of a word can be represented by its location in space-that is, by the coordinates defining its word vector. Moreover, it turns out that using word vectors as numerical inputs to represent words, as opposed to the simple one-hot scheme, greatly improves the performance of neural networks in NLP tasks.

True

23. What analogy did MM present as a segue to her discussion of word vectors giving rise to the notion of thought vectors, as an enchancement of word vector technology for capturing semantics?

"To a person with a hammer, everything looks like a nail; to an AI researcher with a neural network, everything looks like a vector."