Portfolio Assignment: WordNet

Objectives:

- Demonstrate basic skills using WordNet and SentiWordNet
- Learn to identify collocations

Turn in:

- Upload your pdf document to eLearning for grading
- Upload to your portfolio, and include a link on your index page

Instructions:

1. Create a Python notebook with appropriate headings, which you will print-to-pdf to upload. Write a 2-3 sentence summary of WordNet.
2. Select a noun. Output all synsets.
3. Select one synset from the list of synsets. Extract its definition, usage examples, and lemmas. From your selected synset, traverse up the WordNet hierarchy as far as you can, outputting the synsets as you go. Write a couple of sentences observing the way that WordNet is organized for nouns.
4. Output the following (or an empty list if none exist): hypernyms, hyponyms, meronyms, holonyms, antonym.
5. Select a verb. Output all synsets.
6. Select one synset from the list of synsets. Extract its definition, usage examples, and lemmas. From your selected synset, traverse up the WordNet hierarchy as far as you can, outputting the synsets as you go. Write a couple of sentences observing the way that WordNet is organized for verbs.
7. Use morphy to find as many different forms of the word as you can.
8. Select two words that you think might be similar. Find the specific synsets you are interested in. Run the Wu-Palmer similarity metric and the Lesk algorithm. Write a couple of sentences with your observations.
9. Write a couple of sentences about SentiWordNet, describing its functionality and possible use cases. Select an emotionally charged word. Find its senti-synsets and output the polarity scores for each word. Make up a sentence. Output the polarity for each word in the sentence. Write a couple of sentences about your observations of the scores and the utility of knowing these scores in an NLP application.
10. Write a couple of sentences about what a collocation is. Output collocations for text4, the Inaugural corpus. Select one of the collocations identified by NLTK. Calculate mutual information. Write commentary on the results of the mutual information formula and your interpretation.

Grading Rubric:

- 90 and above for exceptional summary and reflection
- 80-89 for good work
- 70-79 for average work
- below 70 for low quality work