Final Exam Notes

Coverage: Linguistic Knowledge / Representations & Algorithms, e.g.,

- Sentiment / Naïve Bayes
- Text Classification / Logistic Regression
- Lexical Semantics / Vector Embeddings
- Lexical Semantics / WordNet
- Shallow Sentence Semantics / Information Extraction & Semantic Role Labeling
- Discourse Phenomena / Coreference Resolution
- Evaluation, Bias

Types of questions:

True/False

- The set of near-synonyms for a WordNet sense is called a *gloss*.
- Yarowsky bootstrapping is a minimally supervised approach to semantic role labeling.
- *Goldfish* is a homonym of *fish*.

Short Answer or similar (conceptual)

- What are two independence assumptions that make the naïve Bayes algorithm naive?
- Discuss one similarity and one difference between Wordnet and Framenet.
- Explain how you would intrinsically versus extrinsically evaluate a word sense disambiguation model.
- What is a difference between WordNet-based and distributional measures of similarity?
- What is a (dis)advantage of semi-supervised WSD, compared to full supervision?
- Explain what specific type of ambiguity in language understanding makes the following dialog from the disaster-movie spoof ``Airplane!'' humorous:
    - Rumack: You'd better tell the Captain we've got to land as soon as we can. This woman has to be gotten to a hospital.
    - Elaine Dickinson: A hospital? What is it?
    - Rumack: It's a big building with patients, but that's not important right biw,

Problem Solving (like hw)

- Train a binarized naive Bayes model on the following document counts for key sentiment words, with positive or negative class as noted. Use the model to assign a

class (pos or neg) to the sentence *A good, good plot but poor acting*.

doc	"good"	"poor"	(class)
d1.	3	0	pos
d2.	2	1	pos
d3.	1	3	neg
d4.	1	5	neg
d5.	0	2	neg

- Represent the semantics of a word using a sparse/dense vector representation
- Compute semantic similarity based on thesaurus/vector representations
- Given an input text, create a gold-standard output for:
  - NER/IE
  - WSD
  - SRL
  - Coreference Resolution