1. Is it possible for a POS tagger to be 95% accurate but still 50% of the sentences have at least one error each. If so, explain how and if not, justify why it is not possible.

2. Use the NLTK or any other POS tagger on the example sentence in Introduction slides with 3 different parse trees. The POS tags are compatible with which parse trees. Explain.

3. Find one unsupervised method and one supervised method in the recent literature (last 5 years or so) for word sense disambiguation and then report in your own words, how they work. Page limit: 1 page each

4. Write code in Python to output the WordNet synset of a given word. Does NLTK have a word sense tagger?

5. For the following situations, consider whether or not the HMM model fits the requirements of the situation? Justify your answer in 1-2 sentences.
   
   (a) Malware detection.
   (b) Likelihood of a person clicking on a phishing email.

6. Read the Esorics 2012 paper of Verma et al. on Phishing Email Detection and write a one page critique. In this critique include what you learned from this paper and suggest at least three major improvements for Phishnet-NLP algorithm.

Academic Honesty Policy: No collaboration with anyone or anything in or outside the course is allowed on any homeworks, exams and programming assignments (yes, that excludes the internet as well) except if it is explicitly allowed on a problem. The appropriate help of the instructor and (if applicable) the TA is of course allowed and encouraged.