CS378 Fall 2020 Midterm Topics

The midterm is a **open book take-home exam**. You are allowed to consult any resources that are helpful. However, **you may not collaborate with other students**!

The exam will have the same broad types of questions (multiple choice, short answer, long answer) as past midterms, with a greater emphasis on open-ended and conceptual questions rather than algorithm execution. **The exam will be designed to take between 2 and 2.5 hours.** Do show your work, but for our sake and yours, please try to refrain from writing an essay as a response unless the question explicitly asks for it.

Submission You will scan and upload your exam into Gradescope. You may type or handwrite responses on a printed exam, or a mix of both. If you type your responses, please try to keep the layout of your pages matching the exam to make our grading easier (e.g., if page 1 consists of 10 multiple-choice questions, try to make your first page consist of 10 multiple-choice questions as well).

Topics Below is the list of topics that will be covered on the midterm, which is the bulk of the course material up to and including October 8.

- Bag-of-words features: how these feature spaces look and how they work for classification
- Perceptron (binary): algorithm, loss function
- Logistic regression (binary): model (the LR formula), training objective, gradient update
- Sentiment analysis: what kinds of features are useful?
- Multiclass classification: how weights and features work in this setting
- Multiclass perceptron: model definition, how to train it, different weights and different features formulations
- Optimization: stochastic gradient descent, impact of step size on optimization, impact of initialization (particularly for FFNNs)
- Feedforward neural networks: definition, initialization
- Training neural networks
- Word embeddings: skip-gram, skip-gram with negative sampling
- Deep averaging networks: model from Assignment 2, limitations of the model
- POS tagging: understanding ambiguities (like the *Fed raises...* example)
- Sequence labeling as classification: how to build position-sensitive features
- Hidden Markov Models: definition, parameter estimation (counting and normalizing), Viterbi algorithm
- Beam search
- Conditional random fields: intuition and general concepts (you won't be expected to know formulas)

- Constituency syntax: what trees look like, ambiguities (*Ban on nude dancing...*, *eat spaghetti with chopsticks*)
- PCFGs: definition, parameter estimation (counting and normalizing), what binarization is (you won't be expected to do it), CKY algorithm
- Dependencies: definition, differences from constituency
- Shift-reduce dependency parsing: how the algorithm works

Other content You should expect to see examples of text and be comfortable reasoning about how these algorithms might work on such examples, as in the assignments so far. We won't expect you to know things like part-of-speech definitions or have encyclopedic knowledge of grammar structures—we will provide the necessary information for such questions.

Readings We won't expect you to know content from the Eisenstein book that hasn't been covered in lecture, even if it's in the posted readings. As for the other assigned readings, we may ask questions pertaining to concepts from these (concepts being at the level discussed in lecture), but we generally won't assume that you've committed the specific approaches to memory.

Practice problems The best source of practice problems is the last couple of midterms, which are posted on the course website. Note that Naive Bayes classification (spring 2019 multiple choice Q1-4) has been dropped as a topic