# CS287: Midterm Guide

## Introduction

The main goals of the first half of the class were to:

1. utilize and reason about standard neural network models;
2. develop familiarity applying these tools to language and related domains;
3. be able to read and understand recent papers written in a variety of styles.

The midterm exam will therefore target these three areas.

## 1 Models

Questions:

- What is the computational and memory complexity of running this model? How might this change during training or test time?
- How many effective parameters does this model have? How does this relate to the input?
- What do the gradients of this model look like? Which components are updated each time?
- What is the receptive field of this network? How can this be verified or tested on real data?
- What are the best cases and worst cases for the features learned by a model? What does it look like when it works or fails?
- In the best case, what features can the model learn? How would you hard-code the parameters to manually construct certain features?
- Model Comparison: When can one model be written in terms of another? Identify linear versions of models, e.g. bag-of-words as conv+sum-over time.

Model Classes:

- Linear Models
- Feed-Forward Neural Networks
- Pooling (mean, max, sum, over-time)

- Convolutional Neural Networks
- Recurrent Neural Networks
- GRU and LSTM
- Soft Attention
- Hard Attention
- VAEs/GANs

## 2 Applications

Questions:

- What underlying aspect of a sentence is a model picking up on, e.g. if we know a convolution works for problem X, what aspects must be important.
- How to use different granularities of sentence? How might you do character versus word versus ngram modeling versus document modeling?
- Why does transfer matter so much in language? What techniques have proven useful for transfer and how do they work?
- How does domain change the properties of the language used? What aspects of a model are dependent on domain versus more general phenomenon?
- What necessary aspects of language understanding are latent, and which techniques have proven useful for handling this? How can we introspect models to determine subdecisions?
- Why is language generation a difficult problem, and how do we measure progress? Which components of neural systems specialize to different aspects of generation?
- How are systems trained differently for different tasks? Which training decisions are motivated by theory and which are practical decisions motivated by convenience?

Applications:

- Sequence Classification (sentiment analysis)
- Sequence Labeling (part-of-speech tagging or named entity recognition)
- Next Step Prediction (language modeling)
- Conditional Next-Step (translation)

Note: You should also have passing fluency with non-text applications including image convolutions and speech inputs.

## 3 Papers

The midterm will also contain question based on the following four papers that we have discussed in the first half of the class. You will be allowed to bring these papers into the exam.

- Efficient Estimation of Word Representations in Vector Space
- CNNs for Sentence Classification
- Sequence-to-Sequence Learning with Recurrent Neural Networks
- Sequence-level Training with Recurrent Neural Networks
- Show, Attend, and Tell

For these questions, we will ask direct questions about choices in the papers and questions about their baselines. You should **not** spend time memorizing the papers, but instead consider each of the decision points and what empirical component of the system they were meant to exercise.