Natural Language Processing (CSE 447/547M): Finale

Noah Smith

University of Washington
nasmith@cs.washington.edu

March 13, 2019

Outline of CSE 447/547M

Revised from the intro lecture.

1. **Probabilistic language models**, which define probability distributions over text passages. (about 2 weeks)
2. **Text classifiers**, which infer attributes of a piece of text by "reading" it. (about 1 week)
3. **Word representations** (about 1 week)
4. **Sequence models** (about 1 week)
5. **Syntax** (about 1.5 weeks)
6. **Machine translation** (about 1 week)
7. **Semantics** (about 1 week)

Some topics we didn't have time for; where to keep learning

- Conversational NLP (EE 596D; last year's course page is here: `https://hao-fang.github.io/ee596_spr2018/`)
- Application areas: information extraction, question answering, language and vision, language and robotics, NLP for social science
- Advanced topics in analysis: natural language inference and paraphrase, discourse, pragmatics
- Advanced topics in machine learning for NLP: graphical models, structured prediction
- Deeper coverage of neural networks (e.g., transformers)
- A great reading list: `https://wammar.github.io/2018sp_uw_cse_599/index.html`
- Keep up with new developments: `https://soundcloud.com/nlp-highlights`

Recurring Themes

- The role of machine learning, and the evolution of techniques (relative frequencies, log-linear models, neural networks)
- The role of data (e.g., annotation, bitext)
- Challenges specific to natural language (ambiguity, the invisibility of $\mathcal{R}$, variation in language)
- The importance and challenge of evaluation
- Useful building blocks, from high-level abstractions (e.g., noisy channel) to low-level tools (e.g., recurrent neural networks, dynamic programming)

Desiderata for NLP Methods

(ordered arbitrarily)

From the intro lecture.

1. Sensitivity to a wide range of the phenomena and constraints in human language
2. Generality across different languages, genres, styles, and modalities
3. Computational efficiency at construction time and runtime
4. Strong formal guarantees (e.g., convergence, statistical efficiency, consistency, etc.)
5. High accuracy when judged against expert annotations and/or task-specific performance
6. Explainable to human users (added in 2019)