Natural Language Processing (CSE 517 & 447): Finale

Noah Smith

University of Washington
nasmith@cs.washington.edu

Winter 2022

What We Covered

1. Classification (MLR)
2. Language models (up to and including transformers)
3. Vector embeddings (up to and including BERT)
4. Morphology (WFSTs)
5. Sequence labeling (CRFs)
6. Syntax and semantics (linguistic structure prediction)
7. Translation (sequence-to-sequence models)

Some topics we didn't have time for

- Conversational NLP
- 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 generation: summarization, captioning, controllable generation
- Advanced topics in machine learning for NLP: graphical models, unsupervised methods, Bayesian modeling
- "Deeper" coverage of neural networks (e.g., theory, learning algorithms)
- Speech processing

Recurring Themes

- The role of machine learning, and the evolution of techniques (relative frequencies, 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., transformers, 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

You Should Now Be Able To

- Learn more if you want to (see resources on class web page)
- Resist the hype
- Find your way to sensible initial solutions to NLP problems
- Apply ideas from NLP to other areas where problems are hard to formalize and messy to solve