

Review and Exam



Exam

- Closed book, one page of notes
- 1 hour and 30 minutes to complete
- Conceptual questions, minor calculation
- Covers everything we've discussed in class so far
- Examples in our slides are very relevant



Text Classification

- Naïve Bayes Classifier
 - How does it work?
 - Compute conditional probabilities given word and class frequencies, determine predicted class for a given model.
- Evaluation Measures: accuracy, confusion matrix
- Text Similarity: bag-of-words representation, TF, TF-IDF
- Machine Learning Models:
 - Basic development cycle
 - Feature Engineering,
 - Logistic Regression, SVMs



Word Embeddings

- Distributional Semantics: representing a word by its context.
- Word Vectors/Embeddings: what are they?
- Word2Vec
 - CBOW vs. Skip-Grams
 - Objective Function
- Count-based Methods
 - Glove: the log-bilinear model
- Evaluation
 - Intrinsic vs. Extrinsic



Deep Learning

- Batch Gradient Descent vs. Stochastic Gradient Descent
- Conceptual Questions:
 - Activation Function, Computational Graphs, Back-Propagation, Multi-Layer Perceptron
- NOT IN THE TEST
 - Gradient Computation
 - Convolution



Sequential Labeling

- Problem Formulation, Potential Applications
- POS Tagging Problem
- Hidden Markov Models:
 - Compute transition and emission probabilities for a given data.
 - Viterbi decoding
 - Model complexity
- Maximum Entropy Markov Models (MaxEnt), CRFs
 - Conceptual understanding: pros and cons, Markov assumption, label bias
- NOT IN THE TEST
 - Specific formulas for MEMM, CRFs
 - Dynamic programming



Parsing

- Constituent Parsing vs. Dependency Parsing
- Basic Concepts: Context-free Grammar (CFG), Chomsky Normal Form, Bracketed Notation, Ambiguities
- CYK Parsing
- Probabilistic CFG: compute probabilities for production rules
- Transition-based Dependency Parsing: arc-standard algorithm
- Evaluation
- NOT IN THE TEST
 - Parent annotation, lexicalized PCFG, non-projective parsing



Information Extraction

- Information Extraction (IE) Pipeline
- Evaluation for different IE tasks: NER, Relation Extraction
- Conceptual Ideas: model generalization beyond lexical matching, kernel methods
- NOT IN THE TEST:
 - Semi-supervised learning, active learning, confidence estimation

