

Summary

LING 570

Fei Xia

Outline

- Main units
- Main techniques
- What's next?

Main units

Unit #0: introduction and summary

- Course overview
- Tokenization
- Introduction to probability theory
- Summary

Unit #1: finite-state machine

- Formal language
- Formal grammar
- Regular expression
- FSA
- Regular relation and FST
- Morphological analyzer

Unit #2: LM, HMM, and n-gram tagger

- LM: n-gram models
- Smoothing
- HMM
- N-gram model
- Viterbi algorithm

Unit #3: Classification

- Classification problem
- Sequence labeling and beam search
- Example tasks:
 - POS tagger
 - NE tagging
 - Chunking

Unit #4: Neural Network

- NN basics: neurons, activation functions, learning weights
- Word embeddings
- Word2vec
- Neural LM

Main techniques

Main techniques (1)

- Probability theory:

- The chain rule:

$$P(X_1, X_2, \dots, X_n) = P(X_1)P(X_2|X_1)\dots P(X_n|X_1, \dots, X_{n-1})$$

- The Bayes' rule: $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$

- The (conditional) independence assumption

$$P(X_n|X_1, \dots, X_{n-1}) = P(X_n|X_{n-k}, \dots, X_{n-1})$$

- ...

Main techniques (2)

- Regexp, regular language and regular grammar, and FSA
- Regular relation and FST
- FSA, FST, and morphological analysis:
 - FSA: DFA, NFA, PFA, WFA
 - Combining simple FSTs in a pipeline can be very powerful.

Main Techniques (3)

- LM and Smoothing
- N-gram model
- HMM
 - The Markov assumption
 - Viterbi algorithm

Main techniques (4)

- Classification and sequence labeling problems:
 - Representing an instance as a feature vector
 - Selecting features is very important
 - Many problems can be treated as classification or sequence labeling problems
 - Beam search

Main techniques (5)

- NN basics:
 - NN can approximate any function with non-linear activation functions.
 - Instead of defining features, design an NN architecture and tune hyperparameters.
 - Words are represented as vectors (aka word embeddings)
- Learn the weights of NN via backpropagation
- Learn word embeddings via “fake” tasks

Tools created

- Prerequisites (programming, FSA, prob): Hw1
- FSA and formal languages: Hw2 - Hw3
- Morphological analyzer with FST: Hw4 – Hw5
- LM, HMM, and Viterbi: Hw6, Hw8, Hw9
- Form feat vectors: Hw10
- NN basics: Hw7
- Using existing packages:
 - Carmel
 - Mallet

What's next?

Other CL courses

- Other tasks → LING 571 (winter)
 - Ex: parsing, semantics, discourse, ...
- Supervised learning → LING 572 (winter)
 - Ex: MaxEnt, Naïve Bayes, SVM, NN, ...
- System → LING 573 (spring)
- LING 575 (winter, spring)

LING 572 in winter 2023

- Unit #0: Introduction
 - 1 week
 - Features, training/testing, ...
 - Classification algorithms
- Unit #1: Simple algorithms
 - 2 weeks
 - kNN
 - Decision tree
 - Naïve Bayes

LING 572 (cont)

- Unit #2: More sophisticated algorithms
 - 5 weeks
 - MaxEnt (*)
 - SVM (**)
 - CRF (**)
- Unit #3: NN: 1-2 weeks

LING 572 (cont)

- More math in ling572:
 - Information theory:
 - Entropy, mutual information
 - Chapter 2 in (Manning & Schutze, 1999)
 - Calculus, derivative of $f(x)$, gradient, Lagrange multipliers

Beyond Ling572

- Advanced machine learning algorithms:
 - Graphical models
 - Bayesian
 - Joint inference
 - Reinforce learning
 - ...
- More applications:
 - Speech
 - MT
 - IR
 - QA
 - ...

Course evaluation

- Course evaluation (for Fei and Yuanhe): see “Announcement” at Canvas.
- Deadline for completing the evaluation: 11:59pm on Dec 9:
 - Fei: <https://uw.iasystem.org/survey/261955>
 - Yuanhe: you should have received an email

LING570 due date

- Hw9: today (2-day free extension)
- Hw10: next Thurs (12/15)

Your final grades

- For hw2-hw10:
 - Remove the largest late penalty
 - Remove the lowest score
 - Calculate the average
- Add the bonus points (if any)
- Map the percentage to a final grade