CS 378 Lecture 8

Today
- Bias in embeddings
- Part-of-speech intro
- Sequence tagging
- Tagging with classifiers

Announcements
[praise in zoom]
- A1
- A2
- Survey:
  Good: rewritable lectures, exercises/breakouts, notes
  Not so good: ① Working in pairs on HW - please do!
  ② More exercises
  ③ Bigger breakouts/more engagement
  ④ More context/big picture
  ⑤ Reading guidance - bolding
Recap: Skip-gram

Input: corpus of text
Corpus \( \Rightarrow (x, y) \) pairs which are \( \leq k \) words apart (\( k \) = window size)

\[
P(\text{context} = y | \text{word} = x) = \frac{e^{v_x \cdot c_y}}{\sum_{y' \in V} e^{v_x \cdot c_{y'}}}
\]

Maximize log likelihood of data \( \Rightarrow \) get useful \( \tilde{v}_s, \tilde{c}_s \)
Use \( \tilde{v}_s, \tilde{c}_s \) or \( \tilde{v} + \tilde{c} \) in downstream tasks

the fish swam quickly \( k = 2 \)
Where we are

Classification: arg\text{max}_{\hat{y}} \ w^\top f(\hat{x})

sent \Rightarrow label

doc \Rightarrow label for each word in that sentence

Part-of-speech tagging

1. Structurally different problem

\begin{align*}
X & \rightarrow Y \\
x_1, \ldots, x_n & \rightarrow y_1, \ldots, y_n \text{ tag for each word}
\end{align*}

2. Syntax
Parts-of-speech

Text to speech: record

Info. extraction: arms V or N?

POS Tags

Open-class: there
can be new words here

Nouns: Proper: IBM,
Common: cat

Verbs: see, registered

Adjs: yellow

Adv: swiftly

Closed-class: fixed set

Determiners: the, a (articles)

Some

DT + N => NP

Conjunctions: and/or

Pronouns

Prepositions: up, on, ...

Particles: made up

Auxiliaries, modals:

Aux: had [V]

Modals: could/should
"real" Penn Treebank

fed NNP proper noun  I fed her...
VB D past tense verb
vBN participial  I had fed...

raises NNS plural noun
VBZ 3rd person present verb

interest NN noun
VBP  I interest you
VB infinitive  I want to

rates NNS
VBZ

0.5 CD cardinal

percent NN

rates are correct

rates are a lively thing

that get interested
Sequence Tagging

Input: $\mathbf{X} = (x_1, ..., x_n)$ $x_i \in V$ words

Output: $\mathbf{Y} = (y_1, ..., y_n)$ $y_i \in Y$ tags

Structured classification: output has structure

Start: use classifiers

$P(y_i = y | \mathbf{X})$ use LR, ... to assign tag $y$ to word $i$ in sentence $\mathbf{X}$

Next: Hidden Markov Models