CS 7650 - lec7-seq2-notes

Monday, February 5, 2024 1:43 PM

Sequential Models

$x = (x_1, x_2, \cdots, x_n)$ $y = (y_1, y_2, \cdots, y_n)$

$\hat{y} = \underset{y' \in \mathcal{Y}(x)}{\operatorname{argmax}} \phi(x, y')$

$|\mathcal{Y}(x)| = |\mathcal{Y}|^n$

$\mathcal{Y} = \{B\text{-PER}, I\text{-PER}, B\text{-ORG}, \cdots O\}$

Scoring func. on pairs of seqs

System #1

I-PER I-PER O - - - -

Barack Obama will - - - -

B-LOC O-O B-ORG - -

Hangzhou - - - G20 - -

$\mathcal{Y}^n \times \mathcal{Y}^n \rightarrow \mathbb{R}$

HMM $\hat{y} = \underset{y}{\operatorname{argmax}} P(y|x) = \underset{y}{\operatorname{argmax}} \frac{P(x,y)}{P(x)}$

CRF: $\hat{y} = \underset{y}{\operatorname{argmax}} P(y|x) = \underset{y}{\operatorname{argmax}} \frac{\exp(\phi(x,y))}{\sum_{y' \in \mathcal{Y}(x)} \exp(\phi(x,y'))}$

$\phi(x,y) = w^T f(x,y)$

Z

LR training — M training examples

$L(x, y^*) = \sum_{j=1}^{M} \log P(y^{(j)*} | x^{(j)})$

$= \sum_{j=1}^{M} \left(w^T f(x^{(j)}, y^{(j)*}) - \log \sum_{y' \in \mathcal{Y}} \exp(w^T f(x^{(j)}, y')) \right)$

$\frac{\partial}{\partial w} L(x, y^*) = f(x^{(j)}, y^{(j)*}) - \sum_{y} f(x^{(j)}, y^{(j)}) P(y | x^{(j)})$

CRF train — $|\mathcal{Y}|^n$

$L(x, y^*) = \sum_{j=1}^{M} \log P(y^{(j)*} | x^{(j)})$

$\frac{\partial}{\partial w} L(x, y^*) = f(x^{(j)}, y^{(j)*}) - \sum_{y \in \mathcal{Y}^n} f(x^{(j)}, y) P(y | x^{(j)})$

exponential!

F1

System #2

B-PER O O - - - B-LOC O O - - - B-ORG O O

Barack Obama will - - - Hangzhou - - - G20 meeting

precision = ? $\frac{2}{4}$

recall = ? $\frac{2}{3}$