Lecture 11/19: Structured Output Prediction - Discriminative Training

Tuesday, November 19, 2019 2:48 PM

Review of 11/14

Structured Output Prediction: $h: X \rightarrow Y$ where Y is combinatorial

Generative Modeling: $h(x) = \arg\max_y \{ P(Y=y, X=x) \} = \arg\max_y \{ P(X=x \mid Y=y) \cdot P(Y=y) \}$

Sequence prediction with HMM

Prior: $P(Y=(y_1 \dots y_\ell)) = P(Y_{start}=y_1) \cdot \prod_{i=2}^{\ell} P(Y_{next}=y_i \mid Y_{prev}=y_{i-1})$

Class cond: $P(X=(x_1 \dots x_\ell) \mid Y=(y_1 \dots y_\ell)) = \prod_{i=1}^{\ell} P(X=x_i \mid Y=y_i)$

$$\rightarrow h(x) = \arg\max_{y_1 \dots y_\ell} \left\{ P(Y_{start}=y_1) \cdot P(X=x_1 \mid Y=y_1) \cdot \prod_{i=2}^{\ell} P(Y_{next}=y_i \mid Y_{prev}=y_{i-1}) \cdot P(X=x_i \mid Y=y_i) \right\}$$

$\rightarrow$ Viterbi Algorithm $O(\ell |Y|^2)$

$$h(x) = \arg\max_{y_1 \dots y_\ell} \{ w \cdot \phi(x,y) \}$$

Example: The/DET bank/N opens/V

$$\log P(X=(\text{The}, \text{bank}, \text{opens}), Y=(DET, N, V)) = \log P(Y_{start}=DET) + \log P(X=\text{The} \mid Y=DET)$$
$$+ \log P(Y_{next}=N \mid Y_{prev}=DET) + \log P(X=\text{bank} \mid Y=N)$$
$$+ \log P(Y_{next}=V \mid Y_{prev}=N) + \log P(X=\text{opens} \mid Y=V)$$

Want $P(x,y) = w \cdot \phi(x,y)$?

$$\phi(x,y) = \begin{pmatrix} \#Y_{start}=DET \\ \#Y_{start}=N \\ \#Y_{start}=V \\ \#Y_{next}=N \wedge Y_{prev}=DET \\ \vdots \\ \#Y_{next}=N \wedge Y_{prev}=N \\ \vdots \\ \#X=\text{The} \wedge Y=DET \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} \qquad w = \begin{pmatrix} \log P(Y_{start}=DET) \\ \vdots \\ \\ \log P(X=\text{The} \mid Y=DET) \\ \vdots \end{pmatrix}$$

Logistic Regression: Max aposteriori

Prior: $P(w) = N(w \mid 0, \lambda I)$

Class probability: $P(y \mid x, w) = \dfrac{\exp(w \cdot \phi(x,y))}{\sum_{y'} \exp(w \cdot \phi(x,y'))}$

$$w = \arg\max \left\{ P(w) \cdot \prod^{n} P(y_i \mid x_i, w) \right\}$$

$$w = \arg\max_{w} \left\{ P(w) \cdot \prod_{i=1}^{n} P(y_i | x_i, w) \right\}$$

$$= \arg\max_{w} \left\{ -w \cdot w + \sum_{i=1}^{n} \left[w \cdot \phi(x_i, y_i) - \log \sum_{y'} \exp\left(w \cdot \phi(x_i, y') \right) \right] \right\}$$

compute via sum-product