

עיבוד שפה טבעית - תרגיל 3.

עומר שחם - 20403472

1 בינואר 2018

1 We saw in class the following regarding forward-backward algorithm:

The marginal $Pr(y_j, y_{j-1} | x_i)$ is the sum over all paths that have edge (y_{j-1}, y_j) between states $j-1$ and j .

Edge weights between y and y' is $M(y, y')$

$\alpha_i(y) = \sum_{y'} M_i(y, y') * \alpha_{i-1}(y')$ where M is the edge (transition) weights, α = sum of weights over all paths of length $i-1$ ends with y .

$\beta_i(y) = \sum_{y'} M_{i+1}(y, y') * \beta_{i+1}(y')$, β = sum of weights over all paths from the end until y .

Let $X = x_1, \dots, x_N$ = sentence of words

$$Pr(y_i, y_{i-1} | X) = \frac{M_i(y_i, y_{i-1}) * \beta_i(y_i) * \alpha_{i-1}(y_{i-1})}{Z(x)} \rightarrow Pr(y_i | X) = \frac{\alpha(y_i) * \beta_i(y_i)}{Z(x)}$$

From the definition of conditional probability: $Pr(y_i | y_{i-1}, X) = \frac{Pr(y_i, y_{i-1} | X)}{Pr(y_{i-1} | X)}$

$$\text{Now we can assign the terms } * \text{ and } **, \text{ we will get that: } Pr(y_i | y_{i-1}, X) = \frac{\frac{M_i(y_i, y_{i-1}) * \beta_i(y_i) * \alpha_{i-1}(y_{i-1})}{Z(x)}}{\frac{\alpha_{i-1}(y_{i-1}) * \beta_i(y_{i-1})}{Z(x)}} =$$

$$\frac{M_i(y_{i-1}, y_i) * \beta_i(y_i)}{\beta_{i-1}(y_{i-1})}$$

let $Y = y_1, \dots, y_N = \text{labels (tags)}$

Let's initialize 2 matrices F, B of size $|Y| * |Y|$, now let's fill them: Matrix F : $\alpha_i(y)$ = value in row i and column j s.t $y = y_j \in Y$ and $\beta_i(y)$ = value in row i and column j s.t $y = y_j \in Y$

Notations $F[i](y) = \alpha_i(y)$ and $B[i](y) = \beta_i(y)$

1.1

input: f- feature function, w- weight vector, x_1, \dots, x_N - sentence of words, i - index

- preprocessing: define a Matrix $A(|Y|*|Y|)$ each row represents the probability distribution on a specific y.
- $i = 0$
- for y in Y:
 - $j = 0$
 - for y' in Y:
 - * $A[i][j] = \frac{M_i(y,y')*B[i](y')}{B[i-1](y)}$
 - * $j++$
 - $i++$
- return A

1.2

Algorithm:

input: f- feature function, w- weight vector, x_1, \dots, x_N - sentence of words, i - index

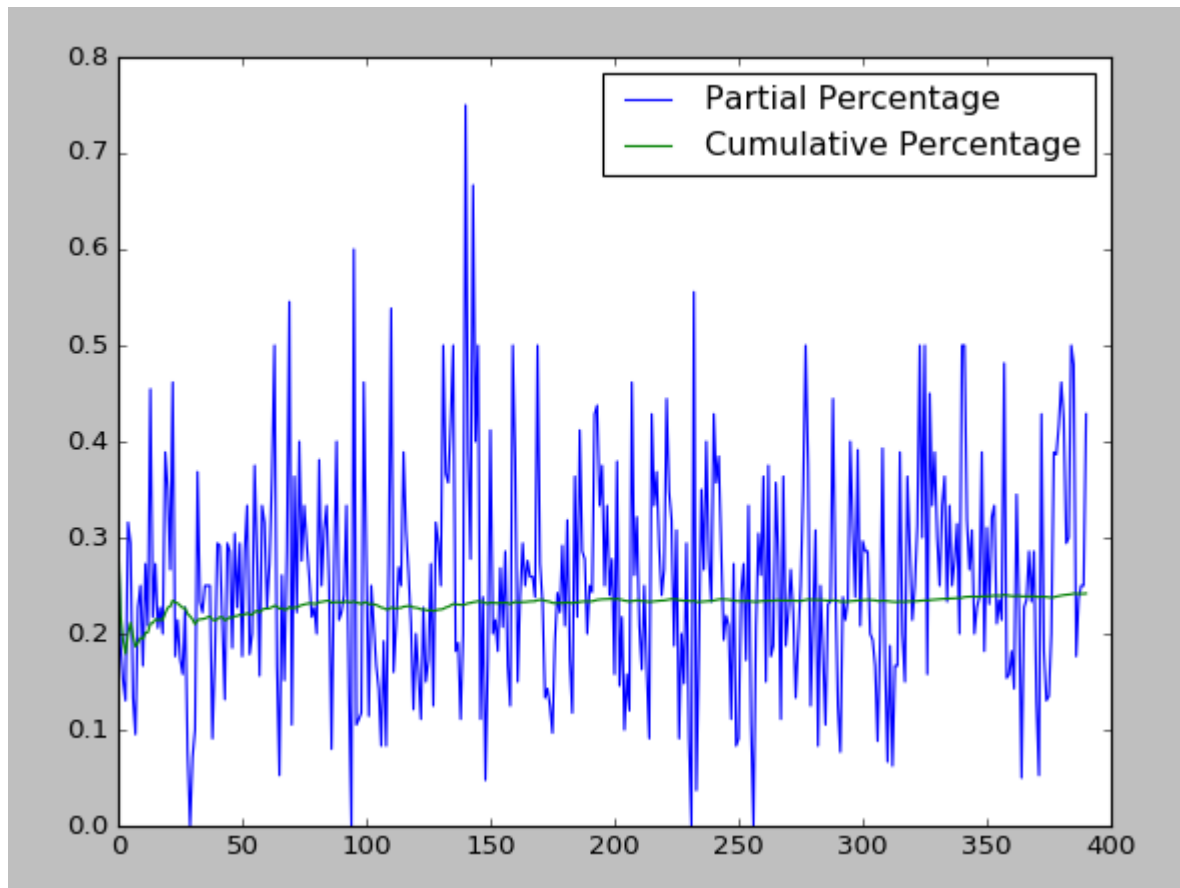
- preprocessing: define v a vector in size $|Y|$ of probabilities.
- $cnt = 0$
- for y in Y:
 - $A[cnt] = \frac{F[i](y)*B[i](y)}{Z(x)}$
- return A

2 Practical Question

2.1

2.2

cumulative score : 0.24. (2 iterations, learning rate = 1)



2.3

2.4

2.5

Cumulative score : 0.28. (2 iteration, adding 4 cells for distance)

