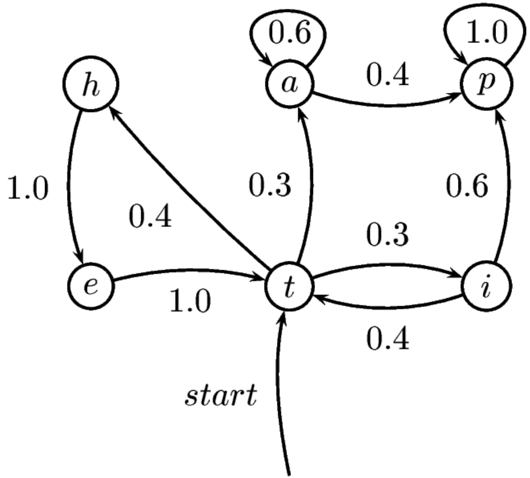
Ukryte modele Markowa

Michał Błauciak

Modele Markowa



Modele Markowa

- Stany: $s_1, s_2, \ldots, s_N$
- Ciąg stanów: $q_1, q_2, \ldots, q_T$
- Własność Markowa: $P(q_t = s_i | q_{t-1} = s_j, q_{t-2} = s_k, \ldots, q_1 = s_l) = P(q_t = s_i | q_{t-1} = s_j)$
- Stacjonarność: $P(q_t = s_i | q_{t-1} = s_j) = (q_{t+u} = s_i | q_{t+u-1} = s_j)$
- Rozkład początkowy Π: $\pi_i = P(q_1 = i)$
- Macierz przejść A: $a_{ij} = P(q_{t+1} = s_j | q_t = s_i)$

Ukryte modele Markowa

- Stany nie są obserwowalne
- Obserwacje pojawiają się z pewnym prawdopodobieństwem w zależności od stanu
- Obserwacje: $k_1, k_2, \ldots, k_M$
- Ciąg obserwacji O: $o_1, o_2, \ldots, o_T$
- Funkcja emisji obserwacji b: $b_i(k) = P(o_t = k | q_t = s_i)$
- Model: $\lambda = (A, b, \Pi)$

Problem pierwszy

Dany jest ciąg obserwacji O i model HMM λ. Jakie jest prawdopodobieństwo $P(O|\lambda)$?

Algorytm forward

- $\alpha_t(i) = P(o_1, o_2, \ldots, o_t, q_t = i|\lambda)$ – prawdpodobieństwo obserwacji takiego $(o_1, o_2, \ldots, o_t)$, że stan q_t jest równy i
- $\alpha_1(i) = \pi_i b_i(o_1)$
- $\alpha_{t+1}(j) = \left(\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right) b_j(o_{t+1})$
- $P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$

Algorytm backward

- $\beta_t(i) = P(o_{t+1}, o_{t+2}, \ldots, o_T | q_t = i, \lambda)$
- $\beta_T(i) = 1$
- $\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)$

Problem drugi

Dany jest ciąg obserwacji O i model HMM λ. Jaki jest najbardziej prawdopodobny ciąg $q = (q_1, q_2, \ldots, q_T)$?

Algorytm Viterbiego

- $\delta_t(i) = \max_{q_1,q_2,\ldots,q_{t-1}} P(q_1, q_2, \ldots, q_t = i, o_1, o_2, \ldots, o_t|\lambda)$
- $\delta_1(i) = \pi_i b_i(o_1)$
- $\psi_1(i) = 0$
- $\delta_t(j) = \max_{i\in\{1,\ldots,N\}} (\delta_{t-1}(i)a_{ij})\, b_j(o_t)$
- $\psi_t(j) = argmax_{i\in\{1,\ldots,N\}} (\delta_{t-1}(i)a_{ij})$
- $\max_{i\in\{1,\ldots,N\}} (\delta_T(i))$
- $argmax_{i\in\{1,\ldots,N\}} (\delta_T(i))$

Problem trzeci

Dany jest ciąg obserwacji O. Jaki jest najbardziej model λ?

Algorytm Bauma-Welcha

- Algorytm iteracyjny – przypadek algorytmu EM
- Prawdopodobieństwo przebywania w stanie i w momencie t i w stanie j w momencie $t+1$:

$$\xi(i,j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)} = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}$$

- Prawdopodobieństwo przebywania w stanie i w momencie t: $\gamma_t(i) = \sum_{j=1}^{N} \xi_i(i,j)$
- Oczekiwany czas przebywania w stanie i: $\sum_{t=1}^{T} \gamma_t(i)$
- Oczekiwana liczba przejść ze stanu i do stanu j: $\sum_{t=1}^{T-1} \xi_t(i,j)$
- Iteracyjne aktualizacje parametrów:

$$\pi_i = \gamma_1(i)$$

$$a_{ij} = \frac{\sum_t \xi_t(i,j)}{\sum_t \gamma_t(i)}$$

$$b_i(k) = \frac{\sum_{t, o_t = k} \gamma_t(i)}{\sum_t \gamma_t(i)}$$

Bibliografia

- C. Manning, H. Schütze, *Foundations of Statistical Natural Language Processing*
- T. Kanungo, *Hidden Markov Models*