

GÉPI TANULÁS 5

Kornai András

KálmánCL 5/11

BESZÁMOLÓK

- A csoport (*Ittész, Molnár, Tóth, Turcsán): prototípus-alapu modellezés jupyter notebookban
- C csoport (Csényi, Hegyi, Jánosy, *Sánta): prototípus-alapu modellezés jupyter notebookban
- B csoport (Goda, Mészáros, Pörtl, *Prótár): beszámoló Leslie Valiant cikkéből (PAC learning)

HMM FOLYT.

A definícióhoz kell

- Állapotok $S = \{s_1, s_2, \dots, s_n\}$
- Bemenő valószínűségek p_1, p_2, \dots, p_n (ez megspórolható csendes kezdőállapot felvételével)
- Átmeneti valószínűségek $t_{i,j}$ ($1 \leq i, j \leq n$)
- Kimeneti valószínűségek $P(f | s_i)$

EGY MENETBEN

Az s_i állapotban kezdünk a p_i -k által diktált módon, kiadunk valami f jelet a $P(f | s_i)$ által diktált módon. Átmegyünk az s_j állapotba $t_{i,j}$ valószínűséggel, s.i.t.

EGY KIMENŐ JELSOROZAT VALÓSZÍNŰSÉGE

$$P(f_1 f_2 \dots f_k) =$$

$$\sum_{\pi \in S^k} p_{\pi(1)} \prod_{j < k} t_{\pi(j), \pi(j+1)} \prod_{j \leq k} P(f_j | \pi(j))$$

Ha csak úgy nekiállunk (brute force) akkor $n^k 2^k$ szorzás kellene, de a Markov-feltétel miatt elég $n^2 2^k$

HOGY HASZNÁLJUK KI A MARKOV FELTÉTELT?

Minden állapothoz külön regiszterben gyűjtjük az adatot

Inicializálás: $R(i, 0) = p_i$ ($1 \leq i \leq n$)

Iteráció: $R(i, k + 1) = \sum_{j=1}^n R(j, k) t_{j,i} P(f_{k+1} | s_j)$
(n szorzás mind az n regiszternél)

Befejezés: $P(f_1 f_2 \dots f_K) = \sum_{j=1}^n R(j, K)$

VITERBI ALGORITMUS

Cél: megtalálni a legjobb $\pi \in S^k$ -t $f_1 f_2 \dots f_k$ -hoz

Brute force: Az összes (n^k) útvonalhoz megszámloljuk, és valami heurisztikával eldobjuk a túl kicsiket

Markov feltételnél: minden állapothoz csak a legjobb odavezető utat tartjuk meg, mert az egyel hosszabb legjobb út ezek valamelyikének folytatása lesz

o	o	o	o	o	o	o
o	o	o	o	o	o	o
o	o	o	o	o	o	o
o	o	o	o	o	o	o
o	o	o	o	o	o	o

HMM FEJLEMÉNYEK

- 1 triphone, quinphone modellek, sőt több (addig amíg van adat)
- 2 kiterjesztés realiztikus adatokra (egybefüggő beszéd, sok beszélő, nagy szókincs, háttérzaj)
- 3 eleve nem volt feltétel, hogy a bemenet grammatikus legyen
- 4 a Baum-Welch stílusú globális optimalizáció ideghálókban nagyon nehéz, alapjában máig megoldatlan
- 5 ezért nincs még mindig rendes morfológia

TUDUNK-E TANULNI?

- Modell-osztályból választunk modellt: nincs olyan, hogy "nyílt modellosztály" vagy "tetszőleges modell"
- Csak a felügyelt tanulás ér el használható eredményeket
- A transzformerek sikere az olcsó felügyelésnek köszönhető
- Nem nagyon értjük a mélytanulósos modelleket, de gyanús, hogy ezek memorizáció+legközelebbi szomszéd alapúak
- Túl sok a paraméter, vagy mégse?