

Primeri napredne uporabe strojnega učenja

Ljupčo Todorovski

Univerza v Ljubljani, Fakulteta za upravo
Institut Jožef Stefan, Odsek za tehnologije znanja (E8)

Maj 2019

Prosojnice uporabljajo slikovne materiale iz članka DOI:10.1126/science.aab3050
in vadbice snap.stanford.edu/deepnetbio-ismb/

Pregled vsebine

Hitro učenje razumljivih konceptov

- Širši nabor nalog strojnega učenja
- Strojno učenje Bayesovih programov

Strojno učenje in omrežja

- Vložitve vozlišč omrežja
- Rekonstrukcija omrežij

Omejitvi strojnega učenja

Učenje iz peščice primerov

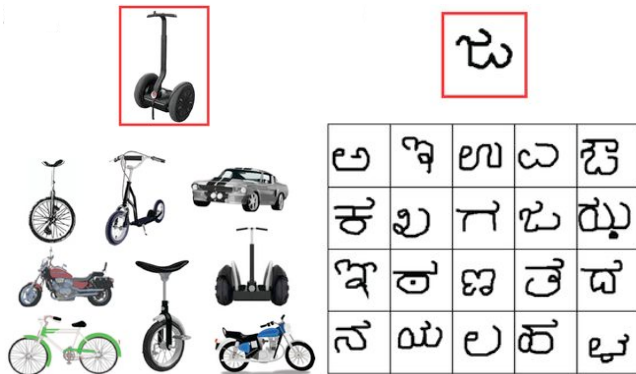
- Strojno učenje rabi tisoče ali milijone primerov
- Ljudje se lahko naučimo iz enega primera

Učenje strukture primerov in konceptov

- Strojno učenje večinoma uporablja vektorski zapis primerov
- Vektorski zapis nima strukture
- Ljudje razumemo (hierarhično) strukturo primerov
- Ljudje se iz strukture primerov naučimo koncepte

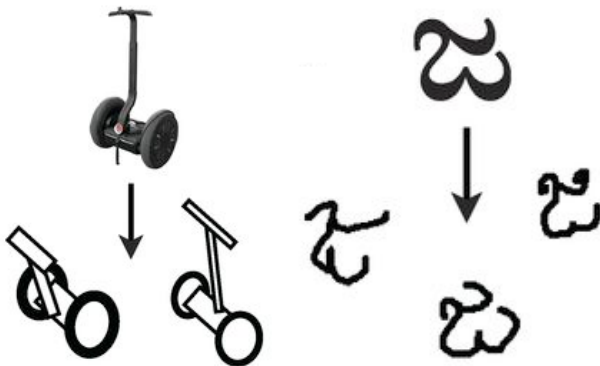
Klasična naloga strojnega učenja: klasifikacija

Razvrščanje primerov v razrede



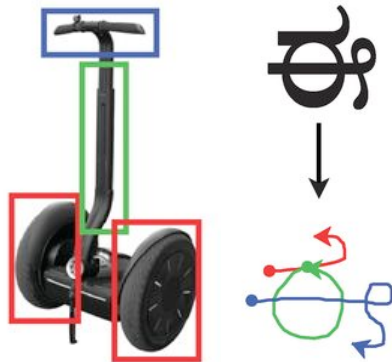
Tvorjenje novih primerov

Iz obstoječih primerov tvorimo nove primere



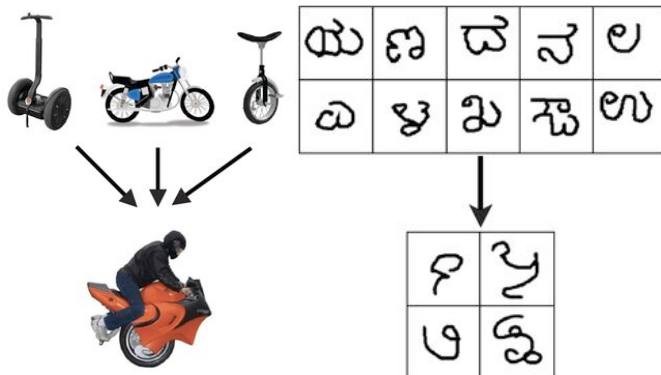
Identifikacija konceptov

Razčlenjevanje primerov v elemente in relacije



Načrtovanje/Kreiranje novih konceptov

Iz obstoječih konceptov tvorimo nove koncepte



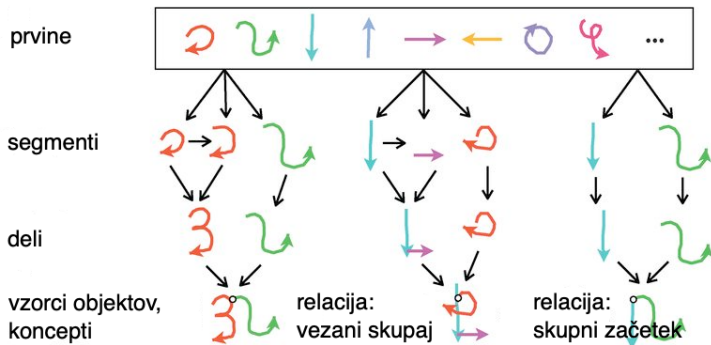
Razlika med konceptom in primerom

splošni koncept/tip črke trije specifični primeri (tipa) črke



Hierarhično urejena knjižnica konceptov (tipov črk)

Od specifičnih prvin do splošnih konceptov (delov sestavljenih z relacijami)



Koncept objekta in verjetnostna distribucija

Koncept objekta ψ je trojica (κ, S, R)

- κ je število delov (potez) objekta (črke)
- $S = \{S_1, S_2, \dots, S_\kappa\}$ je množica delov
- $R = \{R_1, R_2, \dots, R_\kappa\}$ je množica relacij med deli

$$P(\psi) = P(\kappa) \prod_{i=1}^{\kappa} P(S_i) P(R_i | S_1, S_2, \dots, S_{i-1})$$

Bayesov program za tvorjenje konceptov

function GenerateType

$\kappa \leftarrow P(\kappa)$

for $i = 1$ **to** κ **do**

$n_i \leftarrow P(n_i | \kappa)$

$S_i \leftarrow \text{GenerateStroke}(i, n_i)$

$\xi_i \leftarrow P(\xi_i)$

$R_i \leftarrow P(R_i | \xi_i, S_1, \dots, S_{i-1})$

$\psi = (\kappa, \{S_1, \dots, S_\kappa\}, \{R_1, \dots, R_\kappa\})$

return ψ

Vzorči število delov (potez)

Vzorči število segmentov poteze i

Vzorči potezo S_i iz n_i segmentov

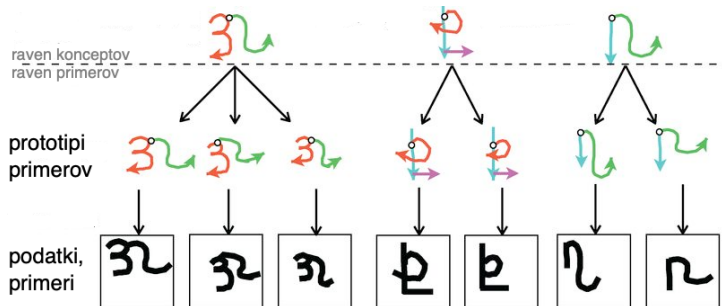
Vzorči relacije do prejšnjih potez

Vzorči podrobnosti relacij

Generativni model

Hierarhično urejena knjižnica primerov (črk)

Od splošnih konceptov (tipov črk) do primerov (končnih, dejanskih črk)



Bayesov program za tvorjenje primerov črk

function GenerateExample(ψ)

for $i = 1$ **to** κ **do**

$s_i = P(s_i | S_i)$

$l_i = P(l_i | R_i, t_1, \dots, t_{i-1})$

$t_i = f(l_i, s_i)$

$A \leftarrow P(A)$

$I \leftarrow P(I | t_1, \dots, t_\kappa, A)$

return I

Vzorči varianco hitrosti vlečenja potez

Vzorči začetno pozicijo poteze i

Sestavi trajektorijo poteze

Vzorči affino transformacijo

Vzorči sliko primera

$$\psi \equiv (\kappa, \{S_1, S_2, \dots, S_\kappa\}, \{R_1, R_2, \dots, R_\kappa\})$$

Strojno učenje: prvine

Standardizacija posnetkov risanja črk

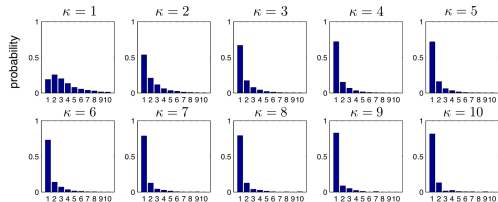
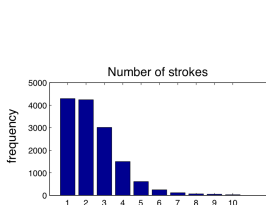
- Vzorčenje posnetkov na 50 ms
- Premik pisala manj kot eno piko označen kot pavza
- Kandidati za prvine so segmenti med pavzami

Od kandidatov do prvin

- Normalizacija razdalje med dvema zaporednima lokacijama pisala: razdalja je vedno ena pika
- Tako dobimo obliko prvine neodvisno od hitrosti premikanja
- Razvrščanje kandidatov v skupine, vsaka predstavlja eno prvino

Strojno učenje: verjetnostne porazdelitve

Število delov (potez) in število segmentov



Začetna pozicija potez glede na vrstni red

Start position for strokes in each position



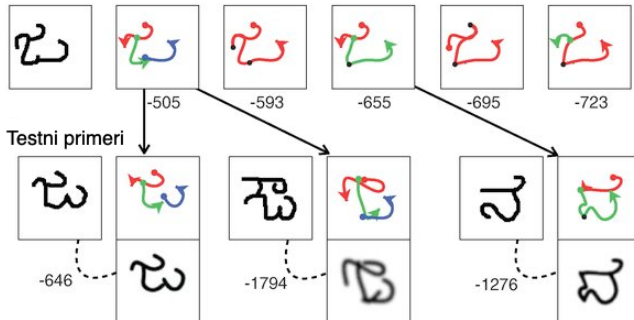
1

2

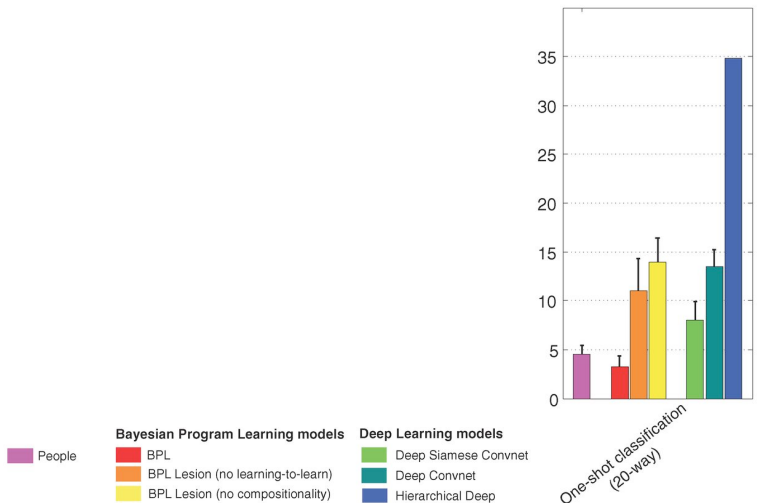
≥ 3

Strojno učenje: verjetnosti razčlenjevanja (konceptov)

Učni primer s petimi najverjetnimi razčlenjevanji (koncepti)

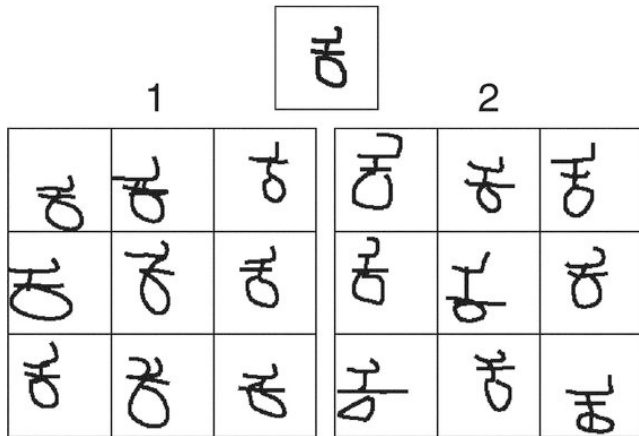


Klasifikacija: klasifikacijska napaka

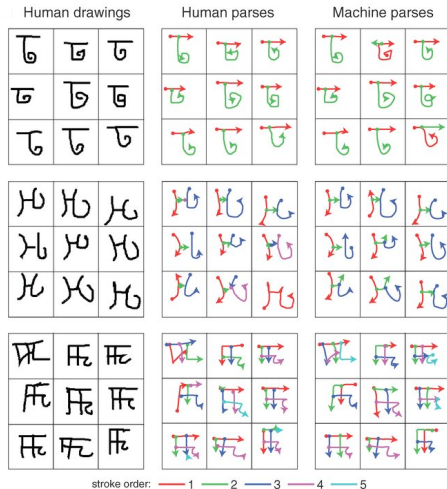


Tvorjenje novih primerov: primerjava s človekom

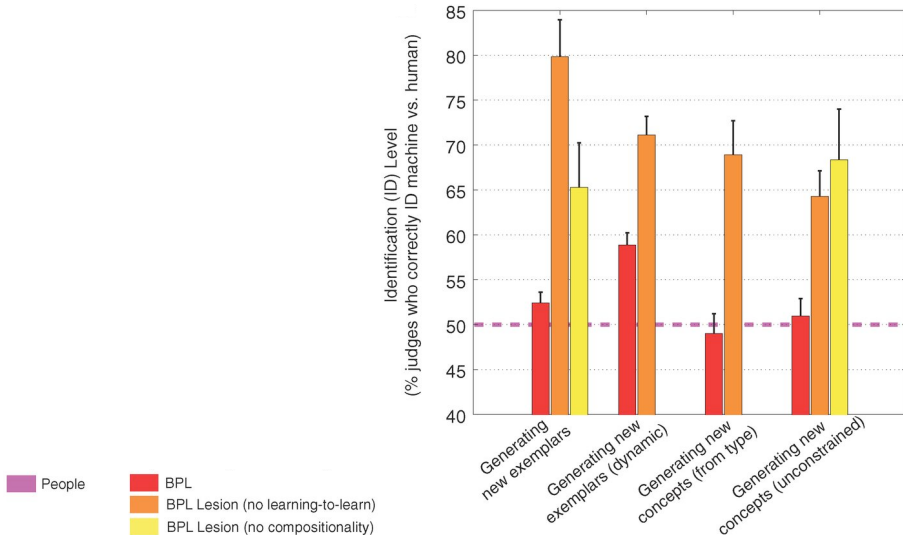
Človek na levi (1).



Identifikacija konceptov: primerjava s človekom



Tvorjenje novih primerov in konceptov: Turingov test



Literatura

(Lake in ost. 2015), DOI:10.1126/science.aab3050
Human-level concept learning through probabilistic program induction

Dve nalogi učenja v omrežjih

- Učenje v vozliščih omrežja
- Rekonstrukcija omrežij iz časovnih vrst

Definicija naloge

Podano

- Omrežje (graf) z vozlišči V in povezavami E
- Spremenljivke X_1, X_2, \dots, X_p, Y , ki odražajo lastnosti vozlišča $v, \forall v \in V$

Najdi napovedni model za napovedovanje vrednosti Y , ki

- Velja v vsakem vozlišču podanega omrežja
- Upošteva strukturo omrežja

Zakaj je naloga zahtevna?

Vozlišča imajo različno število sosedov

- Vsako vozlišče v ima različno število sosedov $d(v)$
- Če bi upoštevali še vrednosti sosedov, bi model za vozlišče v učili iz $(d(v) + 1)p$ napovednih spremenljivk

Kaj pa če hočemo upoštevati sosede od sosedov?

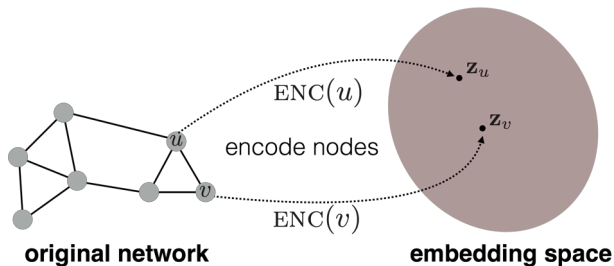
- Število napovednih spremenljivk strmo narašča
- Še bolj se spreminja za različna vozlišča

Kako lahko naslovimo te težave?

- 1 Relacijsko učenje (domača naloga)
- 2 Vložitve in avtokodirniki

Ideja vložitev

$$\text{similarity}(u, v) \approx z_u^T z_v, \quad z_u = \text{ENC}(u), \quad z_v = \text{ENC}(v)$$



Učenje vložitev vozlišč

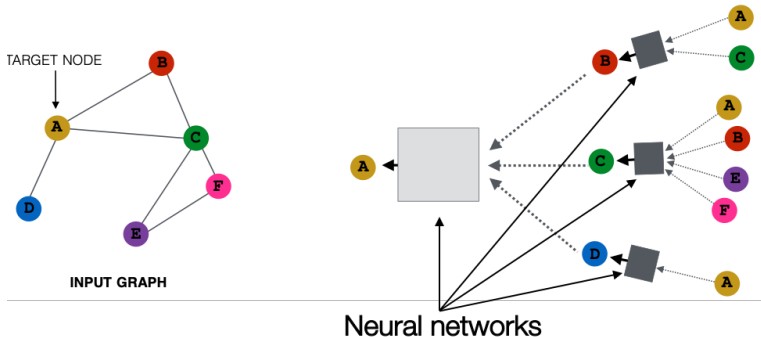
- 1 Izberi koder oziroma obliko funkcije *ENC*
- 2 Izberi mero podobnosti vozlišč v omrežju *similarity*
- 3 Minimiziraj vrednost $\sum_{(u,v) \in V \times V} \|similarity(u, v) - z_u^T z_v\|$

Kot je pri vložitvah običaj, pod 1 izberemo nevronske mreže.

Osnovna ideja: združevanje okolic

Vložitve na osnovi okolice vozlišča

- Vhodni sloj: spremenljivke v vozliščih
- Skriti sloj k : vložitev po k -tih slojih združitve okolic
- Za vsakim skritim slojem povprečno vrednost spremenljivk



Bolj formalno: struktura UNM in ciljna funkcija

$$\begin{aligned}h_v^{(0)} &= \mathbf{x}_v \\h_v^{(l)} &= \phi \left(\mathbf{w}_k^{(l)} \frac{1}{|N_v|} \sum_{u \in N(v)} h_u^{(l-1)} + \mathbf{B}_k^{(l)} h_v^{(l-1)} \right) \\z_v &= h_v^{(L)}\end{aligned}$$

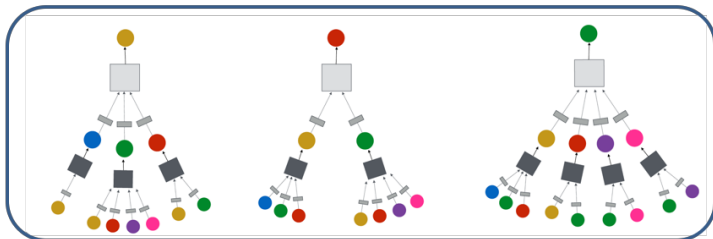
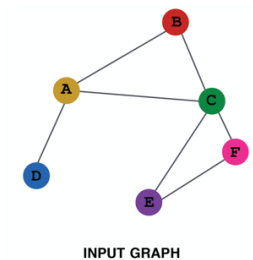
$N(v)$ je okolica vozlišča v

$$N(v) = \{u : (v, u) \in E\}$$

Ciljna funkcija za optimizacijo

$$E = \sum_{v \in V} (y_v \log \hat{y}_v + (1 - y_v) \log(1 - \hat{y}_v))$$

Po ena nevronska mreža za vsako (ciljno) vozlišče



Definicija naloge

Podano

- Množica vozlišč V v omrežju z neznanimi povezavami E
- Spremenljivka X_v v vsakem vozlišču, ki spreminja vrednost skozi čas
- Časovne vrste opazovanih vrednosti $X_v(t)$

Najdi povezave E med vozlišči iz V

Ki najbolj pojasnjujejo opazovane spremembe spremenljivk X_v .

Osnovna predpostavka

Visoka korelacija med

- 1 Verjetjem prisotnosti povezave $e_{(u,v)}$
- 2 Korelaciji med $X_u(t)$ in $X_v(t)$

Iz verjetij do povezav z odločitvenim pragom θ

$$(u, v) \in E \equiv e_{(u,v)} > \theta$$

Dve odprti vprašanji

- Kako lahko merimo korelacijo med $X_u(t)$ in $X_v(t)$?
- Kako izberemo vrednost praga θ ?

Naivni pristop: Pearsonova korelacija

$$r_P(X_u, X_v) = \frac{\mathbb{E}[(X_u - \mathbb{E}[X_u])(X_v - \mathbb{E}[X_v])]}{\mathbb{E}[(X_u - \mathbb{E}[X_u])^2] \mathbb{E}[(X_v - \mathbb{E}[X_v])^2]}$$

Ne upošteva morebitnega vpliva vseh ostalih vozlišč.

Merjenje korelacije s strojnim učenjem

$$M_v : \hat{X}_v(t) = M_v(X_1(t), X_2(t), \dots, X_{v-1}(t), X_{v+1}(t), \dots, X_{|V|}(t))$$

Učimo se model M_v za vsako vozlišče $v \in V$ iz podatkovne množice S_v

- Primeri S_v ustrezajo posameznim časovnim točkam t
- Napovedne spremenljivke so S_v so $X_u(t-1) : \forall u \in V : u \neq v$
- Ciljna spremenljivka je $X_v(t)$

Iz M_v preberemo napovedno moč spremenljivk $p_{(u,v)}$

Napovedna moč spremenljivke $X_u(t-1)$ je enaka verjetju $e_{(u,v)}$.

Kako se določimo vrednost θ ?

Podobno kot odločitveni prag pri klasifikacijskih modelih

- Vsak prag določa prisotnost in odsotnost posameznih povezav
- Če poznamo strukturo omrežja lahko vstavimo prag v prostor ROC
- Nato se odločamo v tem prostoru ali narišemo krivuljo ROC za različne vrednosti praga θ

Literatura

(Kipf in Welling 2016), arXiv:1609.02907

Semi-supervised classification with graph convolutional networks