

JESENSKI ISPITNI ROK

DUBOKO UČENJE 1

29.08.2024.

1. Klasifikacijski konvolucijski model za 1D podatke

Slojevi:

- a) konvolucijski bez nadopunjavanja (2 jezgre, korak 1, ReLU aktivacija)
- b) maxpool
- c) FC s 2 izlaza
- d) Softmax

Početni parametri konvolucijskog sloja:

$$\mathbf{w}_{1A} = [-0.5 \ 0.5], \ b_{1A} = 0$$

$$\mathbf{w}_{1B} = [0.5 \ 0.5], \ b_{1B} = -0.5$$

Početni parametri FC sloja:

$$\mathbf{W}_2 = I, \ \mathbf{b}_2 = 0$$

- a) Jednadžbe modela
- b) Odredi izlaz za ulaze $x_1 = [0 \ 1 \ 1 \ 0]$ $x_2 = [1 \ 1 \ 1 \ 1]$ (konvolucija se provodi korelacijom s nezrecaljenim filtrima)
- c) Odredite gradijente s obzirom na sve konvolucijske parametre u podatku x_1 ako je zadano $y=1$. Pretpostavite da model prediktira $P(y=1|x)$

2. $f(x) = x^4 - 2x^3 + 5$ Raspišite dvije iteracije optimizacije gradijentnim spustom:

$$x_0 = -1, \ \varepsilon = 0$$

Bi li postupak drugačije konvergirao za različite inicijalizacije i ako je korak dovoljno mali? Ako da, navedite rješenja koja bi gradijentni spust mogao naći. Bi li vaš postupak bolje konvergirao s drugim algoritmom? Ako da, kojim? Skicirajte ga i navedite zašto.

3. Razmatramo model kao u 2. labosu. Napišite implementaciju

SoftmaxCrossEntropyWithLogits prema sučelju Layer. Uključite metode forward, backward_inputs, backward_parameters

4. Razmatramo jezični model na razini slova. Model predviđa sljedeće slovo temeljem prethodnih. Rješavamo običnim povratnim modelom koji u povratnoj vezi ima prijenosnu fju \tanh i na izlazu softmax. Na ulaz dovodimo niz znakova kodiranih jednojedinичnim kodom. Model treba moći raditi s 80 različitih znakova. Skriveni sloj ima 200 dimenzija.

Odredite dimenzije svih parametara povratnog modela.

5. Razmatramo obični povratni model koji u povratnoj vezi ima prijenosnu funkciju $\min(1, x)$, a na izlazu identitet bez nelinearnosti. Zadatak je prebrojati jedinice i nule u binarnom broju. Na ulaz dolazi proizvoljno dugačak niz jedinica i nula u jednojedinичnoj reprezentaciji. $W_{hh} = I$, početno stanje $h_0 = [0 \ 0]$, a za preostale parametre modela (za elemente W_{xh}, W_{hy}, b_h, b_o) vrijedi $|t_i| \leq 1, t_i \in \mathbb{Z}$

Odredite vrijednosti ostalih parametara modela.

6. Razmatramo model za ugrađivanje 2D podataka u 1D metrički prostor

Slojevi:

h1: FC s 2 izlaza i ReLU

s2: FC s 1 izlazom

Model učimo trojnim gubitkom $L(a, p, n) = \max(d(a, p) - d(a, n) + m, 0)$

$$W_1 = I, \quad W_2 = [0.4 \ 0.6], \quad b_1 = 0, \quad b_2 = 0$$

d odgovara L1 metrici, $m=1$

a) Jednadžbe modela ako na ulaz dolazi trojka (x_a, x_p, x_n)

b) Iznos gubitka za ulaz $(x_a, x_p, x_n) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

c) Odredite gradijenta gubitka s obzirom na sve parametre modela