

Duboko učenje

završni ispit

1. Binarni klasifikacijski model zadan je nizom slojeva:

- Potpuno povezani sloj ($Wx + b$) s parametrima W_1 i b_1
- Normalizacija nad grupom (eng. batchnorm) bez afixne transformacije
- Aktivacija zgloboicom
- Potpuno povezani sloj ($Wx + b$) s parametrima W_2 i b_2
- Aktivacija sigmoidom

Provedite jedan korak učenja modela gubitkom binarne unakrsne entropije za minigrupu sastavljenu od podataka $x_1 = [1 \ 1]^T$ i $x_2 = [-1 \ 3]^T$ s odgovarajućim točnim oznakama $y_1 = y_2 = 1$.

Početne vrijednosti parametara modela su: $W_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $b_1 = [0 \ 0]^T$, $W_2 = [1 \ 1]$ te $b_2 = -1$. Za izračun statistika koristite pristrane procjenitelje (hint: N u nazivniku). Prilikom unatragnog prolaza kroz sloj normalizacije nad grupom, statistike grupe promatrajte kao konstante. Pretpostavite da je stopa učenja 1.

2. Razmatramo sustav za pretraživanje slika. Zadana su dvodimenzionalna metrička ugrađivanja skupa podataka u obliku matrice X , te njihovi identiteti y . Na ulaz sustava dolazi upit identiteta $y_i = 1$, za kojeg smo izračunali vector ugrađivanja $x_i = [1 \ 1]^T$.

- Izračunajte sličnost podataka x_i s ostalim podacima iz zadanog skupa, ako za mjeru sličnosti koristimo recipročnu vrijednost L1 udaljenosti vektora ugrađivanja
- Koristeći izračunate sličnosti skicirajte krivulju preciznosti i odaziva za zadani ispit i odredite prosječnu preciznost (AP).

$$X = \begin{bmatrix} 1 & 0 \\ 3 & -1 \\ -0.5 & 0.5 \\ 1 & 4 \end{bmatrix} \text{ i } y = [1 \ 1 \ 0 \ 1]^T$$

3. Razmatramo konvolucijski klacifikacijski model zadan nizom slojeva:

- 1D konvolucija bez nadopunjavanja s korakom 1, s dvije jezgre $w_{11} = [1 \ 0 \ -1]$ i $w_{12} = [1 \ 1 \ -1]$ bez pomaka
- Aktivacija zgloboicom
- Globalno sažimanje prosjekom
- Potpuno povezani sloj s parametrima $W_2 = \begin{bmatrix} \ln 2 & 0 \\ 0 & \ln 2 \end{bmatrix}$ i $b_2 = [0 \ \ln 2]^T$
- Aktivacija softmaksom

Na ulazu modela nalazi se podatak $x = [4 \ -1 \ 2 \ 0 \ 1]$ koji pripada razredu $y = 1$.

- a) Napravite unaprijedni prolaz kroz model za zadani podatak x
- b) Izračunajte gradijent gubitka unakrsne entropije po parametrima konvolucijskog sloja

4. Razmotrimo aktivacijsku funkciju $\text{softplus}(x) = \ln(1 + e^x)$.

- a) Skicirajte graf zadane funkcije i njene derivacije
- b) Napišite jakobijan aktivacije zadanom funkcijom za 4-dimenzijski ulaz
- c) Za zadanu aktivacijsku funkciju implementirajte sučelje *Layer* po uzoru na 2. Laboratorijsku vježbu. Implementirani sloj treba moći raditi s tenzorima različitih redova i dimenzija. Ulaz u funkciju *forward* će biti tipa *np.ndarray*.

5. Oblikuj model obične povratne neuronske mreže s prijenosnom funkcijom zglobnice za evidentiranje promjene stanja bankovnog računa. Za taj zadatak koristit će te tradicionalni vector skrivenog stanja, pri čemu će se u pravom elementu pamtići ukupan iznos dostupan za isplatu, na drugom elementu će biti zapisano dopušteno prekoračenje, dok treći element služi kao zastavica koja provjerava jeli račun u “minusu” (0 ako je, inače pozitivan broj). Početno skriveno stanje računa je: $h^{(0)} = [10000 \ 5000 \ 765]^T$.

- a) Odredite preostale parametre (W_{hh} , W_{xh} , b_h) kako bi mreža obavljala navedenu funkcionalnost
- b) Provedite unaprijedni prolaz s navedenim parametrima za iduće ulaze:
 $x^{(1)} = [1000 \ 0]^T$, $x^{(2)} = [1000 \ 3500]^T$, $x^{(3)} = [0 \ 4000]^T$
- c) Razmotrimo pristup istom problemu sa dvoslojnom LSTM mrežom, uz jednake dimenzionalnosti ulazne reprezentacije i skrivenog stanja. Odredite ukupan broj parametara takve mreže.