

## Duboko učenje

### međuispit

#### 1. Razmatramo klasifikacijski model s arhitekturom:

- konvolucijski sloj bez nadopunjavanja: 16 jezgara  $5 \times 5$ , korak 1, aktivacija ReLU;
- sažimanje maksimumom  $2 \times 2$ ;
- konvolucijski sloj bez nadopunjavanja: 32 jezgre  $5 \times 5$ , korak 1, aktivacija ReLU;
- sažimanje maksimumom  $2 \times 2$ ;
- pretvaranje u vektor;
- potpuno povezani sloj dimenzije 10 te aktivacija softmax;

Na ulaz mreže dovodimo RGB slike veličine  $28 \times 28$  piksela.

- (a) Napišite dimenzije ulaznih i izlaznih tenzora za svaki sloj.
- (b) Koliko ukupno parametara ima zadani model?
- (c) Procijenite memorijsko zauzeće u trenutku nakon izračuna gubitka unakrsne entropije. Postupak učenja zadanog modela provodimo s veličinom mini-grupe 8. Tip podataka je 32-bitni float. Pomoć: u memorijsko zauzeće ubrojiti sve u tome trenutku pohranjene tenzore i parametre modela.
- (d) Kolika su receptivna polja aktivacija u prvom sloju sažimanja odnosno aktivacija u potpuno povezanom sloju? Pomoć: receptivno polje računamo s obzirom na piksele ulazne slike.

#### 2. Razmatramo regresijski model zadan nizom slojeva:

- potpuno povezani sloj s matricom težina  $\mathbf{W}_1$  i pomakom  $\mathbf{b}_1$ ;
- normalizacija po grupi, bez pomičnog prosjeka, s fiksnom afinom transformacijom  $\gamma=1$ ,  $\beta=0$  (identitet);
- aktivacija zglobnicom (ReLU);
- potpuno povezani sloj s matricom težina  $\mathbf{W}_2$  i pomakom  $b_2$ .

Zadani su podaci  $\mathbf{X}$  i parametri modela:

$$\mathbf{X} = \begin{bmatrix} 1 & -3 \\ 5 & 3 \\ 2 & 4 \\ 2 & 0 \end{bmatrix}, \mathbf{W}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{W}_2 = [1 \quad -1], b_2 = 1.$$

- (a) Provedite unaprijedni prolaz mreže u fazi treniranja za prvu mini-grupu, ako koristimo veličinu mini-grupe 2 sa slijednim uzorkovanjem.
- (b) Provedite unaprijedni prolaz mreže u fazi eksploatacije za prvi primjer.

#### 3. Potrebno je projektirati program koji na ulazu dobiva binarnu matricu dimenzija $2 \times 3k$ , pri čemu je $k$ prirodan broj. Jedan primjer takve matrice je sljedeći:

$$X = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Gornji redak matrice sadrži niz 3-bitnih binarnih brojeva. Donji redak sadrži niz 3-bitnih jednodijeljeničnih kodova koji određuju način obrade. Zadaća programa je obraditi binarne brojeve u skladu s odgovarajućim jednodijeljeničnim kodom te odrediti rezultat u dekadskom sustavu. Značenja jednodijeljeničnih kodova su:

100 – identitet (rezultat je jednak ulazu);

010 – zbrajanje s tri;

001 – postavljanje na nulu.

Za dani ulazni primjer  $X$ , na izlazu treba proizvesti matricu  $\begin{bmatrix} 10 & 0 & 5 \end{bmatrix}$ .

Prvi broj u gornjem retku ulaza je 7, a ispod njega je zadana operacija 010, te je zato na izlazu broj 10. Nakon toga, u gornjem retku dolazi broj 3, a ispod njega je zadana opcija 001; zato je na izlazu 0. Zadnji broj gornjeg retka je 5, a ispod njega je zapisana opcija 100, te se zato na izlaz propušta broj 5.

Konstruirajte model koji rješava opisani problem koristeći neke od sljedećih slojeva: množenje matricom, softmaks, sigmoida, ReLU, konvolucija, te sažimanje prosjekom ili maksimumom. Ako ste odabrali slojeve s težinama, napišite i vrijednosti tih težina.

4. Razmatramo sloj dubokog modela koji ulazni vektor  $\mathbf{p}$  transformira u izlazni vektor  $\mathbf{q}$  koristeći polinomnu transformaciju koja se može opisati jednadžbom  $q_i = w_0 + w_1 p_i + w_2 p_i^2 + \dots + w_n p_i^n$ . Odredite Jakobijan gubitka po ulazu i Jakobijan gubitka po parametrima ako je poznat Jakobijan gubitka po izlazu.

Napišite kod koji bi omogućio uklapanje sloja u duboki model proizvoljne složenosti. Sloj izrazite razredom koji nasljeđuje sučelje `Layer` iz prvog zadatka druge laboratorijske vježbe te implementira sljedeće metode: `forward(self, inputs)`, `backward_inputs(self, grads)`, te `backward_parameters(self, grads)`. Vaš razred treba kroz konstruktor primiti parametar `n` koji određuje stupanj polinoma.

Izvedba treba osigurati da dimenzionalnost izlaznog vektora bude jednaka dimenzionalnosti ulaznog vektora.

5. Razmatramo klasifikacijski konvolucijski model koji je prilagođen obradi vektora s pet elemenata. Arhitektura modela je:

- konvolucijski sloj bez nadopunjavanja, jezgra  $w_1 = \begin{bmatrix} -0.25 & 0.5 & 0.25 \end{bmatrix}$ , pomak  $b_1 = -1/4$ , korak 1, aktivacija ReLU;
- potpuno povezani sloj, jezgra  $w_2 = \begin{bmatrix} -1 & 1 & 1 \end{bmatrix}$ , pomak  $b_2 = -1$ , sigmoidna aktivacija.

Zadatci:

- Odredite dimenzije svih aktivacijskih tenzora te broj parametara modela;
- odredite predikciju modela za podatak  $x = \begin{bmatrix} 1 & 0 & 6 & 1 & 1 \end{bmatrix}^\top$  te odredite gubitak (negativnu log izglednost) ako je točna izglednost podatka jednaka nuli;
- odredite gradijente gubitka s obzirom na sve parametre modela;
- odredite indeks dimenzije ulaznog podatka čije povećanje maksimalno smanjuje gubitak (BONUS).