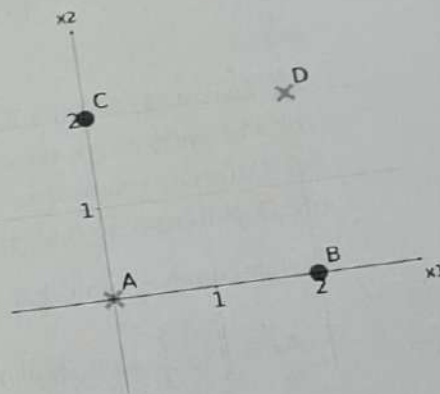


## Duboko učenje međuispit

1. Dizajnirajte dvoslojni potpuno povezani model koji rješava problem sa slike. Zadani problem sadrži četiri podatka koje je potrebno klasificirati u dva razreda: kružić i križić. Aktivacijska funkcija skrivenog sloja neka bude zglobnica, a izlaznog sloja sigmoida. Napišite sve parametre modela i pokažite da Vaš model ispravno klasificira sve podatke pod pretpostavkom da Vaša granica klasa definira vjerojatnost od 50%. Pomoć: razmotrite duboki model nad varijablom  $z = x_2 - x_1$ .



2. Napišite izravnu implementaciju funkcije koja obavlja unaprijedni prolaz i računa gubitak za višerazrednu logističku regresiju koristeći biblioteku numpy. Funkcija na ulazu prima jedan podatak  $x$  dimenzija (D, 1), prirodni broj  $y$  iz intervala [1, C] koji označava točni razred, te ostale potrebne parametre odgovarajućih dimenzija (kojih?). Identificirajte dijelove vašeg koda u kojima može doći do numeričkih pogrešaka. Napišite robusnu implementaciju koja rješava navedene probleme. Uputa: obratite pažnju na mogući preljev u funkcijama exp i log.
3. Zadana je sljedeća inicijalizacija modula, te odgovarajući unaprijedni prolaz. Na ulazu se nalazi jedna RGB slika dimenzija  $32 \times 32$ .

```
conv1 = nn.Conv2d(
    __, 16, kernel_size=7, stride=1, padding=0)
conv2 = nn.Conv2d(
    __, 24, kernel_size=3, stride=1, padding=1)
maxpool1 = nn.MaxPool2d(
    kernel_size=__, stride=__, padding=0)
conv3 = nn.Conv2d(
    __, 32, kernel_size=7, stride=1, padding=0)
conv_skip = nn.Conv2d(
    __, __, kernel_size=3, stride=__, padding=__)
avgpool1 = nn.AvgPool2d(
    kernel_size=2, stride=2, padding=0)
fc1 = nn.Linear(__, 256)
fc2 = nn.Linear(__, 4)
```

```
# img.shape == (1, C, H, W)
x = torch.relu(conv1(img))
x = torch.relu(conv2(x))
x = maxpool1(x)
assert(x.shape[-1] == 22)
assert(x.shape[-2] == 22)
x = torch.relu(conv3(x))
skip = torch.relu(conv_skip(img))
x = x + skip
x = avgpool1(x)
x = x.view(1, -1)
x = torch.relu(fc1(x))
x = fc2(x)
```

Zadaci:

- (a) Dopunite inicijalizaciju zadanih slojeva hiperparametrima koji nedostaju, tako da se unaprijedni prolaz može izvršiti bez grešaka.
- (b) Napišite broj parametara svakoga sloja, te dimenzije tenzora  $x$  nakon izvršene svake li nije unaprijednog prolaza. Podrazumijevana vrijednost argumenta bias u konvolucijskim i potpuno povezanim slojevima je True.
- (c) Odredite receptivno polje izlaza sloja maxpool1.

4. Razmatramo klasifikacijski model u tri klase. Učenje provodimo na podacima koji su označeni indeksima  $y \in \{0, 1, 2, 3\}$ . Indeksi 0, 1, 2 označavaju razrede unutardistribucijskih podataka. Na takve podatke primjenjujemo gubitak negativne log izglednosti  $\mathcal{L}_{\text{NLL}}$ . Indeks 3 označava tzv. negativne podatke za učenje. To su izvandistribucijski podatci koji ne pripadaju niti jednom od 3 poznata razreda. Na takve podatke primjenjujemo gubitak KL divergencije  $\mathcal{L}_{\text{KL}}$ . Takav je model opisan sljedećim jednadžbama:

$$\begin{aligned} s &= \mathbf{W} \cdot \mathbf{x} + \mathbf{b} \\ p &= \text{softmax}(s) \\ \mathcal{L} &= \begin{cases} \mathcal{L}_{\text{NLL}}(p_y), y \in \{0, 1, 2\} \\ \mathcal{L}_{\text{KL}}(p, \text{uniform}), y = 3; \end{cases} \quad \mathcal{L}_{\text{KL}}(p, q) = - \sum_i q_i \log\left(\frac{p_i}{q_i}\right) \end{aligned}$$

Zadaci:

- Izvedite gradijente KL divergencije po logitima  $\frac{\partial \mathcal{L}_{\text{KL}}}{\partial s}$  za zadanu proizvoljnu distribuciju  $q$ .
- Pojasnite odnos između gubitka negativne log izglednosti i gubitka unakrsne entropije.
- Pojasnite odnos između gubitka unakrsne entropije i KL divergencije.
- Pojednostavite funkciju gubitka koristeći prethodna razmatranja.
- Izračunajte gubitak i gradijente po parametrima u podacima  $\mathbf{x}_1 = \begin{bmatrix} -5 \\ 2 \end{bmatrix}$ ,  $y_1 = 0$  i  $\mathbf{x}_2 = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$ ,  $y_2 = 3$ , ako su zadani parametri  $\mathbf{W} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}^T$  i  $\mathbf{b} = 0$ .

5. Razmatramo klasifikacijski model zadan nizom slojeva:

$$\mathbf{s}_1 = \text{conv1D}(\mathbf{x}, \mathbf{w}_1, b_1) \quad (1)$$

$$\mathbf{h}_1 = \text{ReLU}(\mathbf{s}_1) \quad (2)$$

$$\mathbf{g}_1 = \text{concat}(\mathbf{h}_1, \mathbf{x}) \quad (3)$$

$$\mathbf{s}_2 = \mathbf{w}_2^T \mathbf{g}_1 \quad (4)$$

$$p = \sigma(\mathbf{s}_2) \quad (5)$$

Izračunajte gradijente gubitka binarne unakrsne entropije obzirom na sve parametre modela, ako je zadan ulazni podatak  $\mathbf{x} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$ , parametri modela  $\mathbf{w}_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ ,  $b_1 = 0$ ,  $\mathbf{w}_2 = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \end{bmatrix}^T$ , te točna oznaka podatka  $y = 0$ .

Napomena: operacija concat označava konkatenciju dva vektora dimenzija  $n_1$  i  $n_2$ ; rezultat te operacije je novi vektor s  $n_1 + n_2$  elemenata, gdje prvih  $n_1$  elemenata odgovara prvome vektoru, a preostalih  $n_2$  elemenata drugome vektoru.