Napredni algoritmi i strukture podataka

1. jesenski rok

a.g. 2019/2020

Zadatak 1 (4 2)

OZNAKE: **struktura** **stablo** **RB**

Opišite, sažeto, kako u RB stablu nastaje stanje koje modeliramo pomoću dvostruko crnog čvora. Drugim riječima, objasnite kada se u RB stablu pojavljuje dvostruko crni čvor.

Odgovor 1

BODOVI: **4**

To se događa prilikom brisanja. Kako u RB stablu brišemo kopiranjem, onda ćemo čvor kojeg brišemo kopirati u neki list, čvor u kojeg smo upisali ono što brišemo premjestiti (kopirati) u čvor kojeg brišemo. Tada se može dogoditi da je taj čvor kojeg je bilo lako obrisati bio crn, a ako je i njegov roditelj crn, onda dolazimo do stanja dvostruko crnog čvora.

Zadatak 2 (10)

OZNAKE: **neuronska mreža** **backpropagation** **adaline** **tanh**

Potpuno povezana, unaprijedna (feedforward) troslojna neuronska mreža (ANN; *Artificial Neural Network*) ima strukturu $2 \times 3 \times 2$, pri čemu je sloju tangens hiperbolni (tanh), dok je u izlaznom sloju aktivacijska funkcija za izlaz 1 sigmoid, a za izlaz 2 je Adaline. Provedite prvi korak uvježbavanja te mreže (jednom osvježiti sve parametare) algoritmom koračnog uvježbavanja (*on-line learning*) ako se podatci za uvježbavanje uzimaju redom iz sljedeće tablice.

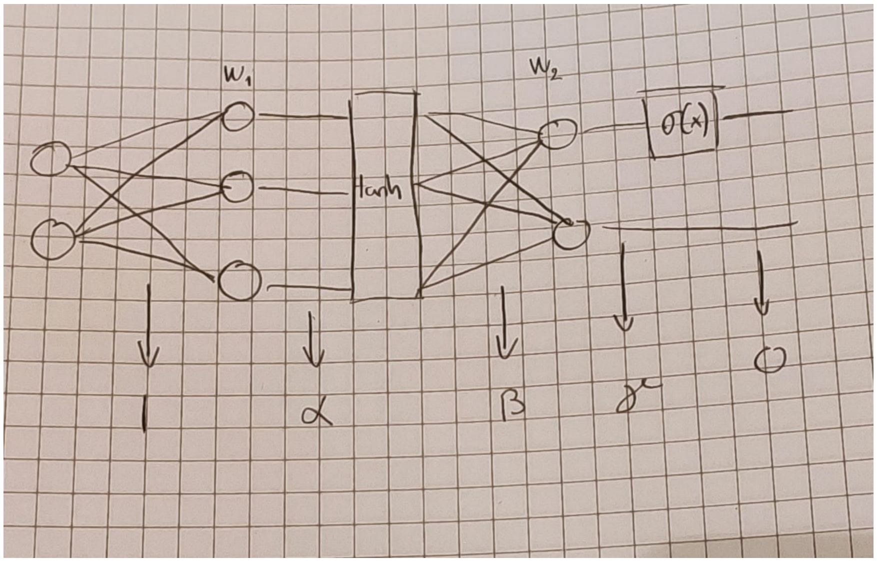
ulaz 1	ulaz 2	izlaz 1	izlaz 2
-1	3	0.4	-2
-1	6	0.2	-6
-9	4	-0.4	8
5	-3	0.4	-9

Početne vrijednosti svih parametara mreže postavite na nula, a zatrebaju li Vam još neke veličine, pridijelite im vrijednosti po vlastitom nahođenju, samo jasno navedite svoj izbor i kratko objasnite ulogu te veličine.

Napomena: $tanh(x) = 2\sigma(2x) - 1$

Odgovor 1 BODOVI: 10

Prvo skicirajmo mrežu:



Pojedine tokove u mreži označili smo s I (input), α, β, γ i O (output). Znamo da nam trebaju gradijenti za svaki tok izuzev ulaza, kao i gradijenti za svaki skup parametara. Osim označenih W_1 i W_2 imamo i b_1 i b_2. Umjesto da pišemo parcijalne derivacije, pisat ćemo gradijent prefiksiran s nablom. Na primjer, umjesto $\frac{\partial \mathcal{L}}{\partial O}$ pišemo ∇O.

Dakle, potrebni su nam sljedeći gradijenti:

- ∇O, $\nabla \gamma$, $\nabla \beta$, $\nabla \alpha$
- ∇W_2, ∇b_2, ∇W_1, ∇b_1

Prvi gradijenti su gradijenti specifičnog toka, tj. točke unutar mreže. Drugi gradijenti su gradijenti po parametrima, koje koristimo za ažuriranje parametara.

Podrazumijevana funkcija gubitka $\mathcal{L}$ je MSE:

$$\mathcal{L}(\text{target}, \text{prediction}) = \frac{1}{2}\left(\text{target} - \text{prediction}\right)^2 \tag{1}$$

Prvi gradijent kojeg možemo izračunati je ∇O. On je derivacija gubitka po izlazu (predictionu) je

$$\nabla O = prediction - target \tag{2}$$

Ako ovo pretvorimo u matrični oblik, dobivamo:

$$\nabla O = \begin{bmatrix} O_0 - Y_0 \\ O_1 - Y_1 \end{bmatrix} \tag{3}$$

nazovemo li prediction vektor $\vec{O}$, a target vektor $\vec{Y}$.

S obzirom na to da je $\nabla\gamma$ kompozitna funkcija, morat ćemo $\frac{\partial O}{\partial \gamma}$ raspisati matrično:

$$\frac{\partial O}{\partial \gamma} = \begin{bmatrix} \sigma(\gamma_0)\,(1-\sigma(\gamma_0)) \\ 1 \end{bmatrix} = \begin{bmatrix} O_0\,(1-O_0) \\ 1 \end{bmatrix} \tag{4}$$

Stoga uz ulančano pravilo vrijedi

$$\nabla\gamma = \begin{bmatrix} \nabla O_0 \cdot O_0\,(1-O_0) \\ \nabla O_1 \end{bmatrix} \tag{5}$$

Sada nas zanimaju gradijenti parametara (∇W_2 i ∇b_2). Prvo trebamo $\frac{\partial \gamma}{\partial W_2}$ i $\frac{\partial \gamma}{\partial b_2}$, a oni su istog oblika kao i W_2 i b_2:

$$\frac{\partial \gamma}{\partial W_2} = \begin{bmatrix} \beta_0 & \beta_0 \\ \beta_1 & \beta_1 \\ \beta_2 & \beta_2 \end{bmatrix} \qquad \frac{\partial \gamma}{\partial b_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{6}$$

pa kad primijenimo ulančano pravilo dobivamo

$$\nabla W_2 = \begin{bmatrix} \beta_0\nabla\gamma_0 & \beta_0\nabla\gamma_1 \\ \beta_1\nabla\gamma_0 & \beta_1\nabla\gamma_1 \\ \beta_2\nabla\gamma_0 & \beta_2\nabla\gamma_1 \end{bmatrix} \qquad \nabla b_2 = \begin{bmatrix} \nabla\gamma_0 \\ \nabla\gamma_1 \end{bmatrix} \tag{7}$$

Zatim tražimo $\nabla\beta$, a to je težinska suma težina po retcima sa zadnjim gradijentom toka (u našem slučaju s γ):

$$\nabla\beta = \begin{bmatrix} W_2^{0,0}\nabla\gamma_0 + W_2^{0,1}\nabla\gamma_1 \\ W_2^{1,0}\nabla\gamma_0 + W_2^{1,1}\nabla\gamma_1 \\ W_2^{2,0}\nabla\gamma_0 + W_2^{2,1}\nabla\gamma_1 \end{bmatrix} \tag{8}$$

Slično kao i prije, trebamo pomnožiti naš gradijent elementwise s gradijentom aktivacijske funkcije. Uz zadatak smo dobili hint:

$$tanh(x) = 2\sigma(2x) - 1 \tag{9}$$

pa stoga možemo reći

$$\frac{\partial tanh(x)}{\partial x} = 4\sigma(x) \cdot^{*} (1-\sigma(x)) \tag{10}$$

intuitivno, kada gledate $tanh(x)$, to je sigmoida koja je samo duplo izdužena u visinu. Povećanje u visinu će kvadratno povećati gradijent, a $2^2 = 4$.

Sada možemo dobiti i $\nabla\alpha$. Uzevši u obzir da vrijedi

$$\frac{\partial \beta}{\partial \alpha} = 4\sigma(\alpha) \cdot^{*} (1-\sigma(\alpha)) \tag{11}$$

uz ulančano pravilo možemo pisati

$$\nabla\alpha = \begin{bmatrix} 4\sigma(\alpha_0)\,(1-\sigma(\alpha_0))\,\nabla\beta_0 \\ 4\sigma(\alpha_1)\,(1-\sigma(\alpha_1))\,\nabla\beta_1 \\ 4\sigma(\alpha_2)\,(1-\sigma(\alpha_2))\,\nabla\beta_2 \end{bmatrix} \tag{12}$$

Finalno, ponovimo sve slično kao u jednadžbi (6):

$$\frac{\partial \gamma}{\partial W_1} = \begin{bmatrix} I_0 & I_0 & I_0 \\ I_1 & I_1 & I_1 \end{bmatrix} \qquad \frac{\partial \gamma}{\partial b_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{13}$$

pa uz ulančano pravilo dobivamo

$$\nabla W_1 = \begin{bmatrix} I_0 \nabla \alpha_0 & I_0 \nabla \alpha_1 & I_0 \nabla \alpha_2 \\ I_1 \nabla \alpha_0 & I_1 \nabla \alpha_1 & I_1 \nabla \alpha_2 \end{bmatrix} \qquad \nabla b_1 = \begin{bmatrix} \nabla \alpha_0 \\ \nabla \alpha_1 \\ \nabla \alpha_2 \end{bmatrix} \tag{14}$$

Time smo izračunali sve što nam treba pa možemo krenuti na prvi korak učenja.

Forward pass

Sve težine su na 0, tj. vrijedi:

$$W_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad b_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{15}$$

$$W_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad b_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{16}$$

Kada propustimo prvi primjerak kroz mrežu, dobivamo sljedeće tokove:

$$I = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \quad \alpha = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \beta = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \gamma = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad O = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \tag{17}$$

Backward pass

Sukladno izračunatim gradijentima, pišemo:

$$\nabla O = \begin{bmatrix} 0.5 - 0.4 \\ 0 - (-2) \end{bmatrix} = \begin{bmatrix} 0.1 \\ 2 \end{bmatrix} \tag{18}$$

$$\nabla \gamma = \begin{bmatrix} 0.1 \cdot 0.5 \cdot 0.5 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.025 \\ 2 \end{bmatrix} \tag{19}$$

$$\nabla W_2 = \begin{bmatrix} 0 \cdot 0.025 & 0 \cdot 2 \\ 0 \cdot 0.025 & 0 \cdot 2 \\ 0 \cdot 0.025 & 0 \cdot 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{20}$$

$$\nabla b_2 = \begin{bmatrix} 0.025 \\ 2 \end{bmatrix} \tag{21}$$

$$\nabla \beta = \begin{bmatrix} 0 \cdot 0.1 + 0 \cdot 2 \\ 0 \cdot 0.1 + 0 \cdot 2 \\ 0 \cdot 0.1 + 0 \cdot 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{22}$$

$$\nabla \alpha = \begin{bmatrix} 4 \cdot 0.5 \cdot 0.5 \cdot 0 \\ 4 \cdot 0.5 \cdot 0.5 \cdot 0 \\ 4 \cdot 0.5 \cdot 0.5 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{23}$$

$$\nabla W_1 = \begin{bmatrix} -1 \cdot 0 & -1 \cdot 0 & -1 \cdot 0 \\ 3 \cdot 0 & 3 \cdot 0 & 3 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{24}$$

$$\nabla b_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{25}$$

Ažuriranje parametara

Pravilom

$$\theta^k = \theta^{k-1} - \eta \nabla \theta^{k-1} \tag{26}$$

ažuriramo težine uz stopu učenja 1 (tj. uz $\eta = 1$):

$$W_1' = \begin{bmatrix} 0-0 & 0-0 & 0-0 \\ 0-0 & 0-0 & 0-0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{27}$$

$$b_1' = \begin{bmatrix} 0-0 \\ 0-0 \\ 0-0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{28}$$

$$W_2' = \begin{bmatrix} 0-0 & 0-0 \\ 0-0 & 0-0 \\ 0-0 & 0-0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{29}$$

$$b_2' = \begin{bmatrix} 0-0.025 \\ 0-2 \end{bmatrix} = \begin{bmatrix} -0.025 \\ -2 \end{bmatrix} \tag{30}$$

Zadatak 3 (5 5)

oznake: **dinamičko programiranje**

Koje su tvrdnje istinite?

a) Dinamičko programiranje je posebna vrsta (grana) linearnog programiranja.

b) Kada je primjenjiva lakoma (*greedy*) strategija, primjenjivo je i dinamičko programiranje.

c) Kada je primjenjivo dinamičko programiranje, primjenjiva je i lakoma (*greedy*) strategija.

d) Nužan uvjet za primjenu dinamičkog programiranja je preklopljenost podproblema (*overlapping subproblems*), a dovoljan optimalna podstruktura (*optimal substructure*) problema.

e) Nužan uvjet za primjenu dinamičkog programiranja je optimalna podstruktura (*optimal substructure*) problema, a dovoljan preklopljenost podproblema (*overlapping subproblems*).

Napomena: u ovom zadatku se može steći najviše 5 bodova, ali i dobiti do 5 negativnih bodova. Vi navodite tvrdnje koje smatrate istinitima, a prilikom bodovanja će se pretpostaviti da tvrdnje koje niste naveli smatrate neistinitima. Time će Vaši odogovori postati vektor s 5 elemenata ISTINA ili NEISTINA, a bodovanje će se provesti kao binarna usporedba s točnim vektorom. Svaka podudarnost elemenata u vektoru Vaših odgovora i odgovarajućih elemenata u točnom vektoru donijet će 1 bod, a nepodudarnost –1 bod. Jedini način da se ovaj zadatak boduje s nula (0) bodova jest da uopće ništa ne napišete.

Odgovor 1

bodovi: **5**

a) točno

b) točno (iako pitanje je što znači **primijenjivo**, dosta greedy strategija ne profitira od dinamičkog programiranja)

c) netočno (npr. 0-1 knapsack)

d) netočno (oba su nužni uvjeti)

e) netočno (oba su nužni uvjeti)

Zadatak 4 (10)

OZNAKE: **stablo** **B stablo** **dodavanje**

U prazno B-stablo 2. reda upišite redom sljedeće elemente:

26, 4, 22, 16, 30, 17, 31, 20, 6, 1, 21

Odgovor 1

BODOVI: **10**

B-stablo drugog reda postoji ako i samo ako je savršeno stablo. S obzirom na to da se radi o on-line dodavanju elemenata, ovo će biti moguće samo za unos 26, a nakon 2. unosa više ne možemo napraviti B-stablo koje zadovoljava sva pravila B-stabla. Nadalje, S obzirom na to da s 12 elemenata ne možemo stvoriti savršeno stablo, čak i da sve elemente upišemo odjednom ne postoji rješenje zadatka. Prema tome, odgovor za sve bodove je: **zadatak je krivo zadan i rješenje ne postoji**.

Komentar: Riješio sam ovaj zadatak kao AVL stablo i dobio **6** bodova.

Zadatak 5 (9)

oznake: **graf** **Hamilton** **Bondy-Chvatal**

Bondy-Chvatalovim algoritmom (tj. koristeći Bondy-Chvatalov teorem) pronađite Hamiltonov ciklus u grafu zadanom sljedećom matricom susjedstva (udaljenosti):

	1	2	3	4	5	6
1		7				2
2	7				1	
3				4		3
4			4		3	1
5		1		3		4
6	2		3	1	4	

Odgovor 1

bodovi: 9

TODO

Zadatak 6 (12)

OZNAKE: **simpleks** **nejednadžba** **skup** **linearni program**

Za skup S zadan sljedećim nejednadžbama:

$$\begin{aligned} z &\geq 3 \\ 2x + \quad y + 2z &\leq 18 \\ -2x + \quad y + 2z &\leq 6 \\ -y + \quad z &\leq 4 \end{aligned}$$

a) (6) Odredite je li skup S neprezan.

b) (6) Kako biste odredili da li je skup S u prvom ortantu (tj. jesu li sve koordinate svih točaka skupa S nenegativne)? Ne trebate provoditi postupak, ali specificirajte sve potrebno za početak postupka te detaljno opišite nastavak postupka.

Napomena: Pod a) i b) se priznaju odgovori nastali na temelju provođenja efikasnih algoritamskih postupaka.

Odgovor 1 BODOVI: 6

Riješio sam zadatak riješavanjem sustava jednadžbi pod rangom. Dakle pronalazio sam rangove varijabli i postavljao parove jednadžbi. Od tih parova sam zbog linearnosti problema dobio granične točke i uzimao sam stroži dobiveni uvjet. Za to sam dobio 6 bodova, a prof. Brčić mi je rekao što je zapravo trebalo napraviti, pa neka netko tko to zna nadopuni (jer ja ne znam xD):

a) dvofazni simpleks, treba pokazati da je optimum sintetičke ciljne funkcije 0

b) 2 linearna programa (iako je prof. Brčić rekao da se može i jednim al da je dosta teže)