Sveučilište u Zagreb
Fakultet elektrotehnike i računarstva

23. lipnja 2022.

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \ldots, H_i, \ldots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

- A Dokazi su međusobno nezavisni, $P(E_j, E_k) = P(E_j)P(E_k)$, i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{n} E_i = X$
- B Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$
- C Za dani dokaz E_j, hipoteze su međusobno uvjetno nezavisne, $P(H_i, H_k|E_j) = P(H_i|E_j)P(H_k|E_j)$
- D Hipoteze su međusobno isključive, $H_i \cap_{i \neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{m} H_i = X$

2 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0, 1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

- A $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5
- B $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%
- C $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak
- D $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak

3 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) i slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.32 B 0.59 C 0.41 D 0.64

4 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \ldots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \ldots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.25/15, 0.75/16, 1/17, 0.75/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

- A $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$
- B $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$
- C $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$
- D $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$

2. Strojno učenje (5 pitanja)

5 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla k možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku $E(k|D_t)$?**

A $k > k_1$ B $k < k_2$ C $k > k_2$ D $k_1 < k \leq k_2$

6 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

- A Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje
- B U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje
- C Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere
- D Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima

7 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Dalmacija}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost $P(y = 1|\mathbf{x})$?**

A 0.318 B 0.706 C 0.685 D 0.237

8 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, P, T B E, P C E, I, P D E, I

9 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

- A Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela
- B Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$
- C Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela
- D Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

3. Umjetne neuronske mreže (4 pitanja)

10 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\{(1,1,1), (1,4,-1), (3,1,1), (3,3,-1)\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 1$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (9, -15, 1)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

A 2 puta, $(33, 28, -10)$ B 1 put, $(11, -13, 3)$ C 4 puta, $(5.5, -1.5, 10)$ D 4 puta, $(13, -25, 12)$

11 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

A 5096 B 2856 C 8726 D 4640

12 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

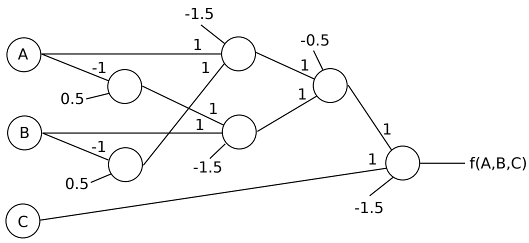
A Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija

B Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom

C Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti

D Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja

13 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



A $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$ C $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$

B $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$ D $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (P) Genetički algoritam pretražuje prostor $[-10, 10]$. Preciznost pretrage mora biti barem 0.0002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

A 9 B 17 C 16 D 14

15 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su smislu križanje i mutacija međusobno komplementarni?**

A Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama

B Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu

C Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima

D Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja

16 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x,y) = 15 - (x-1)^2 - (y-4)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[-1, 6]$ te od y je $[0, 7]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

A 6 B 15 C 2 D 13

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2} = \frac{1}{\sqrt{3}}$, $\tau_{1,4} = 2$, $\tau_{1,6} = 0.5$, $\tau_{2,4} = \frac{1}{2}$, $\tau_{2,5} = \frac{1}{3}$, $\tau_{2,7} = 2$, $\tau_{3,5} = 1$, $\tau_{3,6} = 3$, $\tau_{3,7} = 10$, $\tau_{4,6} = 0.5$, $\tau_{5,7} = \sqrt{3}$, $\tau_{6,7} = 10\sqrt{2}$, $\eta_{1,2} = 3$, $\eta_{1,4} = \frac{1}{\sqrt{2}}$, $\eta_{1,6} = 2$, $\eta_{2,4} = 2\sqrt{2}$, $\eta_{2,5} = 3\sqrt{3}$, $\eta_{2,7} = 0.5$, $\eta_{3,5} = \sqrt{2}$, $\eta_{3,6} = 3\sqrt{3}$, $\eta_{3,7} = \frac{1}{2\sqrt{5}}$, $\eta_{4,6} = 0.5, \eta_{5,7} = 1$, $\eta_{6,7} = 0.1$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j} = \tau_{j,i}$ i $\eta_{i,j} = \eta_{j,i}$. Dodatno, $\alpha = 2$ i $\beta = 2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A 2, 4, 6 B 3, 6, 4 C Mrav neće uspjeti konstruirati ciklus D 1, 2, 4

5. Podržano učenje (3 pitanja)

18 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta stohastička, a okolina neka je deterministička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

A $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s,a) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$
C $v_{k+1}(s) \leftarrow r(s,a,s') + \gamma \cdot v_k(s')$
B $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$
D $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \cdot r(s,a,s')$

19 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha = 0.5$ i $\gamma = 1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost $q(3, desno)$?**

A Veća od 0.9995 B Nije moguće odgovoriti C Manja od 0.75 D Između 0.75 i 0.9995

20 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi -2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa -0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

A $V(3) = q(3, desno)$ i $V(3) < 0.18$
C $V(3) = q(3, gore)$ i $V(3) > 0.15$
B $V(3) = q(3, desno)$ i $V(3) > 0.15$
D $V(3) = q(3, gore)$ i $V(3) < 0.18$

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0, 1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

- A $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak
- B $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5
- C $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak
- D $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%

2 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \dots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \dots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.25/15, 0.75/16, 1/17, 0.75/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

- A $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$
- B $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$
- C $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$
- D $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$

3 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \dots, H_i, \dots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

- A Dokazi su međusobno nezavisni, $P(E_j, E_k) = P(E_j)P(E_k)$, i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{n} E_i = X$
- B Hipoteze su međusobno isključive, $H_i \cap_{i \neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{m} H_i = X$
- C Za dani dokaz E_j, hipoteze su međusobno uvjetno nezavisne, $P(H_i, H_k|E_j) = P(H_i|E_j)P(H_k|E_j)$
- D Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$

4 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) i slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.59 B 0.41 C 0.64 D 0.32

2. Strojno učenje (5 pitanja)

5 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

- A Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima
- B Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje
- C U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje
- D Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere

6 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, I B E, P C E, I, P D E, P, T

7 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Istra}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost $P(y = 1|\mathbf{x})$?**

A 0.685 B 0.237 C 0.318 D 0.706

8 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla k možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku $E(k|D_t)$?**

A $k > k_1$ B $k_1 < k \leq k_2$ C $k < k_2$ D $k > k_2$

9 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

A Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela

B Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela

C Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

D Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$

3. Umjetne neuronske mreže (4 pitanja)

10 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

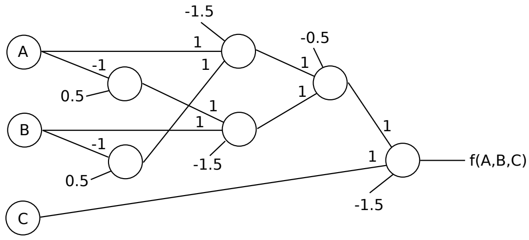
A Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja

B Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija

C Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti

D Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom

11 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



A $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$ C $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$

B $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$ D $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$

12 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\{(1, 1, 1), (1, 4, -1), (3, 1, 1), (3, 3, -1)\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 1$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (10, -10, 10)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

A 2 puta, $(-11, 8, -69)$ B 3 puta, $(-5.2, 0.7, 1.2)$ C 2 puta, $(6, -14, 10)$ D 4 puta, $(13, -25, 12)$

13 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

A 8726 B 2856 C 5096 D 4640

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (P) Genetički algoritam pretražuje prostor $[0, 20]$. Preciznost pretrage mora biti barem 0.0002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

A 9 B 17 C 32 D 10

15 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x, y) = 15 - (x - 1)^2 - (y - 4)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[-1, 6]$ te od y je $[0, 7]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

A 15 B 13 C 6 D 2

16 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su smislu križanje i mutacija međusobno komplementarni?**

- A Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu
- B Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama
- C Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja
- D Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2} = 2$, $\tau_{1,4} = 0.5$, $\tau_{1,7} = \frac{1}{\sqrt{3}}$, $\tau_{2,4} = 0.5$, $\tau_{2,7} = \frac{1}{2}$, $\tau_{3,5} = 1$, $\tau_{3,6} = \sqrt{3}$, $\tau_{3,7} = \frac{1}{3}$, $\tau_{4,5} = 3$, $\tau_{4,6} = 10\sqrt{2}$, $\tau_{5,6} = 10$, $\tau_{6,7} = 2$. $\eta_{1,2} = \frac{1}{\sqrt{2}}$, $\eta_{1,4} = 2$, $\eta_{1,7} = 3$, $\eta_{2,4} = 0.5$, $\eta_{2,7} = 2\sqrt{2}$, $\eta_{3,5} = \sqrt{2}$, $\eta_{3,6} = 1$, $\eta_{3,7} = 3\sqrt{3}$, $\eta_{4,5} = 3\sqrt{3}$, $\eta_{4,6} = 0.1$, $\eta_{5,6} = \frac{1}{2\sqrt{5}}$, $\eta_{6,7} = 0.5$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j} = \tau_{j,i}$ i $\eta_{i,j} = \eta_{j,i}$. Dodatno, $\alpha = 2$ i $\beta = 2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A 7, 6, 3 B Mrav neće uspjeti konstruirati ciklus C 1, 7, 2 D 5, 4, 2

5. Podržano učenje (3 pitanja)

18 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha = 0.5$ i $\gamma = 1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost** $q(3, desno)$**?**

A Manja od 0.75 B Nije moguće odgovoriti C Veća od 0.9995 D Između 0.75 i 0.9995

19 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta stohastička, a okolina neka je deterministička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

- A $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s,a) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$
- C $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \cdot r(s,a,s')$
- B $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$
- D $v_{k+1}(s) \leftarrow r(s,a,s') + \gamma \cdot v_k(s')$

20 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi −2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa −0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

- A $V(3) = q(3, gore)$ i $V(3) < 0.18$
- C $V(3) = q(3, desno)$ i $V(3) > 0.15$
- B $V(3) = q(3, desno)$ i $V(3) < 0.18$
- D $V(3) = q(3, gore)$ i $V(3) > 0.15$

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) i slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.32 B 0.59 C 0.64 D 0.41

2 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \ldots, H_i, \ldots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

A Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$

B Za dani dokaz E_j, hipoteze su međusobno uvjetno nezavisne, $P(H_i, H_k|E_j) = P(H_i|E_j)P(H_k|E_j)$

C Hipoteze su međusobno nezavisne, $P(H_i|H_k) = P(H_i)$, ali dokazi zavise o hipotezi, $P(E_j|H_i) \neq P(E_j)$

D Hipoteze su međusobno isključive, $H_i \cap_{i \neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{m} H_i = X$

3 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \ldots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \ldots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.5/15, 0.75/16, 1/17, 0.5/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

A $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$

B $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$

C $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$

D $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$

4 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0, 1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

A $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%

B $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5

C $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak

D $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak

2. Strojno učenje (5 pitanja)

5 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Dalmacija}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost $P(y = 1|\mathbf{x})$?**

A 0.318 B 0.706 C 0.685 D 0.237

6 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla k možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku $E(k|D_t)$?**

A $k < k_2$ B $k_1 < k \leq k_2$ C $k > k_1$ D $k > k_2$

7 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

A Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$

B Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela

C Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela

D Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

8 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, P B E, I C E, I, P D E, P, T

9 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

A Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima

B U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje

C Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje

D Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere

3. Umjetne neuronske mreže (4 pitanja)

10 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

A 2856 B 8726 C 4640 D 5096

11 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

A Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija

B Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja

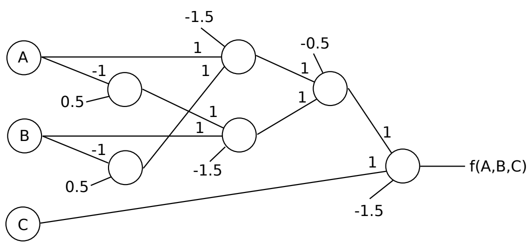
C Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom

D Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti

12 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\big\{(1, 1, 1), (1, 4, -1), (3, 1, 1), (3, 3, -1)\big\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 1$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (10, -10, 10)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

A 3 puta, $(13, 12, -52)$ B 4 puta, $(13, -25, 12)$ C 2 puta, $(6, -14, 10)$ D Postupak ne konvergira

13 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



A $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$ C $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$

B $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$ D $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (P) Genetički algoritam pretražuje prostor $[0, 20]$. Preciznost pretrage mora biti barem 0.002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

A 16 B 14 C 10 D 32

15 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su smislu križanje i mutacija međusobno komplementarni?**

A Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja

B Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima

C Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu

D Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama

16 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x,y) = 13 - (x-2)^2 - (y-3)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[0,7]$ te od y je $[-1,6]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

A 13 B 0 C 4 D 11

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2} = \frac{1}{\sqrt{3}}$, $\tau_{1,4} = 2$, $\tau_{1,6} = 0.5$, $\tau_{2,4} = \frac{1}{2}$, $\tau_{2,5} = \frac{1}{3}$, $\tau_{2,7} = 2$, $\tau_{3,5} = 1$, $\tau_{3,6} = 3$, $\tau_{3,7} = 10$, $\tau_{4,6} = 0.5$, $\tau_{5,7} = \sqrt{3}$, $\tau_{6,7} = 10\sqrt{2}$, $\eta_{1,2} = 3$, $\eta_{1,4} = \frac{1}{\sqrt{2}}$, $\eta_{1,6} = 2$, $\eta_{2,4} = 2\sqrt{2}$, $\eta_{2,5} = 3\sqrt{3}$, $\eta_{2,7} = 0.5$, $\eta_{3,5} = \sqrt{2}$, $\eta_{3,6} = 3\sqrt{3}$, $\eta_{3,7} = \frac{1}{2\sqrt{5}}$, $\eta_{4,6} = 0.5, \eta_{5,7} = 1$, $\eta_{6,7} = 0.1$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j} = \tau_{j,i}$ i $\eta_{i,j} = \eta_{j,i}$. Dodatno, $\alpha = 2$ i $\beta = 2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A 6, 4, 1 B Mrav neće uspjeti konstruirati ciklus C 6, 4, 2 D 2, 7, 5

5. Podržano učenje (3 pitanja)

18 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta deterministička, a okolina neka je stohastička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

A $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s,a) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$
C $v_{k+1}(s) \leftarrow r(s,a,s') + \gamma \cdot v_k(s')$
B $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \cdot r(s,a,s')$
D $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$

19 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi -2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa -0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

A $V(3) = q(3, \textit{gore})$ i $V(3) > 0.15$
C $V(3) = q(3, \textit{desno})$ i $V(3) > 0.15$
B $V(3) = q(3, \textit{desno})$ i $V(3) < 0.18$
D $V(3) = q(3, \textit{gore})$ i $V(3) < 0.18$

20 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha = 0.5$ i $\gamma = 1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost $q(3, \textit{desno})$?**

A Nije moguće odgovoriti B Između 0.75 i 0.9995 C Veća od 0.9995 D Manja od 0.75

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0,1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

- A $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak
- B $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%
- C $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak
- D $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5

2 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) i slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.32 B 0.64 C 0.41 D 0.59

3 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \ldots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \ldots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.25/15, 0.75/16, 1/17, 0.75/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

- A $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$
- B $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$
- C $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$
- D $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$

4 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \ldots, H_i, \ldots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

- A Hipoteze su međusobno nezavisne, $P(H_i|H_k) = P(H_i)$, ali dokazi zavise o hipotezi, $P(E_j|H_i) \neq P(E_j)$
- B Hipoteze su međusobno isključive, $H_i \cap_{i \neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{m} H_i = X$
- C Dokazi su međusobno isključivi, $E_j \cap_{j \neq k} E_k = \emptyset$, a hipoteze nezavisne, $P(H_i, H_l) = P(H_i)P(H_l)$
- D Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$

2. Strojno učenje (5 pitanja)

5 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, I   B E, P   C E, I, P   D E, P, T

6 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Istra}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost $P(y = 1|\mathbf{x})$?**

A 0.706   B 0.318   C 0.685   D 0.237

7 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla k možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku $E(k|D_t)$?**

A $k > k_1$ B $k_1 < k \leq k_2$ C $k < k_2$ D $k > k_2$

8 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

A Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

B Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela

C Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela

D Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$

9 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

A Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima

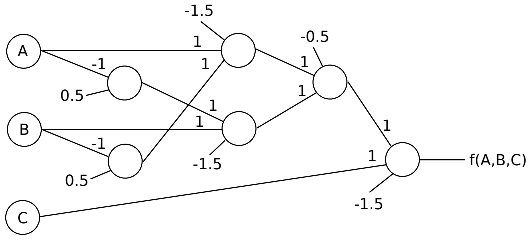
B Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere

C U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje

D Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje

3. Umjetne neuronske mreže (4 pitanja)

10 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



A $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$ C $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$

B $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$ D $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$

11 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

A 4640 B 8726 C 5096 D 2856

12 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\{(1, 1, -1), (1, 4, 1), (3, 1, -1), (3, 3, 1)\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 0.5$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (-10, 20, -5)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

A 2 puta, $(-12, 18, -7)$ B 4 puta, $(5.5, -1.5, 10)$ C 3 puta, $(-5.2, 0.7, 1.2)$ D Postupak ne konvergira

13 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

A Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija

B Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti

C Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom

D Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (P) Genetički algoritam pretražuje prostor $[0, 20]$. Preciznost pretrage mora biti barem 0.0002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

A 17 B 12 C 14 D 10

15 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su smislu križanje i mutacija međusobno komplementarni?**

A Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu

B Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima

C Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama

D Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja

16 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x,y) = 13-(x-2)^2-(y-3)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[0,7]$ te od y je $[-1,6]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

A 0 B 11 C 13 D 4

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2}=2$, $\tau_{1,4}=0.5$, $\tau_{1,7}=\frac{1}{\sqrt{3}}$, $\tau_{2,4}=0.5$, $\tau_{2,7}=\frac{1}{2}$, $\tau_{3,5}=1$, $\tau_{3,6}=\sqrt{3}$, $\tau_{3,7}=\frac{1}{3}$, $\tau_{4,5}=3$, $\tau_{4,6}=10\sqrt{2}$, $\tau_{5,6}=10$, $\tau_{6,7}=2$. $\eta_{1,2}=\frac{1}{\sqrt{2}}$, $\eta_{1,4}=2$, $\eta_{1,7}=3$, $\eta_{2,4}=0.5$, $\eta_{2,7}=2\sqrt{2}$, $\eta_{3,5}=\sqrt{2}$, $\eta_{3,6}=1$, $\eta_{3,7}=3\sqrt{3}$, $\eta_{4,5}=3\sqrt{3}$, $\eta_{4,6}=0.1$, $\eta_{5,6}=\frac{1}{2\sqrt{5}}$, $\eta_{6,7}=0.5$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j}=\tau_{j,i}$ i $\eta_{i,j}=\eta_{j,i}$. Dodatno, $\alpha=2$ i $\beta=2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A Mrav neće uspjeti konstruirati ciklus B 1, 7, 2 C 3, 6, 5 D 6, 4, 1

5. Podržano učenje (3 pitanja)

18 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi -2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa -0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

A $V(3)=q(3,\mathit{desno})$ i $V(3)>0.15$ C $V(3)=q(3,\mathit{gore})$ i $V(3)>0.15$
B $V(3)=q(3,\mathit{gore})$ i $V(3)<0.18$ D $V(3)=q(3,\mathit{desno})$ i $V(3)<0.18$

19 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha=0.5$ i $\gamma=1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost $q(3,\mathit{desno})$?**

A Veća od 0.9995 B Između 0.75 i 0.9995 C Manja od 0.75 D Nije moguće odgovoriti

20 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta deterministička, a okolina neka je stohastička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

A $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s,a)\cdot r(s,a,s')$ C $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s,a)\cdot[r(s,a,s')+\gamma\cdot v_k(s')]$
B $v_{k+1}(s) \leftarrow \sum_a \pi(a|s)\cdot[r(s,a,s')+\gamma\cdot v_k(s')]$ D $v_{k+1}(s) \leftarrow r(s,a,s')+\gamma\cdot v_k(s')$

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \ldots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \ldots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.5/15, 0.75/16, 1/17, 0.5/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

A $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$

B $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$

C $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$

D $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$

2 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0, 1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

A $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5

B $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%

C $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak

D $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak

3 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \ldots, H_i, \ldots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

A Hipoteze su međusobno isključive, $H_i \cap_{i \neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{m} H_i = X$

B Dokazi su međusobno isključivi, $E_j \cap_{j \neq k} E_k = \emptyset$, a hipoteze nezavisne, $P(H_i, H_l) = P(H_i)P(H_l)$

C Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$

D Za dani dokaz E_j, hipoteze su međusobno uvjetno nezavisne, $P(H_i, H_k|E_j) = P(H_i|E_j)P(H_k|E_j)$

4 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) i slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.59 B 0.41 C 0.64 D 0.32

2. Strojno učenje (5 pitanja)

5 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Dalmacija}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost $P(y = 1|\mathbf{x})$?**

A 0.237 B 0.318 C 0.706 D 0.685

6 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

A Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

B Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela

C Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela

D Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$

7 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

A Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje

B Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere

C U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje

D Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima

8 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, I B E, P, T C E, I, P D E, P

9 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla k možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku $E(k|D_t)$?**

A $k < k_2$ B $k_1 < k \leq k_2$ C $k > k_2$ D $k > k_1$

3. Umjetne neuronske mreže (4 pitanja)

10 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

- A Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom
- B Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja
- C Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti
- D Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija

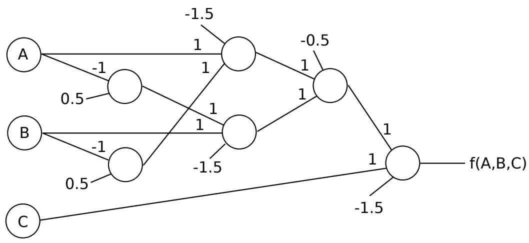
11 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

A 4640 B 5096 C 2856 D 8726

12 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\{(1, 1, 1), (1, 4, -1), (3, 1, 1), (3, 3, -1)\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 1$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (10, -10, 10)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

A 4 puta, $(5.5, -1.5, 10)$ B 2 puta, $(6, -14, 10)$ C 2 puta, $(-11, 8, -69)$ D 3 puta, $(13, 12, -52)$

13 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



A $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$ C $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$

B $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$ D $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (P) Genetički algoritam pretražuje prostor $[-10, 10]$. Preciznost pretrage mora biti barem 0.002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

A 32 B 16 C 14 D 10

15 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su smislu križanje i mutacija međusobno komplementarni?**

- A Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima
- B Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja
- C Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama
- D Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu

16 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x, y) = 18 - (x - 4)^2 - (y - 2)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[2, 9]$ te od y je $[-2, 5]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

A 5 B 9 C 18 D 16

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2} = 2$, $\tau_{1,4} = 0.5$, $\tau_{1,7} = \frac{1}{\sqrt{3}}$, $\tau_{2,4} = 0.5$, $\tau_{2,7} = \frac{1}{2}$, $\tau_{3,5} = 1$, $\tau_{3,6} = \sqrt{3}$, $\tau_{3,7} = \frac{1}{3}$, $\tau_{4,5} = 3$, $\tau_{4,6} = 10\sqrt{2}$, $\tau_{5,6} = 10$, $\tau_{6,7} = 2$. $\eta_{1,2} = \frac{1}{\sqrt{2}}$, $\eta_{1,4} = 2$, $\eta_{1,7} = 3$, $\eta_{2,4} = 0.5$, $\eta_{2,7} = 2\sqrt{2}$, $\eta_{3,5} = \sqrt{2}$, $\eta_{3,6} = 1$, $\eta_{3,7} = 3\sqrt{3}$, $\eta_{4,5} = 3\sqrt{3}$, $\eta_{4,6} = 0.1$, $\eta_{5,6} = \frac{1}{2\sqrt{5}}$, $\eta_{6,7} = 0.5$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j} = \tau_{j,i}$ i $\eta_{i,j} = \eta_{j,i}$. Dodatno, $\alpha = 2$ i $\beta = 2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A 3, 6, 5 B Mrav neće uspjeti konstruirati ciklus C 6, 4, 1 D 1, 7, 2

5. Podržano učenje (3 pitanja)

18 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha = 0.5$ i $\gamma = 1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost $q(3, desno)$?**

A Veća od 0.9995 B Nije moguće odgovoriti C Manja od 0.75 D Između 0.75 i 0.9995

19 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi -2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa -0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

A $V(3) = q(3, desno)$ i $V(3) > 0.15$ C $V(3) = q(3, desno)$ i $V(3) < 0.18$

B $V(3) = q(3, gore)$ i $V(3) > 0.15$ D $V(3) = q(3, gore)$ i $V(3) < 0.18$

20 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta deterministička, a okolina neka je stohastička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

A $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \cdot r(s,a,s')$ C $v_{k+1}(s) \leftarrow r(s,a,s') + \gamma \cdot v_k(s')$

B $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$ D $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s,a) \cdot [r(s,a,s') + \gamma \cdot v_k(s')]$

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) ili slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.32 B 0.41 C 0.64 D 0.59

2 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \ldots, H_i, \ldots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

A Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$

B Za dani dokaz E_j, hipoteze su međusobno uvjetno nezavisne, $P(H_i, H_k|E_j) = P(H_i|E_j)P(H_k|E_j)$

C Hipoteze su međusobno isključive, $H_i \cap_{i\neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{m} H_i = X$

D Dokazi su međusobno nezavisni, $P(E_j, E_k) = P(E_j)P(E_k)$, i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{n} E_i = X$

3 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \ldots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \ldots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.5/15, 0.75/16, 1/17, 0.5/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

A $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$

C $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$

B $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$

D $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$

4 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0, 1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

A $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak

B $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%

C $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak

D $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5

2. Strojno učenje (5 pitanja)

5 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

A Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

B Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela

C Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela

D Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$

6 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla k možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku $E(k|D_t)$?**

A $k < k_2$ B $k > k_1$ C $k > k_2$ D $k_1 < k \leq k_2$

7 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, P, T B E, I, P C E, P D E, I

8 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

A Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima

B Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere

C U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje

D Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje

9 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Istra}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost $P(y = 1|\mathbf{x})$?**

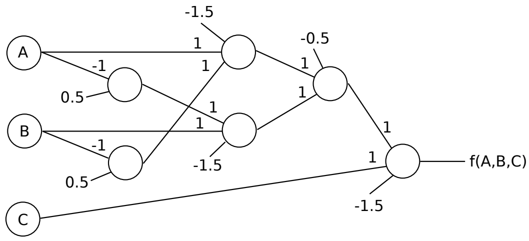
A 0.706 B 0.685 C 0.237 D 0.318

3. Umjetne neuronske mreže (4 pitanja)

10 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\{(1,1,1),(1,4,-1),(3,1,1),(3,3,-1)\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 1$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (10, -10, 10)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

A 3 puta, $(-5.2, 0.7, 1.2)$ B 3 puta, $(13, 12, -52)$ C 2 puta, $(-11, 8, -69)$ D 2 puta, $(6, -14, 10)$

11 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



A $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$ C $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$

B $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$ D $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$

12 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

A 2856   B 5096   C 4640   D 8726

13 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

A Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti

B Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija

C Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom

D Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x, y) = 18 - (x - 4)^2 - (y - 2)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[2, 9]$ te od y je $[-2, 5]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

A 5   B 9   C 16   D 18

15 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su**

smislu križanje i mutacija međusobno komplementarni?

A Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima

B Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama

C Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja

D Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu

16 (P) Genetički algoritam pretražuje prostor $[0, 20]$. Preciznost pretrage mora biti barem 0.002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

A 32 B 9 C 14 D 16

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2} = 2$, $\tau_{1,4} = 0.5$, $\tau_{1,7} = \frac{1}{\sqrt{3}}$, $\tau_{2,4} = 0.5$, $\tau_{2,7} = \frac{1}{2}$, $\tau_{3,5} = 1$, $\tau_{3,6} = \sqrt{3}$, $\tau_{3,7} = \frac{1}{3}$, $\tau_{4,5} = 3$, $\tau_{4,6} = 10\sqrt{2}$, $\tau_{5,6} = 10$, $\tau_{6,7} = 2$. $\eta_{1,2} = \frac{1}{\sqrt{2}}$, $\eta_{1,4} = 2$, $\eta_{1,7} = 3$, $\eta_{2,4} = 0.5$, $\eta_{2,7} = 2\sqrt{2}$, $\eta_{3,5} = \sqrt{2}$, $\eta_{3,6} = 1$, $\eta_{3,7} = 3\sqrt{3}$, $\eta_{4,5} = 3\sqrt{3}$, $\eta_{4,6} = 0.1$, $\eta_{5,6} = \frac{1}{2\sqrt{5}}$, $\eta_{6,7} = 0.5$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j} = \tau_{j,i}$ i $\eta_{i,j} = \eta_{j,i}$. Dodatno, $\alpha = 2$ i $\beta = 2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A 7, 6, 3 B Mrav neće uspjeti konstruirati ciklus C 4, 2, 1 D 4, 2, 7

5. Podržano učenje (3 pitanja)

18 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta stohastička, a okolina neka je deterministička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

A $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \cdot [r(s, a, s') + \gamma \cdot v_k(s')]$

B $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s, a) \cdot [r(s, a, s') + \gamma \cdot v_k(s')]$

C $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s, a) \cdot r(s, a, s')$

D $v_{k+1}(s) \leftarrow r(s, a, s') + \gamma \cdot v_k(s')$

19 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi −2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa −0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

A $V(3) = q(3, \textit{desno})$ i $V(3) > 0.15$

B $V(3) = q(3, \textit{gore})$ i $V(3) < 0.18$

C $V(3) = q(3, \textit{desno})$ i $V(3) < 0.18$

D $V(3) = q(3, \textit{gore})$ i $V(3) > 0.15$

20 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha = 0.5$ i $\gamma = 1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost $q(3, \textit{desno})$?**

A Između 0.75 i 0.9995 B Nije moguće odgovoriti C Veća od 0.9995 D Manja od 0.75

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \dots, H_i, \dots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

- A Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$
- B Za dani dokaz E_j, hipoteze su međusobno uvjetno nezavisne, $P(H_i, H_k|E_j) = P(H_i|E_j)P(H_k|E_j)$
- C Hipoteze su međusobno nezavisne, $P(H_i|H_k) = P(H_i)$, ali dokazi zavise o hipotezi, $P(E_j|H_i) \neq P(E_j)$
- D Hipoteze su međusobno isključive, $H_i \cap_{i \neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^{m} H_i = X$

2 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) ili slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.59 B 0.32 C 0.64 D 0.41

3 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \dots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \dots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.25/15, 0.75/16, 1/17, 0.75/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

- A $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$
- B $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$
- C $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$
- D $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$

4 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0, 1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

- A $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5
- B $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak
- C $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%
- D $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak

2. Strojno učenje (5 pitanja)

5 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla k možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku $E(k|D_t)$?**

A $k > k_1$ B $k > k_2$ C $k < k_2$ D $k_1 < k \leq k_2$

6 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

A Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

B Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela

C Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$

D Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela

7 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, I, P   B E, P, T   C E, P   D E, I

8 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Istra}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost $P(y = 1|\mathbf{x})$?**

A 0.237   B 0.685   C 0.706   D 0.318

9 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

A U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje

B Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje

C Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere

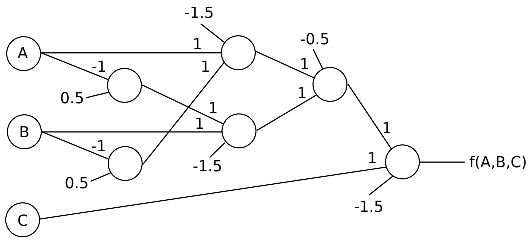
D Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima

3. Umjetne neuronske mreže (4 pitanja)

10 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

- A Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti
- B Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja
- C Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija
- D Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom

11 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



- A $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$
- C $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$
- B $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$
- D $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$

12 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\{(1, 1, 1), (1, 4, -1), (3, 1, 1), (3, 3, -1)\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 1$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (10, -10, 10)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

- A 2 puta, $(6, -14, 10)$
- B Postupak ne konvergira
- C 3 puta, $(-5.2, 0.7, 1.2)$
- D 3 puta, $(13, 12, -52)$

13 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

- A 5096
- B 4640
- C 2856
- D 8726

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x, y) = 15 - (x - 1)^2 - (y - 4)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[-1, 6]$ te od y je $[0, 7]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

- A 13
- B 2
- C 15
- D 6

15 (P) Genetički algoritam pretražuje prostor $[-10, 10]$. Preciznost pretrage mora biti barem 0.002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

- A 32
- B 16
- C 14
- D 8

16 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su smislu križanje i mutacija međusobno komplementarni?**

A Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima

B Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu

C Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja

D Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2} = 2$, $\tau_{1,4} = 0.5$, $\tau_{1,7} = \frac{1}{\sqrt{3}}$, $\tau_{2,4} = 0.5$, $\tau_{2,7} = \frac{1}{2}$, $\tau_{3,5} = 1$, $\tau_{3,6} = \sqrt{3}$, $\tau_{3,7} = \frac{1}{3}$, $\tau_{4,5} = 3$, $\tau_{4,6} = 10\sqrt{2}$, $\tau_{5,6} = 10$, $\tau_{6,7} = 2$. $\eta_{1,2} = \frac{1}{\sqrt{2}}$, $\eta_{1,4} = 2$, $\eta_{1,7} = 3$, $\eta_{2,4} = 0.5$, $\eta_{2,7} = 2\sqrt{2}$, $\eta_{3,5} = \sqrt{2}$, $\eta_{3,6} = 1$, $\eta_{3,7} = 3\sqrt{3}$, $\eta_{4,5} = 3\sqrt{3}$, $\eta_{4,6} = 0.1$, $\eta_{5,6} = \frac{1}{2\sqrt{5}}$, $\eta_{6,7} = 0.5$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j} = \tau_{j,i}$ i $\eta_{i,j} = \eta_{j,i}$. Dodatno, $\alpha = 2$ i $\beta = 2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A 7, 6, 3 B 7, 2, 4 C 3, 6, 5 D Mrav neće uspjeti konstruirati ciklus

5. Podržano učenje (3 pitanja)

18 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta stohastička, a okolina neka je deterministička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

A $v_{k+1}(s) \leftarrow r(s, a, s') + \gamma \cdot v_k(s')$

B $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s, a) \cdot r(s, a, s')$

C $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \cdot [r(s, a, s') + \gamma \cdot v_k(s')]$

D $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s, a) \cdot [r(s, a, s') + \gamma \cdot v_k(s')]$

19 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi -2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa -0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

A $V(3) = q(3, gore)$ i $V(3) > 0.15$

B $V(3) = q(3, desno)$ i $V(3) > 0.15$

C $V(3) = q(3, gore)$ i $V(3) < 0.18$

D $V(3) = q(3, desno)$ i $V(3) < 0.18$

20 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha = 0.5$ i $\gamma = 1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost $q(3, desno)$?**

A Veća od 0.9995 B Nije moguće odgovoriti C Manja od 0.75 D Između 0.75 i 0.9995

Završni ispit iz Uvoda u umjetnu inteligenciju (2021./2022.)

– NEKORIGIRANA VERZIJA –

*Ispit se sastoji od **20 pitanja** i ukupno nosi **20 bodova**. Točan odgovor nosi 1 bod, a netočan -1/3 boda. Trajanje ispita je **150 minuta**. Primjerak ispita trebate predati zajedno sa svojim rješenjima.*

1. Modeliranje neizvjesnosti (4 pitanja)

1 (T) Bayesovo pravilo može se primijeniti na slučaj s više od dvije hipoteze, $H_1, \ldots, H_i, \ldots, H_m$. **Što je pretpostavka Bayesovog pravila s više od dvije hipoteze?**

- A Hipoteze su međusobno nezavisne, $P(H_i|H_k) = P(H_i)$, ali dokazi zavise o hipotezi, $P(E_j|H_i) \neq P(E_j)$
- B Hipoteze su međusobno isključive, $H_i \cap_{i\neq j} H_j = \emptyset$ i pokrivaju cijeli prostor događaja, $\cup_{i=1}^m H_i = X$
- C Za dani dokaz E_j, hipoteze su međusobno uvjetno nezavisne, $P(H_i, H_k|E_j) = P(H_i|E_j)P(H_k|E_j)$
- D Za danu hipotezu H_i, dokazi su međusobno uvjetno nezavisni, $P(E_j, E_k|H_i) = P(E_j|H_i)P(E_k|H_i)$

2 (T) Vrijednost funkcije pripadnosti za element $x \in X$, tj. $\mu(x)$, na prvi se pogled čini istovjetna vjerojatnosti realizacije događaja x, tj. $P(x) \in [0, 1]$. Neka je x skup svih ljudi, a V (neizraziti) skup ljudi s visokim tlakom. Neka $\mu_V(x) = P(x = V) = 0.2$. **Koja je razlika između $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$?**

- A $\mu_V(x) = 0.2$ znači da osoba ima nešto malo povišen tlak, dok $P(x = V) = 0.2$ znači da očekivano jedna od pet osoba ima visok tlak
- B $\mu_V(x) = 0.2$ i $P(x = V) = 0.2$ oboje znače da je osoba u skupu ljudi s visokim tlakom s pripadnošću od 1/5
- C $\mu_V(x) = 0.2$ znači da jedna od pet osoba pripada skupu ljudi s povišenim tlakom, dok $P(x = V) = 0.2$ znači da je šansa visokog tlaka 20%
- D $\mu_V(x) = 0.2$ znači da četiri od pet osoba nemaju visok tlak, dok $P(x = V) = 0.2$ znači da je jedna od pet osoba očekivano ima visok tlak

3 (R) U sustavu neizrazitog zaključivanja modeliramo vezu između dijabetesa i sklonosti depresiji. Skup G je prosječna razina glukoze u krvi u mmol/L, $G = \{1, 2, \ldots, 11\}$, a skup D je stupanj depresije prema Beckovom upitniku depresije, $D = \{0, 1, \ldots, 63\}$. Nad ovim univerzalnim skupovima definirali smo neizrazite skupove "predijabetes" kao $G_p = \{0.3/5, 1/6, 0.1/7\}$ te "umjerena depresija" kao $D_u = \{0.25/15, 0.75/16, 1/17, 0.75/18\}$. U sustavu imamo pravilo "ako manje-više(G_p), onda D_u", modelirano kao neizrazita relacija. Zanima nas kolika je sklonost depresiji osoba čija razina glukoze u krvi odgovara neizrazitom skupu $G' = \{0.7/5, 1/6, 0.3/7, 0.1/8\}$. **Kako glasi neizraziti skup D' koji dobivamo generaliziranim modusom ponensom?**

- A $\{0.3/15, 0.75/16, 0.75/17, \sqrt{0.1}/18\}$
- B $\{0.5/15, 0.75/16, 1/17, 0.5/18\}$
- C $\{\sqrt{0.1}/15, 0.3/16, 0.75/17, 0.3/18\}$
- D $\{0.25/15, 0.75/16, 1/17, 0.75/18\}$

4 (P) Primjenom standardnih (Zadehovih) neizrazitih operacija i jezičnih modifikatora konstruirali smo neizrazite skupove *"manje-više snažan klokan"* i *"vrlo slab klokan"*. Klokan Roger prvom skupu pripada sa 0.8, a drugom sa 0.1. Primjenom istih operatora i modifikatora konstruirali smo neizraziti skup *"ne (vrlo snažan klokan) ili slab klokan"*. **Koja je pripadnost klokana Rogera tom neizrazitom skupu?**

A 0.41   B 0.32   C 0.59   D 0.64

2. Strojno učenje (5 pitanja)

5 (R) Treniramo naivan Bayesov klasifikator za *"Skupo ljetovanje na Jadranu"*. Skup za učenje je sljedeći:

i	Mjesto	Otok	Smještaj	Prijevoz	y	i	Mjesto	Otok	Smještaj	Prijevoz	y
1	Kvarner	da	privatni	auto	1	5	Kvarner	ne	kamp	bus	0
2	Dalmacija	da	hotel	avion	1	6	Dalmacija	ne	privatni	avion	0
3	Istra	da	kamp	auto	0	7	Istra	ne	kamp	bus	1
4	Dalmacija	da	hotel	auto	1						

Procjene radimo Laplaceovim procjeniteljem "dodaj jedan". Klasificiramo primjer $\mathbf{x} = (\text{Istra}, \text{ne}, \text{kamp}, \text{bus})$. **Koliko iznosi aposteriorna vjerojatnost** $P(y = 1|\mathbf{x})$**?**

A 0.685 B 0.237 C 0.318 D 0.706

6 (P) Algoritmom ID3 učimo stablo odluke. Kako bismo spriječili prenaučenost, koristimo unakrsnu provjeru i skup primjera podijelili smo na skup za učenje D_u, skup za provjeru D_p i skup za testiranje D_t. Stablo učimo na D_u i zatim ga podrezujemo na različitim dubinama k, ispitujući pogrešku na D_p. Neka je $E(k|D)$ pogreška stabla dubine k na skupu primjera D. Za neke k_1 i k_2, gdje $k_2 > k_1$, opažamo da $E(k_1|D_u) > E(k_2|D_u)$ i $E(k_1|D_p) > E(k_2|D_p)$. **Za koju dubinu stabla** k **možemo sigurno reći da će ostvariti najmanju očekivanu pogrešku** $E(k|D_t)$**?**

A $k > k_2$ B $k < k_2$ C $k > k_1$ D $k_1 < k \leq k_2$

7 (R) Mali je Ivica svakog svakog ljeta u zadnjih sedam godina naučio jedan novi programski jezik. Svoja vrijedna iskustva sažeo je u listu *"Programski jezik koji mi se sviđa"*, gdje je svaki jezik opisao četirima značajkama, te je naznačio je li mu se dotični jezik svidio ($y = 1$) ili nije ($y = 0$). Ta lista izgleda ovako:

i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y	i	Eval (E)	Izvođenje (I)	Paradigma (P)	Tipovi (T)	y
1	lijena	kompajler	imperativna	statička	0	5	striktna	interpreter	imperativna	statička	1
2	striktna	interpreter	deklarativna	dinamička	0	6	lijena	kompajler	hibridna	dinamička	1
3	lijena	kompajler	imperativna	dinamička	0	7	striktna	kompajler	hibridna	dinamička	1
4	lijena	interpreter	hibridna	statička	0						

Ovog ljeta Mali Ivica opet naučiti novi programski jezik, pa je odlučio izgraditi stablo odluke za klasifikaciju programskih jezika koje još nije naučio. Pomozite Malom Ivici te na gornjem skupu primjera primjenite algoritam ID3. U slučaju da dvije značajke ili više njih imaju jednaku vrijednost informacijske dobiti, prednost dajte značajki koje je u gornjoj tablici navedena prva (ljevija značajka). **Koje se značajke nalaze u drugoj razini stabla odluke dobivenog algoritmom ID3?** (Korijenski čvor je čvor na prvoj razini.)

A E, I, P B E, P C E, P, T D E, I

8 (T) Učenje naivnog Bayesovog klasifikatora svodi se na procjene vjerojatnosti na temelju označenog skupa podataka. Kod procjena uvjetnih vjerojatnosti koje odgovaraju izglednostima klasa, $P(x_j|y)$, preporuča se, umjesto procjene najveće izglednosti (MLE), napraviti zaglađenu procjenu, npr., Laplaceovim zaglađivanjem. **Zbog čega je dobro za izglednosti klasa koristiti zaglađene procjene?**

A Povećava se očekivana točnost modela u situacijama kada postoji stohastička zavisnost između značajke x_j i oznake y, tj. kada $P(x_j|y) \neq P(x_j)$

B Sprječava se da se kombinacija oznake y i značajke x_j, koja se nije pojavila u skupu za učenje, proglasi nemogućom, što bi dovelo do prenaučenosti modela

C Smanjuje se vjerojatnost podljeva kod izračuna brojnika Bayesovog pravila u situacijama kada imamo mnogo značajki s razmjerno malim uvjetnim vjerojatnostima $P(x_j|y)$

D Sprječava se da vjerojatnost $P(x_j|y)$ bude jednaka nuli u situacijama kada se oznaka y nije pojavila u skupu za ispitivanje, što bi moglo narušiti točnost modela

9 (T) Algoritam ID3 rekurzivno dijeli skup primjera za učenje, koristeći u svakom čvoru stabla informacijsku dobit za odabir značajke. **Što bi se dogodilo da, umjesto informacijske dobiti, značajke u čvorovima odabiremo slučajno?** Pritom, naravno, istu značajku ne bismo u stazi stabla iskoristili više od jednom.

A U prosjeku bismo dobili stablo jednake veličine, ali bi ono točnije klasificiralo primjere za učenje

B Dobili bismo pliće stablo, koje bi očekivano imalo manju klasifikacijsku točnost na svim primjerima

C Dobili bismo veće stablo, koje bi savršeno klasificiralo primjere za učenje, ali lošije neviđene primjere

D Dobili bismo prenaučeno stablo, koje bi dobro klasificiralo neviđene primjere, ali loše primjere za učenje

3. Umjetne neuronske mreže (4 pitanja)

10 (T) Prisjetite se TLU-perceptrona i neurona u višeslojnoj unaprijednoj umjetnoj neuronskoj mreži učenoj algoritmom propagacije pogreške unazad (engl. *backpropagation*). Oba ova neurona imaju nelinearnu aktivacijsku funkciju f. **Što bi se dogodilo kada bi aktivacijska funkcija f bila linearna, $f(x) = x$?**

- A Za modeliranje nelinearne funkcije višeslojne mreže s takvim neuronima iziskivale bi više čvorova od mreže s nelinearnim aktivacijskim funkcijama, pa bi ih bilo lakše prenaučiti
- B Izlazi takvih neurona mogli bi se tumačiti kao vjerojatnosti, budući da bi bili ograničeni na interval $[0, 1]$, međutim za takve neurone ne vrijedi Hebbovo pravilo učenja
- C Višeslojne mreže s takvim neuronima ne bi mogle modelirati nelinearne funkcije, jer bi linearna kombinacija linearnih funkcija i dalje bila linearna funkcija
- D Takve neurone ne bismo mogli učiti algoritmom propagacije pogreške unazad, jer je derivacija pogreške konstantna za svaki ulaz u neuron, ali bismo pojedinačni neuron mogli učiti Rosenblattovim algoritmom

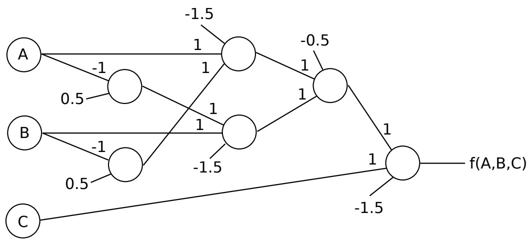
11 (P) Na računalu implementiramo unaprijednu potpuno povezanu slojevitu umjetnu neuronsku mrežu arhitekture $3 \times 40 \times 10 \times 5 \times 2$. Neuroni kao prijenosne funkcije koriste zglobnicu. U memoriji težine mreže čuvaju se kao tip `double` koji zauzima 8 okteta. **Koliki je ukupni utrošak memorije za parametre ove mreže?**

A 5096 B 8726 C 4640 D 2856

12 (R) Skup primjera za učenje $\{(x_2, x_1, y)\}$ je sljedeći: $\{(1,1,1), (1,4,-1), (3,1,1), (3,3,-1)\}$. Učenje se provodi uporabom perceptrona TLU s izlaznim vrijednostima -1 i 1 te stopom učenje $\eta = 1$. Početne vrijednosti težinskih faktora su $(w_2, w_1, w_0) = (10, -10, 10)$. Provedite postupak učenja uporabom Rosenblattovog algoritma. **Koliko se puta tijekom učenja provode korekcije težina te koje su njihove konačne vrijednosti?**

A Postupak ne konvergira B 2 puta, $(6, -14, 10)$ C 3 puta, $(13, 12, -52)$ D 2 puta, $(-11, 8, -69)$

13 (P) Umjetna neuronska mreža na donjoj slici sastoji se od 6 TLU-perceptrona s funkcijom skoka 0-1. Mreža ima tri ulaza (A, B i C) te jedan izlaz. Ako se vrijednosti ulaznih varijabli ograniče na diskretan skup $\{0, 1\}$, tada se svaki neuron može interpretirati kao jedan logički sklop koji obavlja neku elementarnu Booleovu operaciju. U tom slučaju i čitava mreža ostvaruje neku Booleovu funkciju. **Koju Booleovu funkciju implementira mreža sa slike?** (Napomena: u ponuđenim rješenjima ne mora biti ponuđen strukturno jednak zapis.)



A $(A + \bar{B} + C) \cdot (\bar{A} + B + C)$ C $A \cdot B \cdot C + \bar{A} \cdot \bar{B} \cdot C$

B $(A + B + C) \cdot (\bar{A} + \bar{B} + C)$ D $A \cdot \bar{B} \cdot C + \bar{A} \cdot B \cdot C$

4. Prirodom inspirirani optimizacijski algoritmi (4 pitanja)

14 (P) Genetički algoritam pretražuje prostor $[0, 20]$. Preciznost pretrage mora biti barem 0.0002 i koristi se kromosom s binarnim kodiranjem. **Koliko minimalno bitova trebamo upotrijebiti?**

A 14 B 9 C 17 D 12

15 (R) Generacijskim genetskim algoritmom tražimo maksimum funkcije $f(x, y) = 13 - (x-2)^2 - (y-3)^2$. Kao reprezentaciju rješenja koristimo 6-bitovni kromosom, pri čemu se prva tri bita koriste za kodiranje vrijednosti varijable x, a preostala tri bita za kodiranje vrijednosti varijable y. Domena nad kojom se pretražuju vrijednosti od x je $[0, 7]$ te od y je $[-1, 6]$. Populacija se sastoji od četiri jedinke: J1=000110, J2=011001, J3=100111, J4=001110. Pretpostavite da se roditelji biraju proporcionalnom selekcijom te da su u jednom koraku kao roditelji izvučena dva rješenja koja imaju najmanju vjerojatnost odabira. Provedite nad njima postupak križanja s jednom točkom prijeloma (točka prijeloma je nakon prva dva bita); pretpostavite da u ovom koraku operator mutacije svaki puta djeluje na posljednja dva bita kromosoma. **Odredite iznos funkcije f u rješenju koje odgovara djetetu koje će biti poslano u sljedeću generaciju.** (Ako ih operatori križanja i mutacije generiraju više, u novu generaciju treba proslijediti najbolje.)

A 11 B 4 C 0 D 13

16 (T) Selekcija, križanje i mutacija tri su osnovna operatora genetičkih algoritama. Križanje i mutacija međusobno su komplementarni, i da bi optimizacija bila učinkovita tipično su potrebna oba ova operatora. **U kojem su smislu križanje i mutacija međusobno komplementarni?**

A Križanje povećava dobrotu rješenja, dok mutacija osigurava da se ne izgube najbolja rješenja pronađena u prethodnim generacijama

B Križanje potencijalno poboljšava već nađena dobra rješenja, a mutacija potencijalno sprječava zaglavljivanje u lokalnim optimumima

C Mutacija povećava, a križanje smanjuje vjerojatnost selekcije najboljih rješenja

D Mutacija povećava dobrotu rješenja u lokalnom optimumu, a križanje u globalnom optimumu

17 (R) Uporabom mravlje kolonije traži se ciklus kroz graf. Poznati su sljedeći podatci: $\tau_{1,2} = 2$, $\tau_{1,4} = 0.5$, $\tau_{1,7} = \frac{1}{\sqrt{3}}$, $\tau_{2,4} = 0.5$, $\tau_{2,7} = \frac{1}{2}$, $\tau_{3,5} = 1$, $\tau_{3,6} = \sqrt{3}$, $\tau_{3,7} = \frac{1}{3}$, $\tau_{4,5} = 3$, $\tau_{4,6} = 10\sqrt{2}$, $\tau_{5,6} = 10$, $\tau_{6,7} = 2$. $\eta_{1,2} = \frac{1}{\sqrt{2}}$, $\eta_{1,4} = 2$, $\eta_{1,7} = 3$, $\eta_{2,4} = 0.5$, $\eta_{2,7} = 2\sqrt{2}$, $\eta_{3,5} = \sqrt{2}$, $\eta_{3,6} = 1$, $\eta_{3,7} = 3\sqrt{3}$, $\eta_{4,5} = 3\sqrt{3}$, $\eta_{4,6} = 0.1$, $\eta_{5,6} = \frac{1}{2\sqrt{5}}$, $\eta_{6,7} = 0.5$. Također, sve dane vrijednosti su simetrične, tj. $\tau_{i,j} = \tau_{j,i}$ i $\eta_{i,j} = \eta_{j,i}$. Dodatno, $\alpha = 2$ i $\beta = 2$. Prvi mrav kreće iz čvora 1. Kada mrav treba donijeti vjerojatnosnu odluku, pretpostavite da će ishod slučajnog odabira odgovarati najvjerojatnijem. Ako iz nekog čvora mrav ne može dalje, konstrukcija ciklusa se prekida. Uz te pretpostavke odredite ciklus koji će taj mrav konstruirati. **Slijed od koja tri čvora je dio tog ciklusa?**

A 7, 2, 4 B 5, 4, 2 C 3, 5, 4 D Mrav neće uspjeti konstruirati ciklus

5. Podržano učenje (3 pitanja)

18 (R) Rešetkasti svijet sastoji se od 5 ćelija. Ćelije 1 do 4 postavljene su horizontalno s lijeva na desno; ćelija 5 nalazi se ispod ćelije 4. Ćelije 4 i 5 su terminalne: ćelija 4 je izlaz koji agentu donosi 1 bod; ćelija 5 je provalija koja agentu donosi -2 boda. Svijet je ograđen zidovima. Agent može pokrenuti izvođenje akcija pomicanja, i to: *gore*, *dolje*, *lijevo* ili *desno*. Okolina je stohastička, u smislu da kada agent zatraži jednu akciju, ta će biti ostvarena u 60% slučajeva; u preostalih 20%+20% slučajeva provest će se jedna od dvije "nesuprotne" akcije (primjerice, ako zatraži *gore*, to će biti provedeno u 60% slučajeva; u 20% slučajeva provest će se *lijevo* te u 20% slučajeva *desno*). Ako bi pomicanje rezultiralo udaranjem u zid, agent ostaje na ćeliji na kojoj je bio. Također, okolina će agentu za svaki potez (osim onog kojim agent prelazi na terminalnu ćeliju) dodijeliti i nagradu življenja iznosa -0.2. Agent uvijek kreće iz ćelije 1. Potrebno je algoritmom iteracije vrijednosti odrediti optimalne iznose funkcije vrijednosti. Pretpostavite da su početne procjene iznosa funkcije vrijednosti jednake 0, te da je faktor γ jednak 1. Provedite dvije iteracije ovog algoritma. **Što od navedenoga vrijedi u stanju nakon te dvije iteracije?**

A $V(3) = q(3, \mathit{gore})$ i $V(3) > 0.15$ C $V(3) = q(3, \mathit{desno})$ i $V(3) < 0.18$

B $V(3) = q(3, \mathit{gore})$ i $V(3) < 0.18$ D $V(3) = q(3, \mathit{desno})$ i $V(3) > 0.15$

19 (P) Rešetkasti svijet sastoji se od četiri ćelije, označene brojevima 1 do 4 i poredane vodoravno slijeva nadesno. Ćelija 4 je terminalna i agentu donosi 1 bod; nema drugih bodova. Svijet je okružen zidovima. Agent može poduzeti akcije *lijevo* i *desno*. Ako bi ga akcija dovela na zid, ostaje u ćeliji u kojoj je bio. Agent uvijek kreće iz ćelije 1. Provodimo postupak Q-učenja. Početni iznosi Q-funkcije su 0. Parametri su $\alpha = 0.5$ i $\gamma = 1$. Agent najprije igra tri epizode u kojima uvijek bira akciju *desno* i uči. Nakon toga igra još 10 epizoda u kojima u svakoj ćeliji pohlepno odabire potez koji mu se čini optimalan i također uči. **Nakon provedenih svih 13 epizoda, koja je očekivana vrijednost** $q(3, \mathit{desno})$**?**

A Nije moguće odgovoriti B Između 0.75 i 0.9995 C Manja od 0.75 D Veća od 0.9995

20 (T) U modelu podržanog učenja, i politika agenta i okolina mogu biti ili stohastičke ili determinističke. Neka je politika agenta deterministička, a okolina neka je stohastička. **Na što se u tom slučaju svodi nova vrijednost kod učenja algoritmom vrednovanja politike?**

A $v_{k+1}(s) \leftarrow r(s, a, s') + \gamma \cdot v_k(s')$ C $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \cdot [r(s, a, s') + \gamma \cdot v_k(s')]$

B $v_{k+1}(s) \leftarrow \sum_{s'} p(s'|s, a) \cdot [r(s, a, s') + \gamma \cdot v_k(s')]$ D $v_{k+1}(s) \leftarrow \sum_a \pi(a|s) \sum_{s'} p(s'|s, a) \cdot r(s, a, s')$

Grupa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	D	C	A	D	A	C	A	D	A	B	A	A	B	B	C	B	B	B	A	B
B	A	B	B	D	D	A	B	A	B	B	B	C	C	B	A	D	D	C	B	C
C	A	D	B	C	A	C	B	B	D	D	A	C	C	B	B	A	A	A	C	C
D	C	A	D	B	A	D	A	C	B	C	C	A	A	A	B	C	C	A	A	C
E	A	C	A	D	B	C	B	A	D	D	B	B	D	C	A	C	A	A	A	D
F	D	C	A	A	B	B	D	B	C	D	D	B	B	D	A	C	C	A	A	C
G	D	A	B	D	A	B	D	A	C	C	A	A	A	C	C	A	C	C	B	A
H	B	A	D	C	B	C	D	B	C	C	A	B	D	C	D	B	B	D	D	B