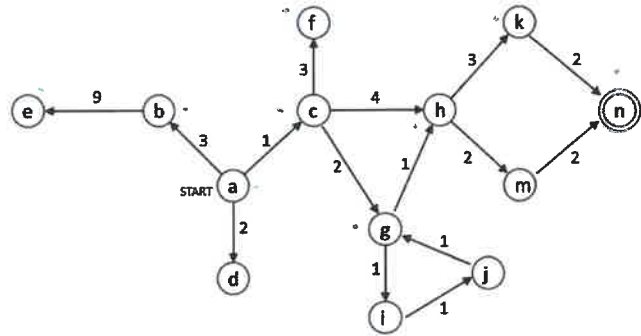


Odgovarjajte kratko in jedrnat, točno na zastavljena vprašanja. Vse odgovore pišite na črto pod vprašanji in **izključno na ta list**, ki ga edinega oddate na koncu! Čas pisanja: 75 minut.

(1) 25%

Na sliki je podan prostor stanj. Začetno vozlišče je a. Nasledniki nekega vozlišča se vedno generirajo po abecednem vrstnem redu (npr. iz c sledijo f, g in šele potem h), bodite pozorni na to! Preiskovalni algoritmi imajo na voljo tudi hevristično funkcijo, za katero vemo le, da je dopustna (angl. *admissible*).



5% (a) Koliko vozlišč razvije iskanje v globino?

9 (a, b, e, c, f, g, h, k, n)

5% (b) Katero rešitev vrne iterativno poglobljanje?

najkrajša, kraj: a → c → h → k → n  
*zaradi abecednega vrstnega reda*

5% (c) Katero rešitev vrne algoritem A\*?

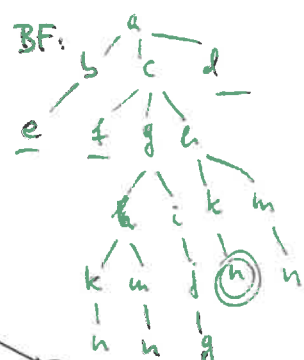
najcenejšo (dopustna hevristika), kraj: a → c → g → m → n (acghmn)

5% (d) Kakšna je cena rešitve, ki jo vrne algoritem IDA\*?

špet najcenejša, 8

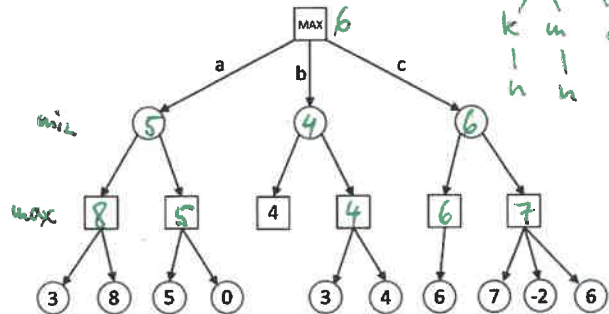
5% (e) Koliko vozlišč generira iskanje v širino?

20 (štejemo tudi začetno vozlišče, a)



(2) 10%

Podano je drevo igre. Številke v vozliščih predstavljajo hevristične ocene pozicije, ki jo vozlišče predstavlja. Na potezi je prvi igralec, MAX. Na voljo ima tri poteze, a, b in c.



5% (a) Katero potezo izbere kot najboljšo algoritem minimax?

c

5% (b) Kakšna je vrednost začetne pozicije, kot jo izračuna algoritem minimax z alfa-beta rezanjem?

6 (isto kot bez rezanja)

(3) 10%

Na podatkih  $X = [2, 3, 4, 8, 9, 11, 12, 15, 16, 17, 18]$  izvedemo algoritem k-voditelj (angl. *k-means*) za  $k=3$ . Začetni centroidi naj bodo  $C_1 = 2, C_2 = 5, C_3 = 11$ .

5% (a) Kakšne so končne vrednosti centroidov?

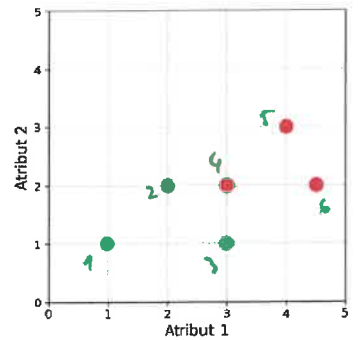
3, 10, 16.5

5% (b) Kakšna je velikost gruč, torej število primerov v posamezni gruči?

3 4 4  
000 1111 2222

(4) 15%

Na sliki je v obliki točk podana učna množica za dvorazredni problem. Primeri so opisani z dvema atributoma (osi X in Y), barva predstavlja razred. Za klasifikacijo uporabljamo algoritem k-NN, ki kot mero razdalje uporablja evklidsko razdaljo in uporablja navadno glasovanje (ne uteženo). »Model« želimo preveriti z uporabo metode izpusti enega (angl. *leave-one-out*; LOO).



2x5% = 10% (a) Kakšna je povprečna klasifikacijska točnost, kot jo izračuna LOO, če je k = 1 in kakšna, če je k = 3?

5% k=1: 50% (3/6) k=3: 67% (4/6)

5% (b) V splošnem, ali z večanjem šuma v podatkih želimo višati ali nižati parameter k?

Višati.

(5) 10%

Pot od začetnega do nekega ciljnega vozlišča (če obstaja) vrnejo kot rešitev preiskovalni algoritmi.

	1 NW	3 NW
1	T	T
2	F	T
3	F	F
4	F	F
5	T	T
6	T 3/6	T 4/6

5% (a) Kaj kot rešitev vračajo algoritmi za planiranje?

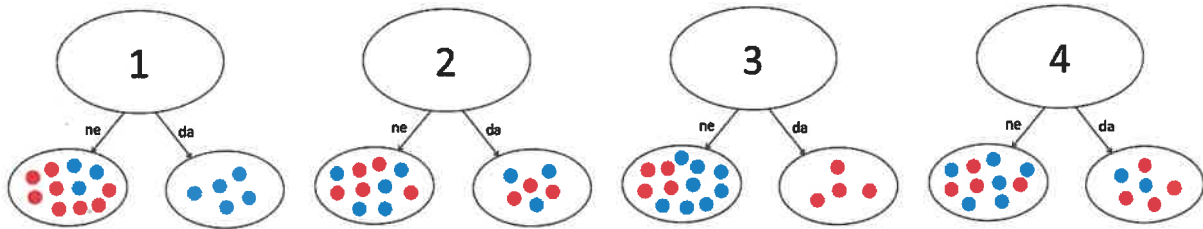
tudi pot, ki ji rečemo plan (= zaporedje akcij)

5% (b) Kaj pa kot rešitev vračajo algoritmi spodbujevanega učenja?

strategijo

(6) 10%

Spodaj so podane štiri možnosti, torej štirje atributi (1-4), ki jih algoritem za gradnjo odločitvenih dreves lahko uporabi za novo delitev (angl. *split*). Vsi atributi so binarni, kot je razvidno iz slike. V resnici obstaja še peta možnost, t.j. da algoritem sploh ne naredi delitve. Kako posamezen atribut razdeli učno množico je prav tako razvidno iz slike. Problem je dvorazreden, imamo rdeč in moder razred.



5% (a) Kateri atribut (ali nobenega) bo izbral algoritem za gradnjo odločitvenih dreves kot najboljši, če uporablja gini prispevek (angl. *gini gain*) kot mero nečistoče?

prvega; alternativno nobenega, če utemeljimo, da algoritem ne deli, če je ≤ 16 primerov v trenutnem vozlišču (razreje) i)

5% (b) Kakšen je gini prispevek drugega atributa v zgornjem primeru?

Obe veji sta 50:50; prispevek = 0

(7) 20%

Sedaj pa si predstavljamo, da slika iz prejšnje naloge predstavlja naključni gozd s štirimi drevesi (1-4).

5% (a) V kateri razred (če sploh lahko) bo ta naključni gozd klasificiral primer pri katerem v prvih dveh drevesih velja »da«, v drugih dveh pa »ne«, če uporablja navadno glasovanje?

Lahko, v modri razred.

5% (b) Sedaj pa imamo primer, kjer pri vseh drevesih velja »da«. Tokrat naključni gozd deluje po principu verjetnosti (soft max). Kakšna je povprečna verjetnost za modri razred?

100%; 50%; 0%; 33% → 123% / 4 = 45,75%

→ za modrega (!) je 45,75% (glej zgornj)

(c) Tudi tokrat naključni gozd deluje verjetnostno. Imamo naslednje štiri primere <da,da,da,da,rdeč>

<da,ne,da,da,moder>, <da,ne,ne,ne,moder> in <da,da,ne,ne,moder>. Pri tem da/ne pomeni v katero vejo pade v pripadajočem po vrsti drevesu, barva pa pomeni pravi razred. Zanima nas modri razred, mejo, da napovemo »modro« pa postavimo na 60%. Za modri razred izračunajte natančnost (angl. *precision*) in priklic (angl. *recall*).

natančnost: 2/2 = 100% priklic: 2/3 = 67%

5%

100%  
50%  
67%  
60% } > 60%

100%  
50%  
67%  
60% } > 60%

Vedno nas zanima verjetnosti za MODER razred

spet izlo kot tu

100%  
50%  
0%  
33%  
45,75%