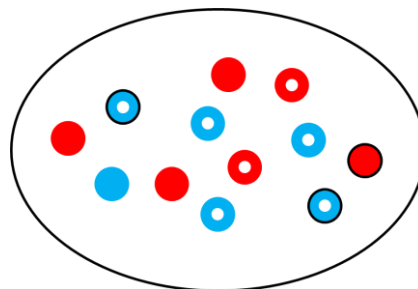


Odgovarjajte kratko in jedrnato, točno na zastavljena vprašanja. Vse odgovore pišite na črto pod vprašanji in **izključno na ta list**, ki ga edinega oddate na koncu! Čas pisanja: 75 minut.

(1)

Imamo dvorazreden problem, kjer nam na sliki razred predstavlja barva. Vsi primeri so opisani z dvema binarnima atributoma (da/ne), prvi atribut je »obroba« in drugi je »luknja«. Problem rešujemo z naključnim gozdom, ki pa ga moramo najprej zgraditi. Pri tem bo gozd majhen, sestavljen le iz dveh klasifikacijskih dreves, uporabljal pa bo gini prispevek kot mero nečistoče. Naključno izbrani učni množici za učenje teh dveh dreves sta prikazani na sliki. Drevesa »porežemo« tako, da omejimo njihovo največjo globino in sicer na globino največ ena (torej le ena delitev, angl. *split*).



- (5%) (a) V kateri razred uvrsti zgrajen naključni gozd primer z obrobo, a brez luknje, če deluje verjetnostno?

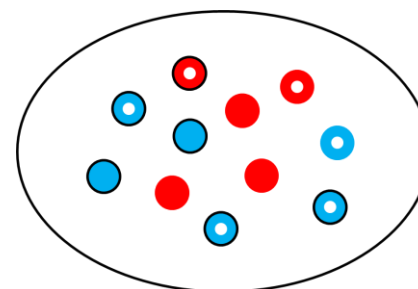
modri (1: 80% R; 2: 83% M)

- (5%) (b) Kakšno verjetnost, da je **rdeč** pripiše gozd, ki spet deluje verjetnostno, primeru brez obrobe in brez luknje?

80% R (1:80% R; 2: 80% R)

- (5%) (c) Kateri atribut (obroba, luknja) bi naključni gozd označil kot najpomembnejši? Ali pa sta enakovredna?

Obroba (pri "ne" je enaka nečistoča, pri "da" je boljša za obrobo)



(2)

V tabeli je podan model, ki smo ga zgradili s pomočjo naivnega Bayesa za isti problem kot v prejšnji nalogi. Napovedali bomo preprosto v bolj verjeten razred, torej je prag 50%.

		Ima obrobo?				Ima luknjo?			
p(M)	p(R)	p(da M)	p(ne M)	p(da R)	p(ne R)	p(da M)	p(ne M)	p(da R)	p(ne R)
60%	40%	90%	10%	30%	70%	80%	20%	20%	80%

- (5%) (a) Prvi primer ima obrobo, nima pa luknje. V kateri razred ga bo uvrstil podani model?

moder (27:24 oz. 9:8)

- (5%) (b) Za drugi primer vemo samo, da nima obrobe. Kaj bo zanj rekel podani model, če sploh lahko da napoved?

rdeč (3:14)

- (5%) (c) Želeli bi dvigniti prag za napoved v **moder** razred. Zato potrebujemo več kot samo primerjati razmerje obeh razredov, ampak potrebujemo dejansko verjetnost, ki jo vrne model. Kakšna je le-ta za prvi primer, torej tistega z obrobo, a brez luknje?

razmerje 9:8 za M, torej $9/(9+8) = 9/17 = 53\%$ M

- (10%) (d) Želimo oceniti klasifikacijsko točnost podanega modela. Uporabili bomo dvakratno prečno preverjanje, (testni) množici naj bosta takšni kot sta na sliki pri prejšnji nalogi. Model, ki bi ga sicer vsakič ponovno zgradili, naj bo kar tak kot je podan. Kakšna je klasifikacijska točnost na prvi množici, kakšna na drugi in kakšno vrne prečno preverjanje?

prva: $7/12 = 58\%$ druga: $9/11 = 82\%$ prečno preverjanje: $185/264 = 70\%$

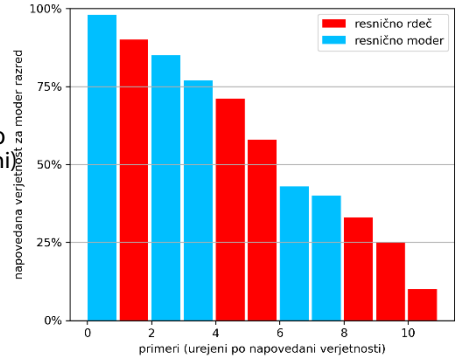
(3%)

(3%)

(6%)

(3)

Na desni sliki so podani primeri, urejeni po *napovedani* verjetnosti za **moder** razred. Za model, ki je podal te napovedi izračunajte:



(5%) (a) Klasifikacijsko točnost (razreda sta dva), če je meja za napoved 50%.
 $6/11 = 55\%$ (nad mejo 50 so prabilno klasificirane 3 modre, pod mejo pa so pravilno klasificirane 3 rdeče; ostali primeri so napačno kalsificirani)

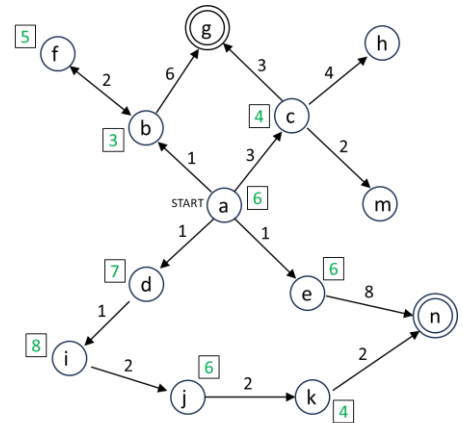
(10%) (b) Natančnost in priklic, ampak ob meji 75% (za napoved »modro«).
natančnost = $3/4 = 75\%$; priklic = $3/5 = 60\%$

(5%) (c) Kakšna je vrednost AUROC, torej ploščine pod ROC krivuljo?
 $22/30 = 11/15 = 73\%$

(Imam:
6 rdečih -> x-os razdelimo na 6 delov;
5 modrih -> y-os razdelimo na 5 delov; --> $6*5=30$
za vsak moder rišemo eno enoto gor, za vsak rdeč primer pa eno enoto desno.)

(4)

Na sliki je podan prostor stanj. Začetno vozlišče je **a**. Nasledniki nekega vozlišča se *vedno* generirajo po abecednem vrstnem redu. Bodite pozorni, povezava **b-f** je dvosmerna; cena je enaka v obe smeri. Preiskovalni algoritmi imajo na voljo tudi hevristično funkcijo, ki je podana z zeleno barvo ob vsakem vozlišču, razen pri ciljnih vozliščih, kjer je enaka 0 in »slepih ulicah«, kjer je vrednost enaka 100.



(5%) (a) Ali iskanje v širino in iterativno poglobljanje vrneta isto rešitev?
Da (a -> b -> g)

(5%) (b) katero rešitev vrne iskanje v globino?
zacikla se, nobene

(5%) (c) katero rešitev vrne algoritem A*, če za »tie-break« pravilo uporablja vrstni red generiranja vozlišč?
a -> c -> g

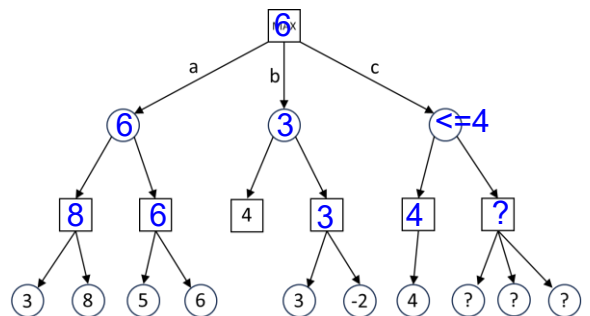
(5%) (d) Katera vozlišča (s ponavljanjem, seveda) in v kakšnem vrstnem redu *razvije* algoritem IDA*?
a,b,a,b,g

(5%) (e) Je podana hevristična funkcija monotona? Če ni, razložite zakaj ni.
Ni monotona; protiprimeri: k-->n; a-->b; c-->g; in še kakšen

(5%) (f) Če algoritem planiranja uporablja iterativno poglobljanje, kakšna je njegova *prostorska* zahtevnost?
ista kot za ID, torej ista kot za DF, torej $O(d)$ ali $O(b*d)$

(5)

Podano je drevo igre. Številke v vozliščih predstavljajo hevristične ocene pozicije, ki jo vozlišče predstavlja, znak »?» pa pomeni, da vrednosti za to vozlišče ne vemo. Na potezi je prvi igralec, MAX. Na voljo ima tri poteze, **a**, **b** in **c**.



(5%) (a) katero potezo izbere kot najboljšo algoritem minimax, če je to sploh možno ugotoviti?
možno; izbere a

(5%) (b) Kakšna je vrednost začetne pozicije, če je to sploh mogoče ugotoviti?
možno, $\max\{a(6), b(3), c(<=4)\} = 6$