

Umetna inteligenca, pisni izpit, 22.1.2014

Literatura ni dovoljena, razen na roko in s svinčnikom napisanega lista A4, ki ga je podpisane potrebno oddati z izpitom. Naloge so enakovredne. Čas pisanja 90 minut. Ustni izpit za študente, ki so dosegli vsaj 50% točk na pisnem izpitu in želijo izboljšati oceno, bo v sredo, 29. 1. 2014 ob 11:00 v sobi prof. Robnika Šikonje na Jadranski.

- 1) a) V čem se razlikuje klasifikacijska točnost na učni množici, od klasifikacijske točnosti na testni množici? Kaj nam pri tem pomaga točnost večinskega klasifikatorja?
b) V čem se razlikuje povprečna kvadratna napaka od relativne povprečne kvadratne napake regresorja? Kakšne so sprejemljive vrednosti za slednjo?
c) V čem se razlikujejo klasifikacijska točnost, Brierjeva mera in informacijska vsebina odgovora klasifikatorja?
- 2) Da bi zagovili likvidnost banke, so začeli zbirati podatke o posojilojemalcih in o tem, ali je bilo posojilo vrnjeno. Zbrali so podatke v spodnji tabeli.

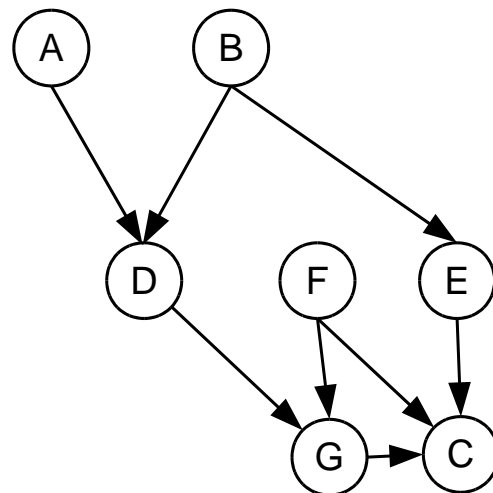
Vrnjeno posojilo	spričevalo OŠ	V kazenskem postopku	znesek posojila > 3 M€	Sorodnik direktorja banke	Barva avta	Spol
DA	DA	DA	NE	NE	zelena	ž
NE	DA	DA	DA	DA	črna	m
DA	NE	NE	NE	NE	črna	m
NE	NE	DA	DA	DA	črna	m
DA	ponarejeno	DA	DA	DA	črna	m
NE	ponarejeno	DA	DA	DA	črna	ž

S pomočjo naivnega Bayesa, naučenega na zgornji učni množici z m-oceno z $m=2$, izračunaj verjetnost:

- A) da bo nečakinja direktorja banke, ki se vozi v črnem avtu vrnila 9 M€ posojila.
 - B) da bo neznani moški, ki ima ponarejeno spričevalo OŠ in je v kazenskem postopku, vozi pa se s črnim BMW, vrnil 50 M€ posojila.
 - C) neobvezno (za dodatnih 10 točk): da direktorjev sin, ki seveda ni vrnil posojila in ima ponarejeno spričevalo OŠ, vozi črni avto.
- 3) Direktor banke se je odločil, da bo napovedoval verjetnost, da bo posojilo vrnjeno, kar s pomočjo odločitvenega drevesa, ki ga je zgradil z uporabo gini-indeksa iz zgornje podatkovne baze o posojilojemalcih.
 - a. Nariši direktorjevo odločitveno drevo.
 - b. Kakšno verjetnost napove direktor za svojo ženo brez izobrazbe in brez kazenskega postopka, ki se vozi v zelenem porscheju in si je sposodila toliko, da bo lahko kupila še enega roza porscheja?

- 4) a) Za podani graf utemeljeno odgovorite, ali bi lahko predstavljal Bayesovo mrežo. Če mreže ne more predstavljati, ga popravite tako, da jo bo predstavljal.

- b) Za po potrebi popravljeni graf, ki predstavlja Bayesovo mrežo, zapišite izraz, kako bi s pomočjo mreže učinkovito izračunali verjetnost $P(\neg ABCD \neg EF \neg G)$. Upoštevajte, da so vse spremenljivke binarne.



- 5) V mednarodni organizaciji različne interese zastopa 100 predstavnikov, ki so razdeljeni v neformalne skupine Zeleni, Rdeči, Rumeni in Kockasti, s številom glasov iz spodnje tabele. Njihove preference glede možnosti A, B, C in D so prav tako opredeljene v tabeli (1 - najmanj zaželena možnost, 4 – najbolj zaželena možnost). Izračunajte, koliko glasov bo posamezna izbira prejela z različnimi načini glasovanja in rezultate vpišite v tabelo.

		Preference glede izbir			
Skupina	Glasov	A	B	C	D
Zeleni	21	3	4	1	2
Rdeči	35	4	3	2	1
Rumeni	17	2	1	3	4
Kockasti	27	1	2	3	4
Večinsko glasovanje					
Bordajevo štetje					
Condorcetovo glasovanje					
Potrjevanje prvih dveh					

- 6) Opišite vektorsko predstavitev dokumentov pri procesiranju naravnega jezika in pojasnite, kako deluje uteževanje posameznih besed na način tf-idf (term frequency – inverse document frequency). Podajte primer.