

1. Za letošnje počitnice smo se odločili ogledati Turčijo. Ker bomo šli s svojim avtom, nas pot vodi preko Bolgarije. Na cesti nas stalno ustavljajo uniformirani tipi – nekateri so policisti, nekateri morda tudi ne – in od nas pobirajo kazni za različne nenavadne prometne prekrške.

avto	uniforma	kapa	brki	policist?
Opel	da	ne	ne	da
Opel	ne	ne	ne	ne
Opel	da	ne	da	da
Opel	da	ne	da	ne
Opel	da	ne	da	da
Golf	ne	ne	ne	ne
Golf	da	ne	da	ne
Golf	da	ne	ne	da
nič	da	ne	da	ne
nič	da	ne	da	da
nič	ne	ne	ne	ne
nič	da	ne	ne	da

- a. Na podlagi zbranih podatkov sestavite naivni Bayesov klasifikator in se odločite, ali boste ustavili naslednjima »policistoma« – oba imata uniformo in nimata kape, prvi pozira pred Oplom in ima brke, drugi nima ne Opla ne brkov.
- b. Kateri kriteriji so bolj in kateri manj uporabni? Če to ne bi bilo očitno – na kakšen način bi lahko prišli do odgovora na to vprašanje?

2. Sovražnik strahopetno obstreljuje civilne cilje v našem glavnem mestu. Da bi se lahko pripravili na napade, smo na njegovo ozemlje poslali šest vohunov. Vohuni ne vedo eden za drugega, nahajajo se na različnih mestih in, predpostavljamo, nimajo istega vira informacij, med sabo so torej povsem neodvisni. Tabela kaže napovedi posameznih vohunov za prvih dvajset dni junija (vsak stolpec predstavlja enega vohuna, × pomeni napad) in ali je tisti dan dejansko prišlo do napada ali ne.

	napovedi vohunov						napad?
1. junij	×	×	×		×	×	×
2. junij			×		×		
3. junij				×	×	×	×
4. junij	×	×	×	×	×	×	×
5. junij	×				×	×	×
6. junij	×	×			×		×
7. junij			×				
8. junij	×				×	×	×
9. junij		×	×	×	×		×
10. junij	×	×					
11. junij		×		×		×	
12. junij		×	×	×	×	×	×
13. junij			×				
14. junij		×				×	
15. junij	×	×		×	×	×	×
16. junij	×	×	×	×	×	×	×
17. junij	×				×		
18. junij					×	×	×
19. junij	×		×	×	×	×	×
20. junij	×				×	×	
23. junij		×	×	×		×	?
24. junij	×	×		×	×	×	?
25. junij	×	×			×	×	?
26. junij	×	×		×	×		?
27. junij	×		×			×	?

- a) Sestavite ustrezen model, s katerim lahko napoveste verjetnosti napadov v naslednjih dneh junija. Kakšne so?
- b) Komunikacija z vohuni je tvegana, zato bo morda potrebno odpoklicati manj zanesljive. Sestavite vrstni red vohunov glede na zanesljivost njihovih napovedi. Katerega bomo odpoklicali prvega? Če bomo obdržali le enega – kateri bo? Pri tem si lahko pomagata s sestavljenim modelom ali pa z zdravo pametjo, odgovor pa v vsakem primeru računsko utemeljite.

3. Banka želi izboljšati merila, na podlagi katerih odobravajo kredite. V ta namen je zbrala podatke o doslejšnjih kreditih. Vsak kreditojemalec je opisan s stopnjo izobrazbe (nizka, srednja, visoka), višino plače (nizka, srednja, visoka) in tem, ali je to njihov prvi kredit (da, ne). Za vsako kombinacijo teh vrednosti so prešteli, koliko kreditojemalcev je kredit redno vračalo in koliko ne. Podatki so zbrani v tabeli. Na primer, med takimi s srednjo izobrazbo, nizko plačo in je to njihov prvi kredit, je 25 klientov vračalo kredit, 12 pa ne.

izobr.	plača	prvi?	ne	da
nizka	nizka	ne	16	19
		da	32	33
	srednja	ne	8	9
		da	7	9
	visoka	ne	1	5
		da	3	11
srednja	nizka	ne	18	23
		da	12	25
	srednja	ne	5	20
		da	5	20
	visoka	ne	0	0
		da	0	4
visoka	nizka	ne	5	3
		da	4	0
	srednja	ne	0	3
		da	1	1
visoka	ne	0	2	
	da	1	5	

Za modeliranje bi radi uporabili naivni Bayesov klasifikator; če boste nalogo reševali tudi na enostavnejši način, vam bodo za to dali manj točk.

- a. Če podeljujemo kredite le tistim, pri katerih je vsaj dve tretjini možnosti, da ga bodo redno vračali – kdo bi ga dobil?
- b. Če opazujemo stranke le glede na izobrazbo: kateri sloj je najmanj tvegan?
- c. Bi kredit raje podelili nekemu s srednjo izobrazbo, visoko plačo, ki je že prejemal kredite ali nekemu z visoko izobrazbo in plačo, ki še ni imel kreditov?