

NALOGA 1

Zaporniki se na otoku lahko prosto gibljejo. Do kopnega je 8 km in pogosto lahko vidimo v morju morske pse. Nekateri zaporniki so poskušali pobegniti s plavanjem ponoči, a samo nekaterim uspe doseči kopno. V tabeli so opisi nekaterih poskusov pobega.

Dober plavalec	Dežuje	Polna luna	Podnevi videni morski psi	So ga opazili pazniki	Je dosegel kopno
DA	DA	DA	DA	DA	DA
DA	NE	DA	NE	NE	DA
DA	NE	NE	NE	DA	NE
NE	NE	DA	NE	DA	NE
DA	NE	NE	DA	NE	NE
NE	DA	DA	DA	NE	NE

Uporabi Naivnega Bayesa z m-oceno in naj bo $m=2$, da oceniš verjetnost:

- da je zapornik, ki je dober plavalec, dosegel kopno, če ni deževalo in je bila polna luna, podnevi ni bilo morskih psov in so ga med begom opazili pazniki.
- da je bil zapornik, ki je dosegel kopno med polno luno, dober plavalec.

NALOGA 2

Na izpit je prišlo 100 študentov. 60 študentov se je učilo veliko, 20 študentov srednje in 20 se jih sploh ni učilo. Recimo, da so to tri skupine študentov. Vsak študent dobi 1 vprašanje. Vprašanje je težko, srednje težko ali lahko. V vsaki od treh skupin študentov jih je polovica dobila težko vprašanje, četrtna srednje težko in četrtna lahko vprašanje. Na težko vprašanje so pravilno odgovorili samo tisti, ki so se veliko učili. Na srednje težko vprašanje so pravilno odgovorili vsi, ki so se veliko učili, 80% tistih, ki so se srednje veliko učili, in 20% tistih, ki se niso nič učili. Na lahko vprašanje je nepravilno odgovorilo samo 50% tistih, ki se niso nič učili.

Kakšna verjetnost, da bo študent, ki se je učil srednje veliko in je dobil srednje težko vprašanje, pravilno odgovoril, če pri izračunu naivni Bayes uporablja m-oceno in je $m = 2$?

NALOGA 3

V morju je potapljač naletel na razbitino gusarske ladje in v notranjosti je bil zaboj poln cekinov. Ker je bil zaboj pretežak, je vzel s seboj samo 101 naključno izbranih cekinov. Na obali je pregledal cekine in ugotovil, da jih je 50 srebrnih, 30 bronastih in 20 zlatih, enega pa je izgubil. Pri srebrnih je ugotovil, da ima polovica vtisnjen simbol Črnega gusarja, pri bronastih tretjina in pri zlatih tri četrtine. Ker je posumil, da so nekateri ponarejeni, jih je še stehal. Od srebrnih jih je bila petina lažjih (ponarejenih) in vsi so imeli vtisnjen simbol gusarja, od bronastih ni bil nobeden ponarejen in od zlatih polovica in vsi so imeli vtisnjen simbol gusarja. Naj bo to naša učna množica (2 atributa: vrsta cekina, simbol gusarja; 2 razreda: pravi/ponarejen; 100 učnih primerov).

- a) Nariši odločitveno drevo, ki vsebuje celotno zgoraj opisano informacijo, torej v vsakem listu mora biti napisano število cekinov, ki ustreza temu listu in verjetnost pravilnega odgovora. Pri gradnji drevesa za izbiro atributa uporabi informacijski prispevek,
- b) Z odločitvenim drevesom klasificiraj 101. cekin, ki ga je potapljač kasneje našel na obali, kjer ga je bil izgubil. Cekin je bil srebrn z vtisnjenim simbolom gusarja,

NALOGA 4

Da bi zagotovili, da študenti pridejo pripravljeni na izpit, so začeli pregledovati vsakotedenske domače naloge. V spodnji tabeli so podatke o dosežkih študentov pri domačih nalogah in pisnem izpitu.

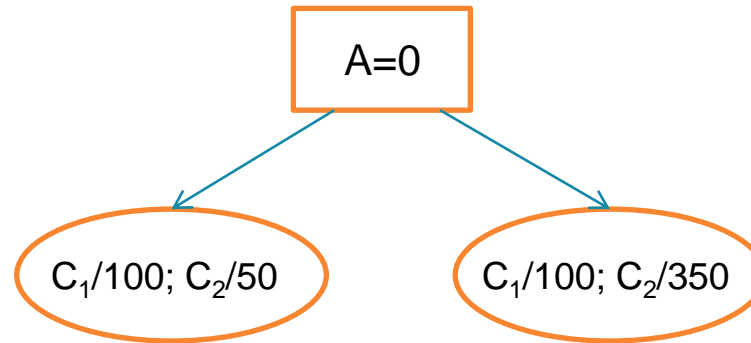
Izdelal vse obvezne DN	Izdelal vse neobvezne DN	Pozitivno ocenjene narejene obvezne	Pozitivno ocenjene narejene neobvezne	Izdelal zadnjo neobvezno	Opravili pisni izpit
DA	DA	DA	DA	DA	DA
DA	NE	DA	DA	DA	DA
DA	NE	NE	NE	NE	NE
NE	DA	DA	NE	DA	NE
DA	NE	DA	DA	DA	DA
NE	NE	NE	DA	DA	NE

Profesor se je odločil, da bo na ustnem izpitu namesto spraševanja študenta raje uporabil odločitveno drevo, ki ga je zgradil na gornji učni množici z uporabo gini-indeksa, s tem da bo odločitev, ki je bila sicer namenjena pisnemu izpitu, uporabil kar za ustni izpit. Pri tem je verjetnost, da bo študent opravil, spremenil v oceno ($<50 \rightarrow 5$, $\geq 50 \rightarrow 6$, $\geq 60 \rightarrow 7$, ..., $\geq 90 \rightarrow 10$).

- Nariši profesorjevo odločitveno drevo.
- Kakšno oceno je na ustnem izpitu dobil študent, ki je izdelal vse domače naloge razen zadnje neobvezne in je imel vse domače naloge, ki jih je izdelal, ocenjene pozitivno?

NALOGA 5

Spodaj je odločitveno drevo s frekvencami primerov v listih za razreda C1 in C2:



Izračunaj pričakovano klasifikacijsko točnost drevesa z m-oceno, če je:

- a) $m = 0$,
- b) $m = 1000$