**Izpit pri predmetu Odločitveni sistemi** (10.9.2015)

**Navodila:**

Izpit se piše 90 minut. Literatura ni dovoljena. Vrednost naloge je označena pri vsaki nalogi.

Ogledi pisnih izpitov: petek, 11.9.2015, ob 10h v laboratoriju za umetno inteligenco

**Naloga 1 (30 točk): Programerji**

Na razpis za delovno mesto programerja se je prijavilo pet kandidatov. Ovrednotili smo jih po kriterijih "prvi vtis" (15 %), "znanje programiranja v Basicu" (10 %), "znanje programiranja v Pythonu" (50 %), "obvladovanje podatkovnih zbirk" (20 %) in "poprečna ocena na fakulteti" (5 %).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **prvi vtis** | **Basic** | **Python** | **pod. zbirke** | **popr. ocena** | | |
| **Jambrek** | 50 | slabo | dobro | povprečno | | 7 |  |
| **Godler** | 20 | slabo | dobro | povprečno | | 8 |  |
| **Bobovnik** | 100 | povprečno | povprečno | povprečno | | 9 |  |
| **Kogovšek** | 10 | povprečno | povprečno | povprečno | | 10 |  |
| **Stefančič** | 50 | slabo | dobro | povprečno | | 6 |  |

(Opomba: podatki so izmišljeni naključno!)

a) Razvrstite variante glede na leksikografsko metodo, za vrstni red upoštevajte uteži pri posameznih kriterijih.

b) Poiščite morebitne manjvredne variante in jih izločite. Za vsako izločeno varianto pojasnite, katera varianta jo izloči.

c) Ocenite variante z modelom uteženih vsot (linearnim modelom).

d) Naredite analizo občutljivosti glede na kriterij "prvi vtis". Ali je model občutljiv? Razložite!

**Naloga 2 (20 točk): Odkrivanje bolezni**

Za ugotavljanje neke bolezni sta na voljo dva testa. Prvi test v 25 % testiranih reče, da je bolezen prisotna,

pri 55 %, da je ni, v 20 % pa da neodločen rezultat. Pri drugem testu so ti deleži 30 %, 60 % in 10 %.

Testa sta neodvisna, zato ju nameravamo uporabiti tako, da najprej naredimo enega, če da le-ta neodločen

rezultat, pa pač uporabimo še drugega. Cena prvega testa je 800 EUR, drugega pa 1000 EUR.

a) Narišite ustrezen model in povejte, v kakšnem vrstnem redu naj izvajamo testa, da bo testiranje čim cenejše? Koliko denarja bomo v poprečju porabili za testiranca?

b) Kakšna je verjetnost, da po obeh testih še vedno ne bomo imeli diagnoze?

**Naloga 3 (25 točk): priporočilni sistemi**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | izdelek 1 | izdelek 2 | izdelek 3 | izdelek 4 |
| uporabnik A | 5 | ? | 4 | ? |
| uporabnik B | 3 | ? | 2 | 1 |
| uporabnik C | 1 | 2 | 4 | 5 |
| uporabnik D | 4 | 4 | 3 | 5 |

1. Na podlagi »item-based collaborative filtering« tehnike napovejte kakšno oceno bi uporabnik A dal izdelkoma 2 in 4. Za mero podobnosti uporabite kosinusno mero za računanje podobnosti.

Formula (a in b sta različna produkta, U je množica uporabnikov, ki so ocenili oba izdelka):

1. Uporabnik A je naknadno ocenil še izdelka 2 in 4; izdelku 2 je dal oceno 4, izdelku 4 pa 5. Na podlagi teh dveh ocen izračunajte MAE in MSE za priporočilni sistem iz točke a).

**Naloga 4 (25 točk)**: Za izboljšanje našega priporočilnega sistema smo se odločili opazovati naslednja povezovalna pravila:

|  |
| --- |
| če gleda komedije, potem gleda mehiške nadaljevanke |
| če gleda mehiške nadaljevanke, potem gleda komedije |
| če komedije, potem gleda nogomet |
| če gleda nogomet, potem gleda komedije |
| če nogomet, potem gleda mehiške nadaljevanke |
| če gleda mehiške nadaljevanke, potem gleda nogomet |

Opravili so raziskavo, v kateri je sodelovalo 580 ljudi. Rezultate kaže spodnja tabela.

|  |  |  |  |
| --- | --- | --- | --- |
| **nogomet** | **komedije** | **mehiške nadaljevanke** | **N** |
| ne | ne | ne | 88 |
| ne | ne | da | 96 |
| ne | da | ne | 115 |
| ne | da | da | 77 |
| da | ne | ne | 114 |
| da | ne | da | 5 |
| da | da | ne | 65 |
| da | da | da | 20 |

a. Za pravila, ki imajo podporo vsaj 10%, izračunajte zaupanje (confidence) in dvig (lift) gornjih pravil!

b. Katero pravilo je najbolj uporabno za priporočilni sistem? Kako bi ga uporabili?