

Istraživanje podataka, teorijski deo ispita, jul 2022. (II termin)

Indeks, ime i prezime studenta: _____

/10 1. Definisati šta je istraživanja podataka i navesti osnovne korake koji se primenjuju u procesu istraživanja podataka.

/10 2. Kako i zašto se vrše standardizacija i normalizacija podataka?

/8 3. Opisati šta se i kako prikazuje pomoću šema sa matricama (toplotne mape). Nacrtati primer.

/10 4. Opisati kako se vrši podela čvora po neprekidnim atributima kod drveća odlučivanja.

/10 5. Opisati šta je model kod naivnih Bajesovskih klasifikatora za skup sa imenskim atributima i kako klasifikuju novu instancu.

/10 6. Opisati matricu konfuzije. Nacrtati primer.

/10 7. Definisati senka koeficijent. Kako se tumače moguće vrednosti senka koeficijenta?

/15 8. Napisati i objasniti pseudokod sakupljajućeg algoritama za klasterovanje. Opisati metod *najkraća veza* za računanje sličnosti klasterovanja i navesti prednosti i mane ove veze.

/9 9. Šta je antimonotonost u kontekstu čestih skupova stavki i kako je koristi algoritam Apriori?

/8 10. Opisati kako i zašto se koristi hijerarhija koncepata kod pravila pridruživanja.