

## Subiectul I

## 1. Completați afirmațiile de mai jos:

- În cazul unei probleme de clasificare a tweet-urilor în pozitive și negative pe baza prezenței sau absenței a 3 *hash tag*-uri (#happy, #fun, #sad), un arbore de decizie poate avea maxim ..... nivele (nodul rădăcină se află pe nivelul 1)
- În algoritmi evolutivi, scopul operatorului de selecție este de a favoriza supraviețuirea celor mai bune soluții, pe când scopul operatorului de mutație este .....
- TF.IDF este un concept care poate fi folosit când se procesează informații de tip .....
- tehnica Greedy efectuează o traversare a spațiului de căutare arborescent folosind metoda .....

2. Se consideră o hartă cu  $n$  orașe, pentru fiecare oraș cunoscându-se numărul de obiective turistice. Să se distribuie toate aceste orașe în  $k$  comunități turistice (o comunitate este format din unul sau mai multe orașe) de aproximativ același volum total al obiectivelor turistice. Un oraș poate să aparțină unei singure comunități. Precizați care dintre următoarele afirmații sunt adevărate în cazul folosirii unui algoritm genetic pentru rezolvarea problemei.

- Un cromozom va avea o reprezentare matriceală cu  $k \times n$  elemente binare, iar suma elementelor de pe oricare coloană este 1
- Un cromozom va avea o reprezentare vectorială cu  $k$  elemente, fiecare element fiind o valoare din  $\{1, 2, \dots, n\}$ , cu condiția ca fiecare valoare din  $\{1, 2, \dots, n\}$  să apară cel puțin o dată.
- Un cromozom va avea o reprezentare vectorială cu  $n$  elemente, fiecare element fiind o valoare din  $\{1, 2, \dots, k\}$ , cu condiția ca fiecare valoare din  $\{1, 2, \dots, k\}$  să apară cel puțin o dată.
- Un cromozom va avea o reprezentare matriceală cu  $n \times k$  elemente binare, iar suma elementelor de pe oricare linie este 1.

## 3. Prezentați o deosebire între

- operatorii de încrucișare și selecție folosiți în algoritmi genetici

- algoritmul de optimizare bazat pe gradient folosit în Regresie și cel folosit în Regresia logistică

4. Care dintre următoarele concepte pot reprezenta derivări (extinderi) ale conceptului de *Metodă de învățare supervizată*:

- Încrucișare
- Rețea Neuronală Artificială
- Algoritm Evolutiv
- K-means

## 5. Precizați câți coeficienți trebuie optimizați în cazul folosirii unei rețele neuronale artificiale cu un strat ascuns cu 2 neuroni pentru predicția consumului de combustibil a unor vehicule pe baza puterii motorului și a anului de fabricație a vehiculului. Exclueți din calcule bias-ul.

- 4
- 6
- 2
- 8

6. Precizați dacă un perceptron poate învăța un model de clasificare bazat doar pe  $x_1$  și  $x_2$  a cărui acuratețe să fie 100% folosind datele din figura alăturată. Justificați răspunsul7. O rețea neuronală artificială trebuie antrenată pentru a prezice numărul de excursii anuale efectuate de o persoană folosind informații despre suprafața locuinței și salariul său, prezentate în tabelul următor. Rețeaua neuronală are un strat ascuns cu doi neuroni, funcția de activare este funcția identică ( $f(x) = x$ ), iar fiecare pondere inițială este 0.5. Care este eroarea medie pătratică după prima epocă de antrenament a rețelei folosind algoritmul de gradient descrescător?

Suprafață locuință	Salariu	#excursii
70	150	4
50	145	5
65	135	3
55	130	4