

Examen TBD - 3 Iulie 2020

1 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Într-un spital se estimează că rata internărilor este 1 pacient/oră. Folosind distribuția Poisson, probabilitatea ca rata reală să fie de 2 pacienți/oră este:

Alegeți o opțiune:

- ☐ a. 0.18
- ☐ b. 0.01
- ☐ c. 0.36
- ☐ d. 0.03

2 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Rolul etapei de **mapare**, caracteristică MapReduce, este de a

Alegeți o opțiune:

- ☐ a. selecta din datele de intrare numai acele valori care sunt reprezentate sub forma <key, value>
- ☐ b. selecta datele de interes sub forma unei reprezentări de tipul <key, value>.
- ☐ c. selecta datele de interes, fără a le altera reprezentarea.
- ☐ d. modela datele de intrare sub forma unei reprezentări de tipul <key, value>.

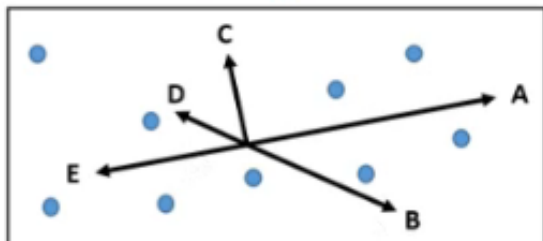
3 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Fie următorul sistem de puncte 2D:



Știind că vectorul A este prima componentă principală, care este a doua?

Alegeți o opțiune:

- ☐ a. D
- ☐ b. E
- ☐ c. B
- ☐ d. C

4 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Pentru proiecția datelor dintr-un spațiu N-dimensional într-un spațiu M-dimensional, $M < N$, transformarea de proiecție este:

Alegeți o opțiune:

- ☐ a. Prima componentă principală a datelor
- ☐ b. Matricea de covarianță a datelor
- ☐ c. Produsul dintre matricea de covarianță a datelor și prima componentă principală
- ☐ d. A matrice formată din componentele principale ale datelor

5 întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Pentru inițializarea numărului de clustere în cadrul K-means, se obțin următoarele valori ale siluetelor medii ale clusterelor: $k=2$: 0.2, 0.5 ; $k=3$: 0.1, 0.6, 0.5 ; $k=4$: 0.3, 0.1, 0.4, 0.4. Numărul optimal al clusterelor este:

Alegeți o opțiune:

- ☐ a. 3
- ☐ b. 4
- ☐ c. infinit
- ☐ d. 2

6 întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Fie următoarea secvență: 2 5 1 2 3 4 5 2 furnizată sub forma unui flux de valori întregi. Valorile pe care le prelucrează procesorul de fluxuri sunt: 2 5 1 2 1 2 3 4 3 4 5 2 . Pentru prelucrarea secvenței s-a utilizat:

Alegeți o opțiune:

- ☐ a. O fereastră cu deplasare (*tumbling window*)
- ☐ b. O fereastră ce presupune degradare
- ☐ c. O fereastră glisantă și una cu deplasare
- ☐ d. O fereastră glisantă

7 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Considerând o adaptare a BFS pentru algoritmul lui Dijkstra, câte rulări complete ale aplicației MapReduce sunt necesare pentru a determina complet toate drumurile de cost minim într-un digraf?

Alegeți o opțiune:

- ☐ a. Aplicația se rulează până când nu mai intervin schimbări ale distanțelor determinate pentru că rezultatele devin consistente atunci când au fost analizate toate posibilitățile.
- ☐ b. Sunt necesare <diametrul digrafului> rulări pentru că în acest mod acoperim toate posibilitățile.
- ☐ c. Aplicația se rulează până când nu mai intervin schimbări ale distanțelor determinate pentru că Dijkstra este un algoritm de tip greedy.
- ☐ d. Aplicația poate fi rulată o singură dată.

8 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Care ar fi numărul total de iterații complete ale unei aplicații MapReduce necesar pentru a implementa parcurgerea BFS a unui digraf?

Alegeți o opțiune:

- ☐ a. Diametrul digrafului runde, pentru digrafuri ponderate.
- ☐ b. O singură iterație completă MapReduce.
- ☐ c. Diametrul digrafului runde, dacă această valoare ar fi cunoscută.
- ☐ d. Diametrul digrafului runde, pentru digrafuri neponderate.

9 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Care dintre următoarele sunt exemple de compresie fără pierderi?

Alegeți o opțiune:

- ☐ a. JPG, MP3
- ☐ b. RLE, JPG
- ☐ c. ZIP, RLE
- ☐ d. ZIP, JPG

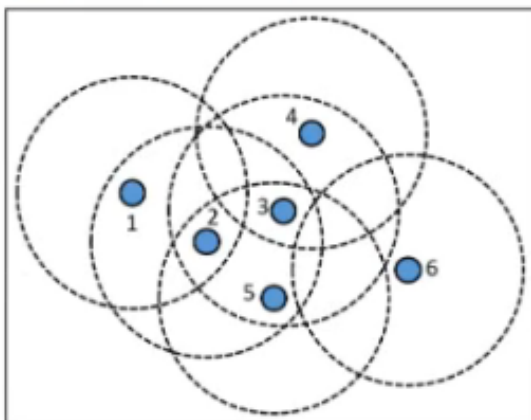
10 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Alegeți afirmația corectă privind clusterizarea DBSCAN din imagine, unde $\text{minP}=3$:



Alegeți o opțiune:

- ☐ a. P1 este accesibil din P4
- ☐ b. P1, P4, P6 sunt puncte frontieră
- ☐ c. P3, P4, P5 sunt puncte nucleu
- ☐ d. P1 este accesibil din P6

11 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Alegeți afirmația corectă legată de clusterizarea expectativă-maximizare (Expectation-Maximization):

Alegeți o opțiune:

- ☐ a. Fiecare punct poate aparține fiecărui cluster indiferent de valorile parametrilor distribuțiilor
- ☐ b. Etapa de maximizare presupune estimarea ponderată a parametrilor distribuțiilor clusterelor
- ☐ c. Cu cât abaterile standard ale distribuțiilor estimate ale clusterelor sunt mai mici, cu atât crește probabilitatea ca punctele să le aparțină
- ☐ d. Etapa de expectativă presupune calculul coeficienților de responsabilitate pentru toate clusterelor

12 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

În fazele inițiale ale rulării unui model MapReduce, nodul cu rol de coordonator divide datele de intrare într-un număr de subseturi disjuncte. Se recomandă ca acest număr de partiții să fie considerabil mai mare decât numărul de worker-i asigurați unei etape. Printre avantajele acestei abordări se numără și

Alegeți o opțiune:

- ☐ a. reducerea complexității algoritmilor implicați în etapele de echilibrare a încărcării sistemului.
- ☐ b. minimizarea timpilor implicați în fazele de sincronizare dintre etapele fundamentale ale modelului.
- ☐ c. minimizarea timpilor implicați în scenariile de recuperare în caz de erori.
- ☐ d. reducerea complexității etapelor de calcul caracteristice etapelor fundamentale ale modelului.

13 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Cum pot fi reprezentate grafurile, conform modelului MapReduce, pentru a putea implementa un algoritm de tip BFS?

Alegeți o opțiune:

- ☐ a. Cheia este id-ul unui nod, iar valoarea cuprinde lista de adiacență a nodului și distanța din nodul sursă.
- ☐ b. Cheia este id-ul unui nod, iar valoarea este distanța determinată până la vecinii nodului.
- ☐ c. Cheia este id-ul unui nod, iar valoarea este distanța din nodul sursă.
- ☐ d. Cheia este id-ul nodului sursă, iar valoarea cuprinde distanța determinată până la fiecare nod din graf.

14 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Un model de clasificare afectat de suprapotrivire (overfitting):

Alegeți o opțiune:

- ☐ a. Are o margine de decizie linară
- ☒ b. Clasifică corect datele de antrenare dar nu și pe cele de test
- ☐ c. Are erori de clasificare prea mari pentru datele de antrenare
- ☐ d. Are acuratețe mare la clasificarea datelor de test

[Clear my choice](#)

15 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Care dintre următoarele NU reprezintă un set valid de coduri Huffman?

Alegeți o opțiune:

- ☐ a. 011, 10, 01, 0001
- ☐ b. 011, 010, 10, 11
- ☐ c. 100, 01, 001, 1010
- ☐ d. 0011, 1101, 101, 010

16 întrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Se consideră următoarea tranzacție izolată dintr-un set supus unei analize de identificare a tiparelor frecvente:

{a, b, c}

Considerând scenariul **unei singure rulări** a unei soluții MapReduce pentru problema menționată, care dintre variantele de mai jos reprezintă rezultate intermediare posibile al etapei de **mapare**?

Alegeți o opțiune:

- ☐ a.
 <{a}, 1>; <{a,b}, 1>; <{a,b,c}, 1>; <{a,c}, 1>
- ☐ b.
 <{a}, NULL>; <{a,b}, NULL>; <{a,b,c}, NULL>; <{a,c}, NULL>
- ☐ c.
 <{a,b,c}, {a}>; <{a,b,c}, {a,b}>; <{a,b,c}, {a,b,c}>; <{a,b,c}, {a,c}>
- ☐ d.
 <{a}, suport+1>; <{a,b}, suport+1>; <{a,b,c}, suport+1>; <{a,c}, suport+1>

17 întrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Modelul **ACID** de reprezentare a datelor

Alegeți o opțiune:

- ☐ a. este adecvat pentru MapReduce deoarece presupune tranzacții atomice.
- ☐ b. este adecvat pentru MapReduce deoarece presupune procesarea izolată a datelor.
- ☐ c. nu este adecvat pentru MapReduce deoarece presupune procesări atomice de date.
- ☐ d. nu este adecvat pentru MapReduce deoarece poate condiționa sincronizările dintre etape.

18 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

O variabilă aleatoare este distribuită normal, cu media 0 și abaterea standard 1. Probabilitatea ca variabila să ia valoarea 2 este:

Alegeți o opțiune:

- ☐ a. 0.066
- ☐ b. 0.01
- ☐ c. 0.054
- ☐ d. 0.2

19 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Următoarea secvență: ABBAABAA este comprimată folosind compresia LZ. La un moment dat în decursul compresiei, dicționarul conține [1 A] [2 B] [3 BA] . Care este următorul element din dicționar?

Alegeți o opțiune:

- ☐ a. [4 ABA]
- ☐ b. [4 BAA]
- ☐ c. [4 AB]
- ☐ d. [4 BA]

20 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Versiunile secvențiale ale algoritmilor BFS/Dijkstra se bazează pe cozi centralizate de explorare. Ce se poate spune despre aceste structuri în cazul MapReduce?

Alegeți o opțiune:

- ☐ a. Sunt implementate la nivelul nodului coordonator și sincronizate cu nodurile worker.
- ☐ b. Nu sunt implementate deoarece nu sunt compatibile cu modelul MapReduce.
- ☐ c. Sunt implementate la nivelul nodurilor worker și sincronizate cu nodul coordonator.
- ☐ d. Nu sunt implementate deoarece implică un consum considerabil de memorie suplimentară.

21 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Considerând modelul MapReduce clasic, divizarea caracteristică etapei de **mapare** se realizează

Alegeți o opțiune:

- ☐ a. fără a se corela dimensiunea datelor de intrare cu numărul efectiv de workeri disponibili.
- ☐ b. astfel încât fiecare worker să primească același volum de lucru.
- ☐ c. corelând dimensiunea datelor de intrare cu numărul de calculatoare disponibile în cluster-ul de lucru.
- ☐ d. corelând dimensiunea datelor de intrare cu numărul efectiv de workeri disponibili.

22 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Care dintre următoarele afirmații este adevărată în ipoteza în care, în cadrul unei soluții MapReduce, **se defectează nodul cu rol de coordonator**?

Alegeți o opțiune:

- ☐ a. Se repornește un nou nod cu rol de coordonator, starea sistemului fiind reconstruită pe baza informațiilor primite de la nodurile cu rol de worker.
- ☐ b. Se repornește un nou nod cu rol de coordonator, starea sistemului fiind consistentă doar la nivel de worker.
- ☐ c. Se repornește un nou nod cu rol de coordonator, starea sistemului fiind reconstruită pe baza ultimelor informații de stare cunoscute.
- ☐ d. Sistemul nu poate fi restaurat și atunci se reinițializează întregul scenariu de procesare.

23 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Fazele MapReduce necesare și scopurile acestor faze pentru determinarea itemset-urilor frecvente dintr-o mulțime de tranzacții sunt următoarele:

Alegeți o opțiune:

☐ a.

1. **mapare**: determinarea candidaților locali, la nivel de partiție
2. **amestec și sortare/combiner**: determinarea suportului global
3. **reducere**: filtrarea itemset-urilor frecvente

☐ b.

1. **mapare**: determinarea candidaților locali, la nivel de partiție, și a suportului la nivel de partiție
2. **reducere**: determinarea suportului global și filtrarea itemset-urilor frecvente

☐ c.

1. **mapare**: determinarea candidaților locali, la nivel de partiție
2. **amestec și sortare/combiner**: determinarea suportului local, la nivel de partiție
3. **reducere**: determinarea suportului global și filtrarea itemset-urilor frecvente

☐ d.

1. **mapare**: determinarea candidaților locali, la nivel de partiție, și a suportului la nivel de partiție
2. **reducere**: filtrarea itemset-urilor frecvente

24 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

În cadrul modelului MapReduce, **echilibrarea încărcării** se realizează prin

Alegeți o opțiune:

- ☐ a. retrimiteră task-urilor întârziate către aceleași noduri care au întâzieri în procesare.
- ☐ b. reassignarea task-urilor întârziate către noduri libere; se acceptă primul set de rezultate primit de către coordonator.
- ☐ c. reassignarea task-urilor întârziate către noduri libere; setul de rezultate primit în astfel de cazuri este marcat special.
- ☐ d. retrimiteră task-urilor voluminoase către nodurile cu cele mai multe resurse de calcul libere.

25 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Pentru determinarea itemset-urilor frecvente pe baza unei aplicații MapReduce, noțiunea de **itemset candidat** se poate defini ca

Alegeți o opțiune:

- ☐ a. un super set al unor itemset-uri frecvente care sunt incluse în tranzații analizate în faza de mapare.
- ☐ b. orice subset, de orice dimensiune, al unei tranzații analizate în faza de mapare.
- ☐ c. orice subset, compus din cel puțin două elemente, al unei tranzații analizate în faza de mapare.
- ☐ d. orice subset al unei tranzații analizate în faza de mapare, subset compus exclusiv din itemi frecvenți.

26 ÎntrebareNu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Se dorește aplicarea modelului MapReduce **pentru problema determinării numărului de apariții ale cuvintelor întâlnite într-un text**. Care dintre variantele de mai jos reprezintă **ieșiri** valide pentru etapele de mapare și reducere (x reprezintă o valoare numerică întreagă, strict mai mare decât 1)?

Alegeți o opțiune:

- ☐ a. **mapare:** <cuvant, NULL>
reducere: <cuvant, x>
- ☐ b. **mapare:** <cuvant, x>
reducere: <cuvant, list(x)>
- ☐ c. **mapare:** <x, cuvant>
reducere: <cuvant, x>
- ☐ d. **mapare:** <cuvant, 1>
reducere: <cuvant, x>

27 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Fie un set de date ale căror elemente au două atribute A, B. Matricea de covarianță a setului de date este:

$\begin{bmatrix} 3 & -3 \\ -2 & 4 \end{bmatrix}$

Ce se poate spune despre relația dintre cele două atribute?

Alegeți o opțiune:

- ☐ a. Nimic, deoarece matricea de covarianță este greșită
- ☐ b. Dacă valorile atributului A cresc, tind să crească și valorile atributului B
- ☐ c. Dacă valorile atributului A scad, tind să scadă și valorile atributului B
- ☐ d. Valorile celor două atribute nu sunt corelate

28 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Fie următoarea mulțime de valori: 2 4 1 3 4 5 3 2 . Folosind estimarea probabilității maxime (Maximum Likelihood Estimation), valoarea estimată a mediei distribuției normale a valorilor este:

Alegeți o opțiune:

- ☐ a. 1
- ☐ b. 3
- ☐ c. 0.33
- ☐ d. 0.66

29 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Dacă se consideră o partiție formată din ***m tranzații***, fiecare tranzacție fiind compusă din ***n itemi***, atunci numărul maxim de itemset-uri candidat selectate va fi (caracterul $^$ înseamnă ridicare la putere):

Alegeți o opțiune:

- ☐ a. $m \cdot (2^n)$
- ☐ b. $2^{(n \cdot m)}$
- ☐ c. n^m
- ☐ d. $n \cdot (m^2)$

30 Întrebare

Nu a primit
răspuns încă

Marcat din 1,00

🚩 Întrebare cu
flag

Alegeți afirmația corectă privind prelucrarea fluxurilor:

Alegeți o opțiune:

- ☐ a. Datele sunt de dimensiune teoretic infinită
- ☐ b. Prelucrarea datelor din memoria de lucru necesită timpi îndelungați
- ☐ c. Procesarea datelor are loc după stocarea lor în memoria de lucru
- ☐ d. Memoria de lucru este în general de mari dimensiuni