

Started on	Wednesday, 24 November 2021, 2:03 PM
State	Finished
Completed on	Wednesday, 24 November 2021, 2:18 PM
Time taken	14 mins 26 secs
Marks	10.00/12.00
Grade	8.33 out of 10.00 (83%)

Question **1**

Correct

Mark 1.00 out of 1.00

Regularizarea regresiilor liniare și logistice presupune

Select one:

- ☒ a. Adăugarea unui termen în funcția de cost care penalizează ponderile cu valori mari
- ☐ b. Creșterea dimensiunii setului de antrenare prin adăugarea de noi exemple
- ☐ c. Selectarea acelor attribute ale datelor de intrare care sunt mai informative
- ☐ d. Transformarea spațiului de intrare într-unul mai complex

Răspunsul dumneavoastră este corect.

The correct answer is: Adăugarea unui termen în funcția de cost care penalizează ponderile cu valori mari



Question **2**

Correct

Mark 1.00 out of 1.00

Care dintre urmatoarele sunt bune practici pentru reducerea overfitting-ului?

- (a) Utilizarea unei funcții de cost cu 2 componente, care include un regulator pentru a penaliza complexitatea modelului
- (b) Utilizarea unui optimizator bun pentru a reduce erorile pe datele de antrenare
- (c) Construirea unei structuri de subseturi de modele imbricate, antrenarea pe fiecare subset pornind de la cel mai mic, si c de cross-validare incepe sa creasca
- (d) Eliminarea aleatoare a 50% din datele de antrenare

Select one:

- ☒ a. (a) si (c)
- ☐ b. (c)
- ☐ c. (a) si (b)
- ☐ d. (b) si (c)
- ☐ e. (a) si (b) si (c)

Răspunsul dumneavoastră este corect.

The correct answer is: (a) si (c)



Question **3**

Correct

Mark 1.00 out of 1.00

Avem urmatoarea matrice de confuzie pentru o problema de clasificare binara (pe coloane sunt valorile reale, iar pe linii sunt valorile predictibile).
Care afirmatie este adevarata?

	Positive	Negative
Positive	23	1
Negative	12	556

Select one:

- ☐ a. Accuracy=568/592
- ☐ b. Recall=23/24
- ☐ c. Accuracy=557/556
- ☒ d. Precision=23/24

Răspunsul dumneavoastră este corect.

The correct answer is: Precision=23/24

Question **4**

Correct

Mark 1.00 out of 1.00

Intr-un arbore de decizie pentru clasificarea unor exemple care contin attribute cu valori continue, are sens sa repetam pe o cale din arborelui (cale de la radacina la o frunza) un acelasi atribut.

Select one:

- ☒ True ✓
- ☐ False

The correct answer is 'True'.



Question **5**

Correct

Mark 1.00 out of 1.00

In cazul algoritmului C4.5 daca avem valori necunoscute pentru un atribut in anumite exemple de invatare trebuie sa

Select one:

- ☒ a. atribuim acelui atribut valoarea cu cea mai mare frecventa din exemplele de invatare
- ☐ b. atribuim acelui atribut valoarea cu cea mai mica frecventa din exemplele de invatare

Răspunsul dumneavoastră este corect.

The correct answer is: atribuim acelui atribut valoarea cu cea mai mare frecventa din exemplele de invatare



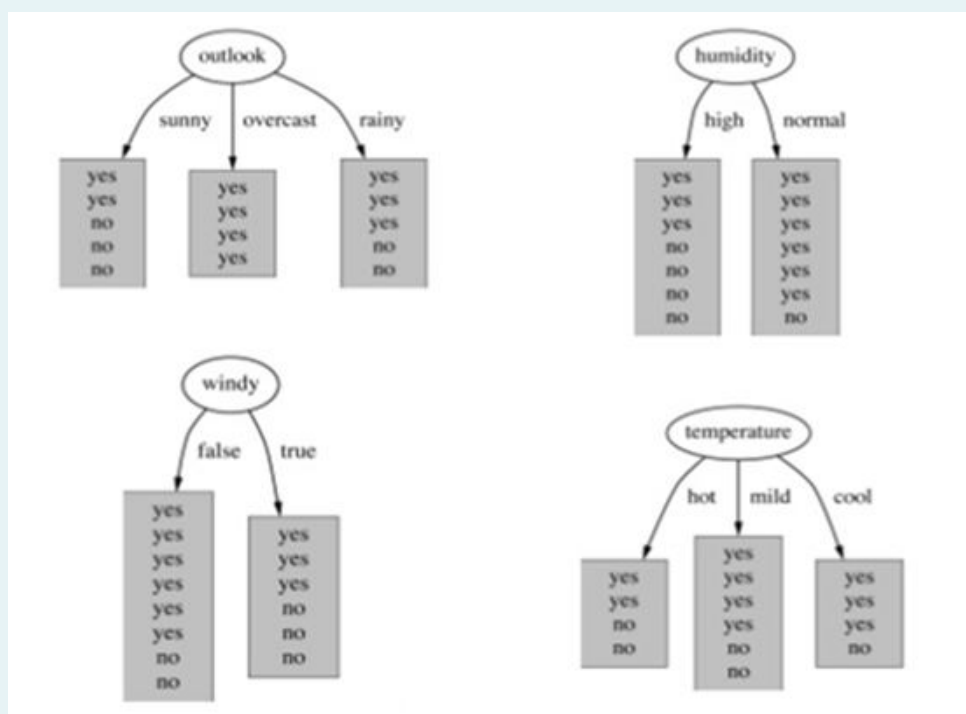
Question **6**

Correct

Mark 1.00 out of 1.00

Se antrenează un arbore de decizie având "maximum information gain" (câștigul maxim informațional, bazat pe diferența criteriu de selecție a atributelor).

Considerând acest criteriu și următoarea situație de antrenare (cf. imaginii), care va fi atributul selectat pentru a "împărți" r



Select one:

- ☒ a. Outlook
- ☐ b. Humidity
- ☐ c. Temperature
- ☐ d. Windy

Răspunsul dumneavoastră este corect.

The correct answer is: Outlook

Question **7**

Correct

Mark 1.00 out of 1.00

Care dintre următoarele reprezintă criterii de oprire pentru construcția arborelui de decizie cu algoritmul ID3. Alegeți opțiunile corecte:

1. Toate exemplele din subset fac parte din aceeași clasă
2. Câștigul informațional ajunge să fie negativ
3. Nu mai există atribute valide cu care să creăm un nod test
4. Adâncimea maximă este atinsă
5. Câștigul informațional depășește un anumit prag

Select one:

- ☒ a. 1, 3 și 4
- ☐ b. 4 și 5
- ☐ c. 1 și 4
- ☐ d. 1, 2 și 3
- ☐ e. 1

Răspunsul dumneavoastră este corect.

The correct answer is: 1, 3 și 4

Question **8**

Correct

Mark 1.00 out of 1.00

Ce tip de regresie ar trebui folosită pentru o problemă în care dorim să împărțim un set de date în 2 categorii? (presupunem că antrenăm un algoritm pentru care avem setul de date de antrenare)

Select one:

- ☐ a. Regresie liniară deoarece datele pot fi separate cu ajutorul unei drepte
- ☒ b. Regresia logistică deoarece folosește funcția logistică pentru a modela probabilitatea ca un punct din setul de date să aparțină într-o anumită clasă.
- ☐ c. Regresie liniară deoarece cele 2 categorii (din setul de antrenare) pot fi etichetate cu 0 și 1. În etapa de antrenare se găsește o dreaptă care trece prin cele 2 puncte (0 și 1)
- ☐ d. Regresie logistică deoarece nu știm cum să codificăm "categoriile" în nume reale necesare regresiei liniare.

Răspunsul dumneavoastră este corect.

The correct answer is: Regresia logistică deoarece folosește funcția logistică pentru a modela probabilitatea ca un punct din setul de date să aparțină într-o anumită clasă.



Question **9**

Correct

Mark 1.00 out of 1.00

Un set de date utilizat pentru învățare supervizată are o valoare mare a entropiei informaționale dacă:

Select one:

- ☐ a. Foarte multe dintre exemple aparțin unui număr mic de clase, în timp ce restul de clase au un număr mic de exemple
- ☒ b. Numărul de exemple din fiecare clasă este relativ similar
- ☐ c. Toate exemplele fac parte din aceeași clasă

Răspunsul dumneavoastră este corect.

The correct answer is: Numărul de exemple din fiecare clasă este relativ similar

Question **10**

Incorrect

Mark 0.00 out of 1.00

Arborii de decizie construiți cu algoritmul ID3 prezintă robustețe crescută la „outliers”? Care dintre răspunsuri este cel mai corect?

Select one:

- ☒ a. Da, întrucât acest lucru este garantat prin calculul entropiei.
- ☐ b. Nu, pentru că vom obține valori negative pentru câștigul informațional în cazul lor.
- ☐ c. Nu, pentru că ele (outliers) influențează sub-arborii generați ca fii ai unui nod test.
- ☐ d. Da, pentru că unele atribute nu sunt luate în considerare.

Răspunsul dumneavoastră este incorrect.

The correct answer is: Nu, pentru că ele (outliers) influențează sub-arborii generați ca fii ai unui nod test.



Question **11**

Incorrect

Mark 0.00 out of 1.00

Dorim să prezicem prețul unui apartament din zona Pipera. Pentru a face acest lucru ce putem folosi dintre:

Select one:

- ☐ a. Arbori de decizie
- ☐ b. Regresie logistică
- ☐ c. Regresie liniară
- ☐ d. Regresie liniară sau logistică
- ☒ e. Regresie liniară sau arbori de decizie

Răspunsul dumneavoastră este incorect.

The correct answer is: Regresie liniară

Question **12**

Correct

Mark 1.00 out of 1.00

Ce presupune "lama lui Occam" în contextul arborilor de decizie?

Select one:

- ☐ a. Un arbore corect de decizie trebuie să cuprindă toate atributele din setul de date
- ☒ b. Dacă sunt mai mulți arbori de decizie corecți, se preferă cel mai simplu
- ☐ c. Nu se poate aplica principiul în contextul arborilor de decizie
- ☐ d. Nu există mai mulți arbori corecți pentru aceeași problemă

Răspunsul dumneavoastră este corect.

The correct answer is: Dacă sunt mai mulți arbori de decizie corecți, se preferă cel mai simplu

◀ [Lab 7 - Soluții Finale](#)

Jump to...

