

1. Care dintre aceste argumente este fals cu referire la metoda celor mai apropiati vecini?

- A. Suprafata de decizie este neliniara
- B. Calitatea rezultatelor creste atunci cand avem mai multe date de antrenare
- C. Poate fi aplicat numai pe probleme de clasificare
- D. Eroarea de clasificare pe antrenare creste odata cu k , dar suprafata de decizie devine mai neteda

2. Care este scufundarea asociata functiei nucleu liniare?

- A. $f(x) = x$
- B. $k(x,y) = \langle x,y \rangle$, unde \langle, \rangle denota produsul scalar
- C. Nu exista
- D. $f(x) = \sqrt{x}$

3. Care dintre acestea nu este o regula pentru algoritmul DTW (Dynamic Time Warping)?

- A. Ultimele elemente formeaza ultima pereche
- B. Putem sari peste elemente
- C. Monotonia (nu ne putem intoarce)
- D. Primele elemente formeaza prima pereche

4. In cazul in care intampinam overfitting intr-o problema de invatare, care dintre aceste optiuni nu vor rezolva problema?

- A. Adaugarea regularizarii
- B. Simplificarea modelului de invatare
- C. Marirea setului de date
- D. Micșorarea setului de date

5. Eroare de modelare provine din:

- A. Alegerea unui spatiu de ipoteze prea larg
- B. Alegerea unui algoritm de optimizare nepotrivit pentru model
- C. Alegerea unui model cu capacitate slaba de modelare
- D. Alegerea unui algoritm de optimizare nepotrivit pentru setul de date

6. Pentru a detecta contururile dintr-o imagine RGB, cel mai potrivit model ar fi:

- A. Un clasificator SVM
- B. O retea de tip feedforward
- C. Un model Naive Bayes
- D. O retea convolutionala

7. Regresia logistica este:

- A. O regresia liniara la care se adauga o componenta regularizare
- B. Echivalentul clasificatorului cu vectori suport, in forma primala
- C. Echivalentul clasificatorului cu vectori suport, in forma duala
- D. O metoda de clasificare a carei functie de pierdere are valori intre 0 si $+\infty$

8. Cati parametri invatabili va avea un perceptron daca dimensiunea vectorilor de intrare este de 30 de elemente?

- A. 29
- B. 31

- C. 1
- D. 30

9. Este necesar sa impartim datele in seturi separate de antrenare si testare?

- A. Da, pentru a putea ajunge la acuratete 100% pe setul de antrenare
- B. Da, pentru a putea calcula diferenta de performanta pe cele doua seturi de date
- C. Nu, pentru ca este mai bine sa avem mai multe date de antrenare
- D. Da, pentru a putea estima corect eroare de generalizare

10. Care este volumul de iesire al unui strat convolutional cu intrarea de dimensiune $15 \times 15 \times 3$, avand 16 filtre de dimensiune 3, stride 2, padding 2?

- A. $2 \times 2 \times 15$
- B. $3 \times 3 \times 15$
- C. $3 \times 3 \times 16$
- D. $9 \times 9 \times 16$

1. Pentru o matrice 2D in care pe linie avem exemple de antrenare, iar pe coloana avem trasaturile exemplilor, standardizarea datelor presupune:

- A. Calcul medie pe coloana
- B. Calcul varianta pe linie
- C. Calcul medie si varianta pe coloana
- D. Calcul medie si varianta pe linie

2. Care din urmatoarele nu este o functie de evaluare a performantei?

- A. MAE
- B. Acuratetea
- C. F1-score
- D. ReLU

3. Cu cat este egal PReLU(x) cand $\alpha=0.5$ si $x = [-2, 2, 1, 0]$?

- A. $[-1, 2, 1, 0]$
- B. $[0, 1, 0.5, 0]$
- C. $[-2, 1, 0.5, 0]$
- D. $[-1, 1, 0.5, 0]$

4. Cand este mai eficient sa folosim reprezentarea duala a datelor?

- A. Cand avem o problema de clasificare cu foarte multe clase (mai mult de doua)
- B. Cand numarul de trasaturi este mai mare decat numarul de exemple
- C. Cand numarul de trasaturi este mai mic decat numarul de exemple
- D. Cand avem o problema de clasificare binara (cu doua clase)

5. Care din urmatoarele modele este un model binar?

- A. Naive Bayes
- B. KNN
- C. SVM
- D. MAE

6. Cati clasificatori binari vor fi antrenati pentru a rezolva o problema cu 9 clase folosind metoda de combinare ONE-vs-ALL?

- A. 12
- B. 36
- C. 9
- D. 18

7. Functia kernel PQ provine din:

- A. Functia de corelatie Kendall Tau
- B. Calcularea anumitor perechi concordante si discordante
- C. Diferenta media intre histograme de valori
- D. Functia kernel liniara aplicata histogramei de valori

8. Daca un nod dintr-un graf computational reprezinta operatia $z = 2 * x * y$, intrarile sunt $x = -0.5$ si $y = 2$, iar gradientul $dL/dz = -5$, atunci gradientii in raport cu intrarile dL/dx si dL/dy sunt:

- A. $dL/dx = -40$ si $dL/dy = 10$
- B. $dL/dx = 10$ si $dL/dy = -5$
- C. $dL/dx = 40$ si $dL/dy = -5$
- D. $dL/dx = -20$ si $dL/dy = 5$

9. Cum se poate calcula eroare de generalizare?

- A. Nu se poate calcula
- B. Se poate calcula prin integrarea functiei obiectiv in fiecare punct din setul de date de testare
- C. Se poate calcula prin integrarea functiei obiectiv in fiecare punct din setul de date de antrenare
- D. Se poate calcula prin insumarea functiei obiectiv in fiecare punct din setul de date de antrenare

10. Ce este functia Softmax?

- A. Functie kernel
- B. Functie de masurare a performantei
- C. Functie de pierdere
- D. Functie de activare