IA - Învățare automată

8 iunie 2022

Examen – Varianta 4

Oficiu (0.5p)

1. (0.25p) Care este scufundarea asociată funcției nucleu Hellinger?
A. $f(x) = x$
B. $f(x) = \sqrt{x}$
C. $k(x,y) = \langle \sqrt{x}, \sqrt{y} \rangle$, unde $\langle \blacksquare, \blacksquare \rangle$ este produsul scalar
D. Nu există

2. (0.25p) Ce determină capacitatea de modelare neliniară a rețelelor neuronale?
A. Funcțiile de activare B. Ponderile de tip bias C. Matricile de ponderi D. Funcția de pierdere

3. (0.25p) Care este dimensiunea tensorului de ieșire al unui strat convoluțional cu 20 filtre de 3x3 aplicate la stride de 2 pe un tensor de dimensiune 65x65x40?
A. 32x32x20 B. 31x31x20 C. 16x16x40 D. 31x31x40

4. (0.25p) La ce se referă teorema de aproximare universală?
A. Orice model de învățare poate aproxima orice funcție
B. Parametrii unei rețele se pot aproxima din date
C. Hiperparametrii se pot aproxima pe validare
D. Rețelele neuronale pot aproxima orice funcție

5. (0.25p) Eroarea de modelare provine din alegerea:
A. unui spațiu de ipoteze prea restrâns
B. unui model cu capacitate prea mare de modelare
C. unui algoritm de optimizare nepotrivit
D. unui spațiu de ipoteze circular

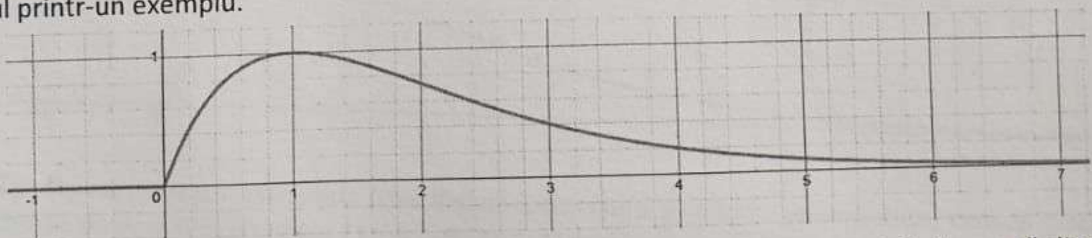
6. (0.25p) Care este numărul de parametri al unui strat convoluțional cu 40 de filtre de dimensiune 3x3 aplicate cu un stride de 3 pe un tensor de dimensiune 128x128x20?
A. 400 B. 360 C. 7240 D. 7220

7. (0.25p) Când ar putea să apară fenomenul Hughes?
A. Când algoritmul de optimizare nu converge
B. Când avem prea puține exemple
C. Când avem un model cu capacitate de modelare prea mică
D. Când avem prea puține trăsături

8. (0.25p) Dacă un nod dintr-un graf computațional reprezintă operația $z = 2xy^2$, intrările sunt $x = -0.5$ și $y = 2$, iar gradientul $\partial L/\partial z = -5$, atunci gradienții în raport cu intrările $\partial L/\partial x$ și $\partial L/\partial y$ sunt:
A. $\partial L/\partial x = 20$ și $\partial L/\partial y = -5$
B. $\partial L/\partial x = -20$ și $\partial L/\partial y = 5$
C. $\partial L/\partial x = -40$ și $\partial L/\partial y = 10$
D. $\partial L/\partial x = 40$ și $\partial L/\partial y = -10$

9. (1p) Considerând funcția kernel $k(x,y) = \sum_i \min\{x_i, y_i\}$, unde x și y sunt histograme cu frecvența de apariție a unor cuvinte, având frecvența maximă egală cu 5, definiți o funcție de scufundare ϕ, care aplicată lui x și y, produce: $k(x,y) = \sum_i \min\{x_i, y_i\} = \langle \phi(x), \phi(y) \rangle$. Exemplificați aplicarea funcției de scufundare pe $x = (4,0,3,2)$ și $y = (3,2,5,4)$, demonstrând egalitatea $\sum_i \min\{x_i, y_i\} = \langle \phi(x), \phi(y) \rangle$ pe acest exemplu.

10. (0.5p) Fiind dată funcția de activare $y = f(x) = max(0,x) \cdot exp(-x+1)$ cu graficul de mai jos, precizați dacă un singur neuron artificial bazat pe această funcție de activare ar putea rezolva problema XOR. Justificați răspunsul printr-un exemplu.



11. (1p) Configurați o rețea neuronală (specificând arhitectura și valorile parametrilor) care să discrimineze între mulțimile de puncte A (eticheta 1) și B (eticheta -1) reprezentate în figura de mai jos.