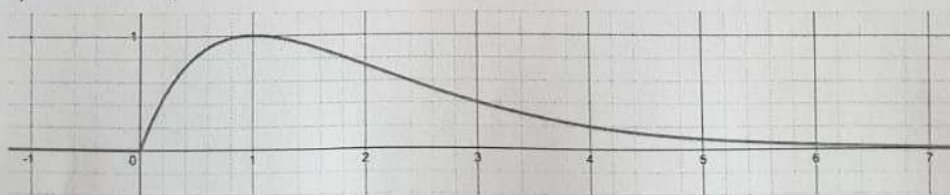


Oficiu (0.5p)

1. (0.25p) Cum se descompune eroarea unui model de învățare automată?
 A. Eroare de modelare, optimizare și testare ☒ B. Eroare de modelare, estimare și optimizare
 C. Eroare de modelare, învățare și testare D. Eroare de modelare, optimizare și generalizare
2. (0.25p) Când ar putea să apară fenomenul Hughes?
 A. Când avem prea multe exemple B. Când algoritmul de optimizare nu converge
☒ C. Când avem prea multe trăsături D. Când avem un model cu capacitate de modelare prea mică
3. (0.25p) Care este numărul de parametri al unui strat convoluțional cu 20 de filtre de dimensiune 3x3 aplicate cu un stride de 3 pe un tensor de dimensiune 128x128x40?
 A. 200 B. 180 C. 7240 ☒ D. 7220
4. (0.25p) Care este scufundarea asociată funcției nucleu RBF?
☒ A. Nu există B. $f(x) = x$
 C. $k(x, y) = e^{-\frac{\Delta(x, y)}{2\sigma}}$, unde $\Delta(x, y)$ este distanța Euclidiană D. $f(x) = \sqrt{x}$
5. (0.25p) Ce determină capacitatea de modelare neliniară a rețelelor neuronale?
 A. Matricile de ponderi ☒ B. Funcțiile de transfer C. Funcția de pierdere D. Ponderile de tip bias
6. (0.25p) Dacă un nod dintr-un graf computațional reprezintă operația $z = 2xy$, intrările sunt $x = -0.5$ și $y = 2$, iar gradientul $\partial L / \partial z = 5$, atunci gradientii în raport cu intrările $\partial L / \partial x$ și $\partial L / \partial y$ sunt:
 A. $\partial L / \partial x = -40$ și $\partial L / \partial y = 10$ B. $\partial L / \partial x = 40$ și $\partial L / \partial y = -10$
☒ C. $\partial L / \partial x = 20$ și $\partial L / \partial y = -5$ D. $\partial L / \partial x = -20$ și $\partial L / \partial y = 5$
7. (0.25p) Care este dimensiunea tensorului de ieșire al unui strat convoluțional cu 40 filtre de 3x3 aplicate la stride de 2 pe un tensor de dimensiune 65x65x20?
 A. 16x16x20 B. 31x31x20 ☒ C. 32x32x40 D. 31x31x40
8. (0.25p) La ce se referă teorema „no free lunch”?
 A. Rețele neuronale sunt în general soluția optimă B. Lansarea procesului de antrenare este costisitoare
 C. Rețelele neuronale aproximează orice funcție ☒ D. Nu există un singur model bun pe toate task-urile
9. (1p) Considerând funcția kernel $k(x, y) = 2PQ$, unde $P = |\{(i, j): 1 \leq i < j \leq n, (x_i - x_j) \cdot (y_i - y_j) \geq 0\}|$ și $Q = |\{(i, j): 1 \leq i < j \leq n, (x_i - x_j) \cdot (y_i - y_j) < 0\}|$, definiți o funcție de scufundare ϕ , care aplicată lui x și y , produce: $k(x, y) = 2PQ = \langle \phi(x), \phi(y) \rangle$. Exemplificați aplicarea funcției de scufundare pe $x = (5, 2, 1, 0)$ și $y = (1, 3, 5, 1)$, demonstrând egalitatea $2PQ = \langle \phi(x), \phi(y) \rangle$ 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.

