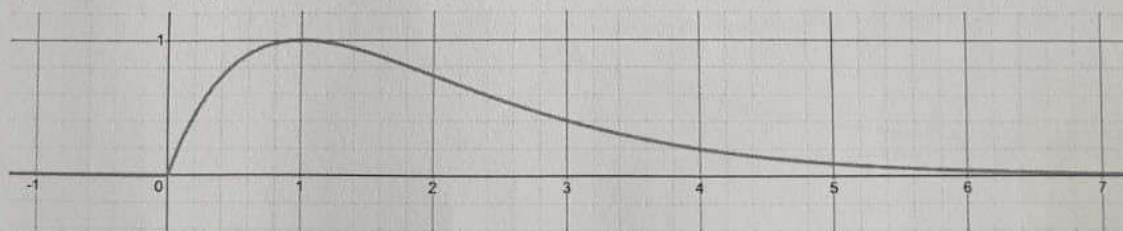


Oficiu (0.5p)

1. (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. $\partial L / \partial x = -20$ și $\partial L / \partial y = 5$ 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. $\partial L / \partial x = 40$ și $\partial L / \partial y = -10$
2. (0.25p) Care este scufundarea asociată funcției nucleu RBF?
 A. $f(x) = x$ B. $f(x) = \sqrt{x}$
 C. $k(x, y) = e^{-\frac{\Delta(x, y)}{2\sigma}}$, unde $\Delta(x, y)$ este distanța Euclidiană ☒ D. Nu există
3. (0.25p) Cum se descompune eroarea unui model de învățare automată?
☒ A. Eroare de modelare, estimare și optimizare B. Eroare de modelare, învățare și testare
 C. Eroare de modelare, optimizare și generalizare D. Eroare de modelare, optimizare și testare
4. (0.25p) La ce se referă teorema „no free lunch”?
 A. Lansarea procesului de antrenare este costisitoare B. Rețelele neuronale aproximează orice funcție
☒ C. Nu există un singur model bun pe toate task-urile D. Rețele neuronale sunt în general soluția optimă
5. (0.25p) Când ar putea să apară fenomenul Hughes?
 A. Când avem prea multe exemple ☒ B. Când avem prea multe trăsături
 C. Când algoritmul de optimizare nu converge D. Când avem un model cu capacitate de modelare prea mică
6. (0.25p) Care este numărul de parametri al unui strat convoluțional cu 30 de filtre de dimensiune 5x5 aplicate cu un stride de 2 pe un tensor de dimensiune 64x64x10?
 A. 7510 ☒ B. 7530 C. 750 D. 780
7. (0.25p) Ce determină capacitatea de modelare neliniară a rețelelor neuronale?
 A. Matricile de ponderi B. Funcția de pierdere ☒ C. Funcțiile de transfer D. Ponderile de tip bias
8. (0.25p) Care este dimensiunea tensorului de ieșire al unui strat de pooling cu filtre de 2x2 aplicate la stride de 4 pe un tensor de dimensiune 66x66x40?
 A. 16x16x40 B. 31x31x40 ☒ C. 17x17x40 D. 32x32x40
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 $x = (0, 4, 1, 2)$ 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.

