

1. - Vor 4  
 a) Ce se poate fi aplicată pt a transforma dist  $L_2$  în fc de nucleu?  
 kernel intersecție; kernel polinomial; kernel RBF; kernel liniar
2. Curolatza Kendall  $\tau$  pt etichetele  $[0,5; 0,4; 0,3; 0,2; 0,1]$  și predicțiile  $[0,8; 0,6; 1,0; 0,2; 0,4]$  este:  $0,6; 1,5; 0,5; 0,4$
3. Atunci când m-șor, eroare empirică derivă echivalentă cu: eroarea de modelare; eroarea de optimizare; micșura din variante; eroarea de generalizare
4. Cum ar putea fi descrisă formula:  $\max(0, \frac{e^x - e^{-x}}{e^x + e^{-x}})$ ?  
 $\sigma(\tanh(x))$ ,  $\text{ReLU}(\tanh(x))$ ,  $\text{ReLU}(\sigma(x))$ ,  $\tanh(\text{ReLU}(x))$
5. Prez. apl. fc. string kernel bazat pe left de prezență folosind 2- grame pe stivă "ana are mere" și "dan are banani" este:  $10; 7; 3; 1$
6. Care este efectul aplicării operației de normalizare batch pe un batch format dintr-un exemplu? norma exemplului derivă 1; norma exemplului derivă 0; exemplul rămâne nemodificat; nu se poate aplica pt că implică împărțirea la 0.
7. Ce este condiția de independență a rețelei neuronale pt a înțelege & generaliza de separare: minim 2 straturi ascunse; într-un strat rețetă mare, continuă într-un strat ascuns cu activare liniară, să conțină al puțin un strat ascuns.
8. Care din urm. exp. produce o fc kernel validă? minimă;  $k_1(k_2(x,z), k_2(x,z))$ ,  $k_1(x) + k_2(y)$ ,  $k_1(x,y) + k_2(x,y)$ ,  $k_3(x,y)$
9. Ce se întâmplă dacă setăm momentum egal cu 1 în alg SGD cu momentum? Alg devine cu SGD standard; Alg diverge; Alg se blochează într-un minim local; optimizarea merge la 0
10. Care din urm variabile nu permite să conținem hiperplanul de margine maximă în cazul unei conf. de puncte liniar separabile, ignorând arundele pet?  
 $\gamma$ ,  $\xi$ ,  $w$ , nu este cazul.
11.  $(0,7)$  pt # multime de exemple necorespunzătoare  $S$  din  $\mathbb{R}^2$ : a) numită și derivată de geom. pe care se proiectează pt. din  $S$  după normalizarea datelor folosind norma  $L_2$ .  
 b) Se poate construi un neuron care să repore pt normalizate de din  $\mathbb{R}^2$  cu  $x > 0$  de toate celelalte? 12(1,3 pt) Fiind dată multimea de exemple  $S = \{(1,1,0,1), (1,1,1,0), (1,1,0,0), (1,1,1,1)\}$ : a) apl neuronul cu ponderile  $w = [-1, 0, 0]$  și  $b = 0,1$  aruncă la activare  $\text{ReLU}$  pe datele din  $S$ . b) Ce procl apare în cazul optimizării acestui neuron? Derivatele nu sunt

val pt le. a. și să rez. problema intermilitară. Din că marea val rez într-aderat prosl  
~) Se poate conștii un neuron cu fe. de activare masant care să presă coact  
(cu erare o) etichetele exemplilor din s? justifică resp.