A

Zadanie	1	2	3	4	5	6	7	8	9	10	$\sum$
Punkty (maks)	(4)	(4)	(8)	(10)	(8)	(10)	(10)	(8)	(8)	(10)	(80)

Uwaga: część A zostanie wypełniona przez egzaminatora!

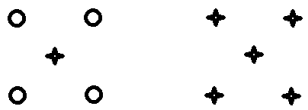
Nazwisko i Imię: Nr. indeksu:

Systemy Uczące Się – SUS 2017-2018

— Warszawa, 26/06/2018. Czas pisania pracy: 90 minut —

1. [4 punkty] Podaj wymiar $VCdim$ dla klasy klasyfikatorów $k-NN$ dla $k=1$. Odpowiedź krótko uzasadnij.

2. [4 punkty] Chcemy oszacować błąd klasyfikatorów typu kNN za pomocą algorytmu walidacji Leave-One-Out (1-CV). Używamy do tego celu następującego zbioru danych w $\mathbb{R}^2$ (z metryką euklidesową).

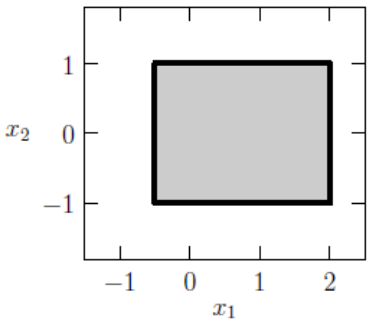


Wybierz T (tak) lub N (nie) przy następujących stwierdzeniach:

- ☐ a) 1NN jest lepszym klasykatorem niż 3NN;
- ☐ b) 1NN nie jest gorszy niż 9NN;
- ☐ c) Na tych danych nie istnieje gorszy algorytm niż 1NN;
- ☐ d) Na tych danych nie ma lepszego klasyfikatora typu kNN niż 9NN ;

3. [8 punktów]

Narysuj dwu-warstwową sieć neuronową z jedną wartwową ukrytą i z jednym neuronem w warstwie wyjściowej. Wszystkie neurony są aktywowane funkcjami progowymi, a neuron wyjściowy powinien zwracać wartość 1 wtedy i tylko wtedy, gdy punkt na wjeściu (x_1, x_2) znajduje się w zaciemnionym obszarze.
Zaznacz na rysunku dokładnie wartości progowe neuronów oraz wag krawędzi.



4. [10 punktów] Rozpatrujemy algorytm 3-centroidów (k-means dla $k = 3$) dla zbioru S zawierającego 6 punktów $a = (0, 0), b = (8, 0), c = (16, 0), d = (0, 6), e = (8, 6), f = (16, 6)$. W każdej iteracji stosujemy metryką euklidesową, aby przypisać obiekty do najbliższego centroida, a w przypadku nietrzygnięcia obiekt jest przypisany do lewego lub dolnego centroida. Konfiguracja początkowa składa się z 3 punktów ze zbioru S.

   (a) Ile jest możliwych konfiguracji początkowych składających z 3 punktów ze zbioru S?

   (b) Wymień wszystkie stabile podziały zbiory S na 3 zbiory

   (c) Dla każdego podziału stabilnego wyznacz liczbę konfiguracji początkowych, z których algorytm 3-centroidów prowadzi do tego podziału?

   (d) Jaka jest maksymalna liczba iteracji w algorytmie k-centroidów prowadzących konfigurację początkową do jednego ze stabilnych podziałów?

5. [8 punktów] Pogrupowano pewien zbiór danych algorytmem *k-means*. Centra grup po 3 iteracjach algorytmu to $p_1, \ldots, p_k$, zaś centra po 7 iteracjach to $q_1, \ldots, q_k$. Pokaż, że

$$Conv(p_1, \ldots, p_k) \cap Conv(q_1, \ldots, q_k) \neq \emptyset$$

   gdzie $Conv(F)$ to otoczka wypukłą zbioru punktów F.

6. [10 punktów] Rozpatrzmy następującą tablicę decyzyjną z 7 atrybutów i 8 obiektów. Obiekty o numerach 31 i 32 są przykładami testowym o nieznanych wartociach decyzyjnych.

LP.	e1	e2	e3	e4	e5	e6	e7	dec
11	1	1	1	0	1	1	1	T
12	1	0	1	1	1	0	1	T
13	0	1	1	1	0	1	0	T
14	1	1	0	1	1	1	1	T
21	1	0	1	1	0	1	1	N
22	1	1	0	1	0	1	1	N
23	1	0	1	0	0	1	1	N
24	1	1	1	1	0	1	1	N
31	0	0	1	0	0	1	0	?
32	1	1	1	1	1	1	1	?

(a) Chcemy skonstruować binarne drzewo decyzyjne dla tej tablicy. Jeśli używasz miary Entropy lub Discernibility (rozróżnialności) do wyznaczania najlepszego testu podczas konstrukcji drzewa, to który atrybut powinien być wybrany? Odpowiedź uzasadnij.

(b) Czy wszystkie atrybuty są potrzebne do konstrukcji drzewa decyzyjnego dla tej tablicy? Odpowiedź uzasadnij.

(c) Klasyfikuj obiekty nr 31 i 32 metodą Naive-Bayes. Przedstaw najważniejsze kroki obliczeń.

7. [10 punktów] Klasyfikator postaci $L_{a,b}(x,y) = 1 \iff ax + by > 0$ nazywamy *dwuwymiarowym klasyfikatorem liniowym*. Niech $\mathcal{F}_{lin} = \{L_{a,b} : a, b \in \mathbb{R}\}$ oznacza przestrzeń wszystkich dwuwymiarowych liniowych klasyfikatorów.

   (a) Wyznacz $VCdim(\mathcal{F}_{lin})$.

   (b) Wyznacz $m_{\mathcal{F}_{lin}}(3)$.

   (c) Niech f_1 i f_2 będą ustalonymi klasyfikatorami ($f_1, f_2 \notin \mathcal{F}_{lin}$). Wykazać, że $VCdim(\mathcal{F}_{lin} \cup \{f_1, f_2\}) \leqslant 4$

8. [8 punktów]

   Rozpatrzmy ukryty łańcuch Markowa (π, A, B) nad przestrzenią stanów $\{E_1, E_2, E_3\}$ i zbiorem symboli $\{S_1, S_2\}$ zdefiniowany jak następuje:

$$\pi = (0, 0, 1)$$

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 2/9 & 4/9 & 3/9 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}$$

   Jakie jest prawdopodobieństwo emisji symbolu S_1 po dwustu krokach?

9. [8 punktów]

   (a) Niech $K_1 : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ będzie dowolną funkcją jądrową. Udowodnij, że funkcja

   $$K(\mathbf{x}, \mathbf{y}) = a \cdot K_1^2(\mathbf{x}, \mathbf{y}) + b$$

   gdzie a, b są dodatnimi liczbami rzeczywistymi, też jest funkcją jądrową.

   (b) Dla funkcji jądrowej $K(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle^2 \cdot (4\langle \mathbf{x}, \mathbf{y} \rangle + 1)$, gdzie $\mathbf{x}, \mathbf{y}$ są wektorami w przestrzeni dwuwymiarowej, znaleźć wartość k oraz odpowiednie zanurzenie $\Phi : \mathbb{R}^2 \to \mathbb{R}^k$ tak, aby

   $$\langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle = K(\mathbf{x}, \mathbf{y})$$

10. [10 punktów] Przypomnijmy, że dla danej macierzy błędu (confusion matrix) zawierające TP, FP, TN i FN,

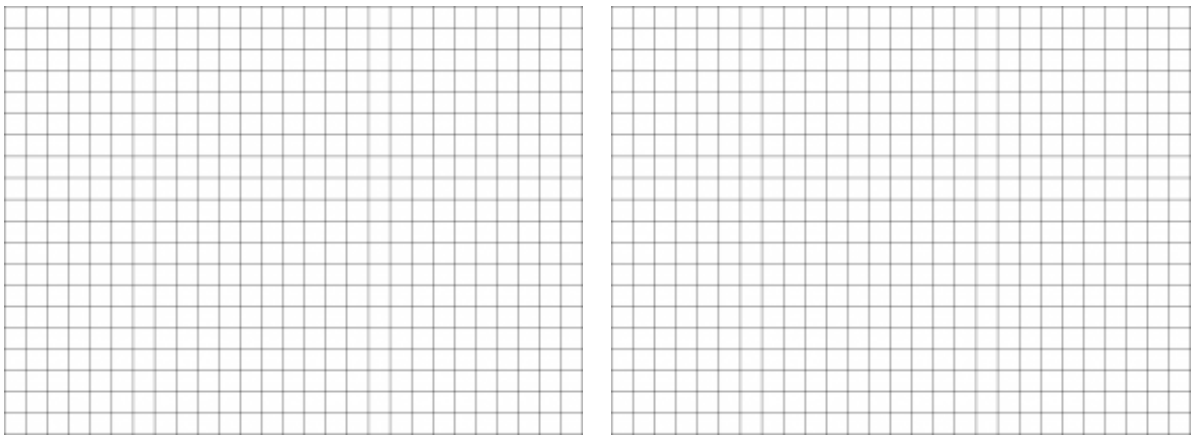
		Predykcja	
		1	0
Oryginalna	1	TP	FN
	0	FP	TN

możemy zdefiniować dodatkowe miary skuteczności, takie jak $Precision = \dfrac{TP}{TP + FP}$, $Recall = \dfrac{TP}{TP + FN}$ i $TPR = Recall$, $FPR = \dfrac{FP}{TN + FP}$. Krzywa ROC jest parametryzowaną krzywą z TPR na osi y oraz FPR na osi x.

Pewien algorytm podejmowana decyzji produkuje hipotezę, która uporządkuje przykłady w zbiorze testowym na liście rankingowej od najwyższej (na lewej stronie) do najniższej (na prawej stronie). Znaki + i − oznaczają klasę decyzyjną tych obiektów:

$$+ + + - + + - + + + - + - - - - + - --$$

Narysuj krzywą ROC oraz krzywą Precision-Recall dla tego zbioru przykładów.



— BRUDNOPIS —