National Technical University of Athens

School of Electrical and Computer Engineering

Artificial Intelligence

Semester 7 - Flow L

Analytic Assignment 3
Machine Learning & Uncertainty

Dimitris Dimos - 031 17 165

Athens
January, 2021

1 Άσκηση 1: Perceptron

1.1 ερώτημα (α)

Εκτελούμε τον αλγόριθμο εκμάθησης του *perceptron*, διαθέτοντας τα εξής διανύσματα δεδομένων:

$$\begin{aligned}
\mathbf{x_1} &= \big[0,1,3\big]^T \in B \\
\mathbf{x_2} &= \big[3,0,-1\big]^T \in A \\
\mathbf{x_3} &= \big[1,2,0\big]^T \in B \\
\mathbf{x_4} &= \big[3,-1,0\big]^T \in A \\
\mathbf{x_5} &= \big[-2,1,-2\big]^T \in B \\
\mathbf{x_6} &= \big[0,-2,-1\big]^T \in A
\end{aligned}$$

Αρχικό διάνυσμα βαρών: $\mathbf{w}[0] = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$, βήμα μάθησης $\beta = 0.2$ και συνάρτηση ταξινόμησης: $f(x) = \begin{cases} 1 & ,x \geq 0 \\ 0 & ,x < 0 \end{cases}$. Έξοδος 0 $\longrightarrow$ κλάση A, έξοδος 1 $\longrightarrow$ κλάση B.

<u>Επανάληψη 1</u>

$$\mathbf{w^T}[0]\mathbf{x_1} = \begin{bmatrix} w_0 & w_1 & w_2 & w_3 \end{bmatrix} . \begin{bmatrix} 1 \\ 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix} . \begin{bmatrix} 1 \\ 0 \\ 1 \\ 3 \end{bmatrix} = -3 < 0$$

Άρα, το $\mathbf{x_1}$ ταξινομείται λανθασμένα στην κλάση A.

$$\mathbf{w}[1] = \mathbf{w}[0] + \beta\big(y_1 - f(\mathbf{x_1})\big)\mathbf{x_1} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} + 0.2\big(1-0\big)\begin{bmatrix} 1 \\ 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1.2 \\ 1 \\ -0.8 \\ -0.4 \end{bmatrix}$$

<u>Επανάληψη 2</u>

$$\mathbf{w^T}[1]\mathbf{x_2} = \begin{bmatrix} 1.2 & 1 & -0.8 & -0.4 \end{bmatrix} . \begin{bmatrix} 1 \\ 3 \\ 0 \\ -1 \end{bmatrix} = 4.6 > 0$$

Άρα, το $\mathbf{x_2}$ ταξινομείται λανθασμένα στην κλάση B.

$$\mathbf{w}[2] = \mathbf{w}[1] + \beta\big(y_2 - f(\mathbf{x_2})\big)\mathbf{x_2} = \begin{bmatrix} 1.2 \\ 1 \\ -0.8 \\ -0.4 \end{bmatrix} + 0.2\big(0-1\big) \begin{bmatrix} 1 \\ 3 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.4 \\ -0.8 \\ -0.2 \end{bmatrix}$$

Επανάληψη 3

$$\mathbf{w^T}[2]\mathbf{x_3} = \begin{bmatrix} 1 & 0.4 & -0.8 & -0.2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix} = -0.2 < 0$$

Άρα, το $\mathbf{x_3}$ ταξινομείται λανθασμένα στην κλάση A.

$$\mathbf{w}[3] = \mathbf{w}[2] + 0.2\big(1-0\big)\mathbf{x_3} = \begin{bmatrix} 1 \\ 0.4 \\ -0.8 \\ -0.2 \end{bmatrix} + 0.2\big(1-0\big) \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.2 \\ 0.6 \\ -0.4 \\ -0.2 \end{bmatrix}$$

Επανάληψη 4

$$\mathbf{w^T}[3]\mathbf{x_4} = \begin{bmatrix} 1.2 & 0.6 & -0.4 & -0.2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -1 \\ 0 \end{bmatrix} = 3.4 > 0$$

Άρα, το $\mathbf{x_4}$ ταξινομείται λανθασμένα στην κλάση B.

$$\mathbf{w}[4] = \mathbf{w}[3] + 0.2\big(0-1\big)\mathbf{x_4} = \begin{bmatrix} 1.2 \\ 0.6 \\ -0.4 \\ -0.2 \end{bmatrix} + 0.2\big(0-1\big) \begin{bmatrix} 1 \\ 3 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -0.2 \\ -0.2 \end{bmatrix}$$

Επανάληψη 5

$$\mathbf{w^T}[4]\mathbf{x_5} = \begin{bmatrix} 1 & 0 & -0.2 & -0.2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 1 \\ -2 \end{bmatrix} = 1.2 > 0$$

Άρα, το $\mathbf{x_5}$ ταξινομείται ορθά στην κλάση B. Άρα, $\mathbf{w}[5] = \mathbf{w}[4]$.

Επανάληψη 6

$$\mathbf{w^T}[5]\mathbf{x_6} = \begin{bmatrix} 1 & 0 & -0.2 & -0.2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -2 \\ -1 \end{bmatrix} = 1.6 > 0$$

Άρα, το $\mathbf{x_6}$ ταξινομείται λανθασμένα στην κλάση B.

$$\mathbf{w}[6] = \mathbf{w}[5] + 0.2(0-1)\mathbf{x_6} = \begin{bmatrix} 1 \\ 0 \\ -0.2 \\ -0.2 \end{bmatrix} + 0.2(0-1)\begin{bmatrix} 1 \\ 0 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0 \\ 0.2 \\ 0 \end{bmatrix}$$

Αυτό είναι το πέρας της πρώτης εποχής. Συνεχίζουμε στις επόμενες, με πιο συνοπτική εκτέλεση του αλγορίθμου εκμάθησης.

2η Εποχή

x_k	y_k	$f(w^Tx)$	$y_k - f(x_k)$	*update*	
7	$[1,0,1,3]$	$f(1)=1$	$1-1=0$	0	$[0.8,0,0.2,0]$
8	$[1,3,0,-1]$	$f(0.8)=1$	$0-1=-1$	$[-0.2,-0.6,0,0.2]$	$[0.6,-0.6,0.2,0.2]$
9	$[1,1,2,0]$	$f(0.4)=1$	$1-1=0$	0	$[0.6,-0.6,0.2,0.2]$
10	$[1,3,-1,0]$	$f(-1.4)=0$	$0-0=0$	0	$[0.6,-0.6,0.2,0.2]$
11	$[1,-2,1,-2]$	$f(1.6)=1$	$1-1=0$	0	$[0.6,-0.6,0.2,0.2]$
12	$[1,0,-2,-1]$	$f(0)=1$	$0-1=-1$	$[-0.2,0,0.4,0.2]$	$[0.4,-0.6,0.6,0.4]$

3η Εποχή

x_k	y_k	$f(w^Tx)$	$y_k - f(x_k)$	*update*	
13	$[1,0,1,3]$	$f(2.2)=1$	$1-1=0$	0	$[0.4,-0.6,0.6,0.4]$
14	$[1,3,0,-1]$	$f(-1.8)=0$	$0-0=0$	0	$[0.4,-0.6,0.6,0.4]$
15	$[1,1,2,0]$	$f(1)=1$	$1-1=0$	0	$[0.4,-0.6,0.6,0.4]$
16	$[1,3,-1,0]$	$f(-1)=0$	$0-0=0$	0	$[0.4,-0.6,0.6,0.4]$
17	$[1,-2,1,-2]$	$f(0.4)=1$	$1-1=0$	0	$[0.4,-0.6,0.6,0.4]$
18	$[1,0,-2,-1]$	$f(-1.2)=0$	$0-0=0$	0	$[0.4,-0.6,0.6,0.4]$

$$\boxed{w = \begin{bmatrix} 0.4 \\ -0.6 \\ 0.6 \\ 0.4 \end{bmatrix}}$$

1.2 ερώτημα (β)

Έχουμε το διάμνυσμα [3,-1,3].

$$w^Tx = [0.4,-0.6,0.6,0.4] \cdot [1,3,-1,3]^T = -0.8$$

Επομένως, ταξινομείται στην κλάση A.

2 Άσκηση 2: Ταξινομητές Πλησιέστερων Γειτόνων

Ευκλείδια απόσταση: $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$.

$$d_1 = \sqrt{(3-0)^2 + (-1-1)^2 + (3-3)^2} = \sqrt{13} \approx 3.6$$

$$d_2 = \sqrt{(3-3)^2 + (-1-0)^2 + (3+1)^2} = \sqrt{17} \approx 4.12$$

$$d_3 = \sqrt{(3-1)^2 + (-1-2)^2 + (3-0)^2} = \sqrt{22} \approx 4.7$$

$$d_4 = \sqrt{(3-3)^2 + (-1+1)^2 + (3-0)^2} = 3$$

$$d_5 = \sqrt{(3+2)^2 + (-1-1)^2 + (3+2)^2} = \sqrt{54} \approx 7.35$$

$$d_6 = \sqrt{(3-0)^2 + (-1+2)^2 + (3+1)^2} = \sqrt{26} \approx 5.1$$

Επομένως, ένας ταξινομητής πλησιέστερου γείτονα θα ταξινομούσε το δεδομένο εισόδου στην κλάση στην οποία ανήκει το δεδομένο με το οποίο έχει τη μικρότερη απόσταση, δηλαδή αυτή του x_4. Άρα, ταξινομείται στην A.

Ένας ταξινομητής 3 πλησιέστερων γειτόνων θα ταξινομούσε το δεδομένο εισόδου στην κλάση στην οποία ανήκει η πλειοψηφία των 3 πλησιέστερων γειτόνων. Δηλαδή την A.

3 Άσκηση 3: Ασαφής Λογική

$$p : \text{Αν η } \mathcal{X} \text{ είναι } A_1 \text{ και η } \mathcal{Y} \text{ είναι } A_2\text{, τότε η } \mathcal{Z} \text{ είναι } B$$

Ο ασαφής κανόνας p αποτελείται από μια συνθέτη ασαφή πρόταση (δηλαδή ασαφή σχέση) και μια ασαφή ατομική πρόταση που συνδέονται μέσω ασαφούς συνεπαγωγής.

Η σύνθεση πρόταση ερμηνεύεται ως:

$$\text{Το } \langle A_1, A_2 \rangle \text{ ειναι } R_1\text{, όπου } R_1(x,y) = i(A_1(x), A_2(y))$$

Χρησιμοποιώντας ως τομή, τον συνήθη τελεστή τομής min, η ασαφής σχεση παράγει το σύνολο:

$$A_3 = 0.4/x_1, y_1 + 0.6/x_1, y_2 + 0.4/x_2, y_1 + 0.8/x_2, y_2 + 0.4/x_3, y_1 + 1/x_3, y_2$$

Η ασαφής συνεπαγωγή ερμηνεύεται ως:

$$\text{Το } \langle A_3, B \rangle \text{ ειναι } R_2\text{, όπου } R_2(x,z) = J(A_3(x), B(z))$$

Χρησιμοποιώντας συνεπαγωγή $Mamdani$, δηλαδή $J(a,b) = min\{a,b\}$, εν τελεί η συνολική πρόταση παράγει το ασαφές σύνολο:

$$\begin{aligned}
J(A_3, B) &= 0.4/x_1, y_1, z_1 + 0.6/x_1, y_2, z_1 + 0.4/x_2, y_1, z_1 \\
&+ 0.8/x_2, y_2, z_1 + 0.4/x_3, y_1, z_1 + 1/x_3, y_2, z_1 \\
&+ 0.4/x_1, y_1, z_2 + 0.5/x_1, y_2, z_2 + 0.4/x_2, y_1, z_2 \\
&+ 0.5/x_2, y_2, z_2 + 0.4/x_3, y_1, z_2 + 0.5/x_3, y_2, z_2
\end{aligned}$$

Θεωρώντας ότι ο ασαφής κανόνας αποτελεί ασαφές σύστημα, βλέπουμε ότι αν η τιμή της εισόδου είναι x_2 για τη μεταβλητή $\mathcal{X}$ και y_1 για τη μεταβλητή $\mathcal{Y}$, τότε το σύστημα, παράγει στην έξοδο του το ασαφές σύνολο:

$$\boxed{T(p) = 0.4/z_1 + 0.4/z_2}$$