

חלק 2 [75 נקודות]

- יש להציג את כל התוצאות עם **שלוש ספרות אחרי נקודה עשרונית** אלא אם צוין אחרת!
- יש לרשום את כל התשובות **על-גבי שאלון הבחינה בלבד**
- טיטות החישוב ייגרסו **ללא בדיקה**

כל השאלות בחלק הזה מבוססות על בסיס הנתונים ICU הנמצא במאגר UCI Machine Learning Repository. הנתונים מתייחסים לתינוקות בת 8.5 חודשים שאושפזה ביחידה לטיפול נמרץ (Intensive Care Unit) בעקבות זיהום חמור שלא הגיב לטיפול אנטיביוטי ופגע במערכת הנשימה. בסיס הנתונים כולל מספר פרמטרים שנמדדו בתדירות גבוהה במשך כ-12 שעות בהן החולה הייתה מחוברת למכונת הנשמה מלאכותית. תקופת הניטור חולקה על-ידנו לתשעה אינטרוולים כאשר בכל אינטרוול חושבו המשתנים הבאים:

- Average7 – ממוצע דופק במהלך האינטרוול (משתנה קלט)
- Slope7 – מגמת שינוי של דופק במהלך האינטרוול (משתנה קלט)
- Average19 – ממוצע לחץ דם במהלך האינטרוול (משתנה קלט)
- Slope19 – מגמת שינוי של לחץ דם במהלך האינטרוול (משתנה קלט)
- Balance – מאזן נוזלים בסיום האינטרוול (הוגדר על-ידנו כמשתנה המטרה)

להלן ערכי המשתנים שחושבו בכל אינטרוול:

Interval	Average7	Slope7	Average19	Slope19	Balance
1	176.8	-0.023	68.8	-0.012	74.0
2	168.0	-0.113	63.7	-0.119	96.0
3	167.9	0.000	65.6	0.046	123.9
4	163.5	-0.046	72.8	0.149	140.9
5	163.5	0.012	76.7	-0.117	138.4
6	163.6	-0.034	68.7	-0.098	139.6
7	163.6	-0.080	62.9	-0.111	137.1
8	161.7	-0.002	70.3	0.295	136.8
9	168.8	0.307	78.6	-0.129	86.0

א. יש לחשב את האנטרופיה של המשתנה Average7 לאחר שחולק לשלושה אינטרוולים עפ"י שתי שיטות דיסקרטיזציה: רוחב שווה ועומק שווה. (15 נקודות)

Interval	Average7	Interval ID Equal width	Interval ID Equal depth
1	176.8		
2	168.0		
3	167.9		
4	163.5		
5	163.5		
6	163.6		
7	163.6		
8	161.7		
9	168.8		

Interval	Prob. Equal width	Prob. Equal depth
1		
2		
3		
Entropy		

ב. יש לחשב את הערך החזוי של משתנה המטרה (Balance) בשיטת Distance-weighted k-NN עבור תצפיות מס' 7 – 9 תוך שימוש בתצפיות מס' 1 – 6 כנתוני האימון (**20 נקודות**). יש להשתמש בנתונים המנורמלים המופיעים בטבלה הבאה וב- $k = 2$.

Record	Average7	Slope7	Average19	Slope19	Balance
1	1.000	0.214	0.376	0.276	74.0
2	0.417	0.000	0.051	0.024	96.0
3	0.411	0.269	0.172	0.413	123.9
4	0.119	0.160	0.631	0.656	140.9
5	0.119	0.298	0.879	0.028	138.4
6	0.126	0.188	0.369	0.073	139.6
7	0.126	0.079	0.000	0.042	137.1
8	0.000	0.264	0.471	1.000	136.8
9	0.470	1.000	1.000	0.000	86.0

	Euclidean Distance			Weight		
	Rec 7	Rec 8	Rec 9	Rec 7	Rec 8	Rec 9
Dist_1						
Dist_2						
Dist_3						
Dist_4						
Dist_5						
Dist_6						
Predicted Balance						
Actual Balance						
Absolute Error						

ג. יש להשתמש במודל SVM with Polynomial Kernel, Degree = 2 ע"מ לסווג את משתנה המטרה (שעבר דיסקרטיזציה) בתצפיות מס' 1 ו-2 (15 נקודות). להלן פירוט המודל:

Record i	Balance	Alpha	$K(x_i, x_1)$	$K(x_i, x_2)$	$\text{Alpha} \cdot y_i \cdot K_{X_1}$	$\text{Alpha} \cdot y_i \cdot K_{X_2}$
1	-1	0.902				
2	-1	0.902				
3	1	0.902				
4	1	0.000				
5	1	0.902				
6	1	0.615				
7	1	0.000				
8	1	0.000				
9	-1	0.615				
Predicted Balance						

$$b = 1.134$$

מה מספר וקטורי התמיכה במודל?

ד. יש לבנות מודל IFN בעל שתי שכבות. השכבה הראשונה תפוצל עפ"י המשתנה Slope7 והשכבה השנייה – עפ"י המשתנה Slope19. יש לפצל כל אחד מהמשתנים ללא קשר לתוצאות הבדיקות של מובהקות סטטיסטית. נקודות הפיצול של כל משתנה מפורטות להלן. 15 נקודות.

Record	Slope7	Slope19	Balance
1	-0.023	-0.012	-1
2	-0.113	-0.119	-1
3	0.000	0.046	1
4	-0.046	0.149	1
5	0.012	-0.117	1
6	-0.034	-0.098	1
7	-0.080	-0.111	1
8	-0.002	0.295	1
9	0.307	-0.129	-1

Slope7

j' / j	-1	Cond.	Joint	1	Cond.	Joint	Total	Cond.
< -0.002								
>=-0.002								
Total								

Conditional Mutual Information

j' / j	-1			1			Total
< -0.002							
>=-0.002							
Total							

Total Layer 0

Slope19

j' / j	-1	Cond.	Joint	1	Cond.	Joint	Total	Cond.
< 0								
>= 0								
Total								

Conditional Mutual Information

j' / j	-1			1			Total
< 0							
>= 0							
Total							

j' / j	-1	Cond.	Joint	1	Cond.	Joint	Total	Cond.
< 0								
>= 0								
Total								

Conditional Mutual Information

j' / j	-1			1			Total
< 0							
>= 0							
Total							

Total Layer 1



ה. יש להציג את התרשים של המודל IFN שנתקבל כולל ההסתברות של כל אחד מהסיווגים בכל אחד מהקודקודים הטרמינליים. **10 נקודות.**



בהצלחה!