

**בית ספר "אפי ארזי" למדעי המחשב המרכז הבינתחומי**  
**The Efi Arazi school of computer science**  
**The Interdisciplinary Center**

סמסטר ב' תשע"ט  
Spring 2019

**מבחן מועד א בלמידה ממוכנת**  
**Machine Learning Exam A**

**Lecturer:** Prof Zohar Yakhini  
**Time limit:** 3 hours

**מרצה:** פרופ זהר יכני  
**משך המבחן:** 3 שעות

Answer 5 out of 6 from the following  
question (each one is 20 points)  
Good Luck!

יש לענות על 5 מתוך 6 השאלות הבאות  
לכל השאלות משקל שווה (20 נקודות)  
בהצלחה!

You can use calculators and the formulae sheets provided  
Justify all your answers and show your calculations  
All answers should be written in exam notebooks

### שאלה 1 (4 סעיפים)

1. ברגרסיה לינארית, למה שנרצה לבצע נרמול לערכי המאפיינים לפני שנתחיל את תהליך הלמידה באמצעות gradient descent?  
האם מקובל לבצע נרמול לפני שימוש בפתרון האנליטי לרגרסיה לינארית  
$$\vec{\theta} = \text{pinv } X \cdot \vec{y}$$
הסבר
2. כחלק מעבודתך כמהנדס למידה חישובית, אתה מפתח מודל רגרסיה לינארית באמצעות מידע,  $(x_i, y_i)$  כאשר  $x_i \in \mathbb{R}, y_i \in \mathbb{R}$ , שהצוות שלך השיג בתנאי מעבדה.  
כל קבוצת נתונים שלך  $(x)$  מגיע מטווח מוגבל, ספציפית  $x_i \in [0, 10]$ . אחרי שאימנת את המודל שלך והשגת שגיאה נמוכה על קבוצת האימון, אתה מתחיל לבדוק את המודל שלך על מידע חדש אשר אינו נאסף בתנאי מעבדה (מגיע מהשטח). מהר מאוד אתה שם לב שנתונים שנאספים מהשטח מגיעים מטווח  $[100, 200]$ .  
א. נניח שבמציאות הפונקציה אותה אתה מנסה למדל היא לינארית.  
כמה טוב הייתה מצפה שהמודל שלך יעבוד על נתונים מהשטח? הסבר.  
ב. אחד מהעמיתים שלך מציע להשתמש ב kNN עם  $k=5$  בשביל לנבא תוצאות עבור נתונים מהשטח. האם זה רעיון טוב? הסבר.
3. חברת נדל"ן מעוניינת לשערך מחירי בתים באמצעות 10 מאפיינים כמותיים.  
בנוסף, יש בתים להם יש חשיבות גבוהה יותר ועל כן חשוב יותר שהטעות בהן תהיה קטנה יחסית לבתים אחרים.  
כיצד היית משנה את הניסוח של משימת הרגרסיה הלינארית כך שהוא יתחשב בדרישות החשיבות הנ"ל?
4. אתה מבצע רגרסיה לינארית על קבוצת נתונים  $(x_i, y_i)$ . כיצד אתה מצפה שה MSE יתנהג תחת התנאים הבאים:  
א. קורלציית פירסון (Pearson correlation) בין  $\vec{x}$  ו  $\vec{y}$  היא 1.  
ב. קורלציית ספירמן (Spearman correlation) בין  $\vec{x}$  ו  $\vec{y}$  היא 1.

## שאלה 2 (5 סעיפים)

1. תהי  $(A_1, A_2, \dots, A_i)$  קבוצת מחלקות המכילות אלמנטים עם מאפיינים  $(x_1, x_2, \dots, x_j)$ .  
כתבי את נוסחת הסיווג עבור:

א. Naïve Bayes

ב. Full Bayes

לפניכם ההתפלגות המשותפת של זריקת שני קוביות (כל קובייה בעלת 6 פאות) בשני בתי קזינו שונים.

קזינו א

Dice1 \ Dice2	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

קזינו ב

Dice1 \ Dice2	1	2	3	4	5	6
1	$\frac{1}{18}$	$\frac{1}{18}$	0	0	0	$\frac{1}{18}$
2	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	0	0	0
3	0	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	0	0
4	0	0	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	0
5	0	0	0	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
6	$\frac{1}{18}$	0	0	0	$\frac{1}{18}$	$\frac{1}{18}$

ההסתברות הפריורית לשחק בקזינו א היא  $\frac{3}{5}$ .

בהינתן תוצאה של זריקת 2 קוביות ברצוננו לנבא האם התוצאה הגיע מקזינו א או ב.

2. צפינו בתוצאה הבאה: קובייה ראשונה – 6, קובייה שנייה – 1. באיזה קזינו יבחר מסווג מסוג Naïve Bayes? הציגי את חישוביך.

3. איזה קזינו יבחר מסווג מסוג Full Bayes עבור אותה תוצאה? הציגי את חישוביך.

4. מהו הפריור המינימלי אשר אם נקצה אותו לקזינו א, קזינו א ייבחר בכל תוצאה אפשרית של הקוביות.

5. מותר לך כעת לשנות שני כניסות במטריצת ההתפלגות המשותפת של קזינו ב. בהינתן אותה תוצאה מסעיף 2, בצעי שינוי במטריצה ככה שקזינו ב ייבחר תחת הפריור שמצאת בסעיף 4.

### שאלה 3 (4 סעיפים)

1. נתון  $X$  משתנה Poisson. נזכיר שאז  $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$  עבור  $\lambda > 0$  כלשהי. נניח שעבור סדרת הגרלות בלתי תלויות מהמשתנה  $X$  קיבלנו את הערכים  $x_1, \dots, x_n$ . כתוב ביטוי המייצג את ה-likelihood של ה-data הנצפה כפונקציה של הפרמטר  $\lambda$ .

2. הוכיח שהערך של  $\lambda$  שיתקבל ע"י MLE במקרה זה הוא:

$$\lambda = \frac{1}{n} \sum_{i=1}^n x_i$$

3. נתון המידע הנצפה הבא הלקוח משתי התפלגויות Poisson שונות עם פרמטרים  $\lambda_1, \lambda_2$ . כל שורה נוצרה מחמש הגרלות בלתי תלויות מאחת משתי ההתפלגויות לעיל:

2	1	11
5	11	5
2	0	2
2	1	2
6	4	8

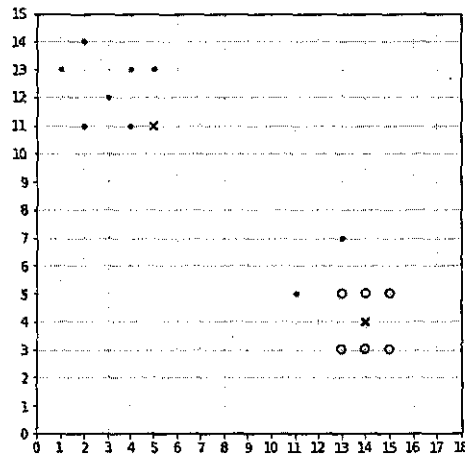
עבור כל שורה הוטל מטבע שבהסתברות  $w_1$  הוביל לשלוש הגרלות עם הפרמטר  $\lambda_1$  ובהסתברות  $w_2$  הוביל לשלוש הגרלות עם הפרמטר  $\lambda_2$ . באיזה אלגוריתם היית משתמש בשביל להעריך את ההתפלגות המשותפת שבבסיס הנתונים המתוארים לעיל. הסביר את בחירתך.

4. השתמש באלגוריתם מסעיף 3 לחשב את הצעד הראשון, כאשר האתחול הינו:

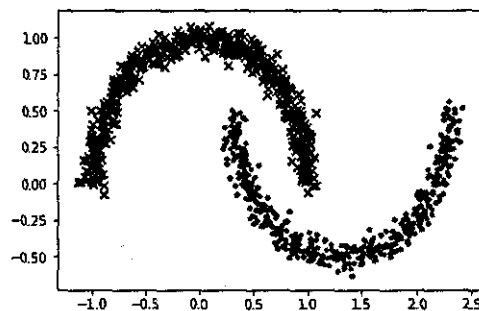
$$\lambda_1 = 1, \quad \lambda_2 = 5, \quad w_1 = 0.75$$

### שאלה 4 (5 סעיפים)

1. נסחי את הפונקציה,  $J$ , אותה מנסה אלגוריתם  $k$ -means הסטנדרטי להביא למינימום.
2. הביטי על הסטטוס המוצג אשר התקבל במהלך ריצת  $k$ -means על מידע ב  $\mathbb{R}^2$ . הנקודות (עיגול ריק) שייכות לאשכול (cluster) 1 בעוד שהעיגולים שייכים לאשכול 2. האיקסים מייצגים את המרכזים של כל אשכול. חשבי את השינוי המדויק ב  $J$  לאחר עדכון האשכולות (לפני עדכון המרכזים).



3. האם הערך של הפונקציה  $J$  יכול לעלות בזמן ריצת אלגוריתם ה  $k$ -means? הסבירי
4. הביטי במבנה האשכולות (clustering) המצורף למטה. האיקסים שייכים לאשכול 1 בעוד שהעיגולים שייכים למקבץ 2. האם אלגוריתם ה  $k$ -means יכול להתכנס למצב שבו הנקודות יסווגו למקבצים הנ"ל?



5. הציעי אלגוריתם ניתוח אשכולות אשר יכול לייצר את מבנה האשכולות הנ"ל כפלט שלו בהינתן הנתונים המוצגים.

## שאלה 5 (4 סעיפים)

1. בהינתן מידע ב  $\mathbb{R}^{10}$  אשר אינו ניתן להפרדה לינארית, סטודנט מציע למפות את המידע באמצעות מיפוי לתבנית רציונלית מלאה מדרגה 10 (full rational variety of degree 10) ואז לנסות להפריד את הנתונים בממד החדש. איזה בעיה יש בגישה הנ"ל וכיצד היית מתגבר עליה?
2. הבט בקבוצת הנתונים הבאה:

X1	X2	Y
+1	0	+1
-1	0	+1
0	+2	+1
0	+1	-1

השתמש בלמה הבאה על מנת להוכיח שהקבוצה אינה ניתנת להפרדה לינארית.  
למה:

הנח שמפריד לינארי  $h$  מסווג לאותו הערך  $y \in \{-1, +1\}$  שתי נקודות  $z, z' \in \mathbb{R}^2$  כלומר  $h(z) = h(z')$ . אזי  $h$  יסווג באותה צורה כל נקודה על הקו שבין  $z'$  ו  $z$ :  
 $\forall \alpha \in [0,1] \quad h((1-\alpha)z + \alpha z') = y$

3. מצא מיפוי  $\varphi$  לתוך ממד לבחירתך שממפה את קבוצת הנתונים מסעיף 2 לקבוצת נתונים אשר ניתנת להפרדה לינארית והגדר את המפריד הלינארי.

4. יהי  $X = \mathbb{R}^2$ . נתון מיפוי  $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , כך ש:  
 $\varphi(x_1, x_2) = (\sqrt{2}x_1x_2, x_1^2, x_2^2)$   
א. מצאי את הקרנל המתאים למיפוי זה.  
ב. נתון מרחב ההיפותזות הבא:

$$H = \left\{ h: \mathbb{R}^2 \rightarrow \{-1, +1\} \mid \begin{array}{l} \text{שתי הקבוצות} \\ \{\varphi(x) \mid h(x) = -1\} \\ - \\ \{\varphi(x) \mid h(x) = +1\} \\ \text{ניתנות להפרדה ב- } \mathbb{R}^3 \end{array} \right\}$$

מהו ה-VC-dimension של  $H$ ? נמקי את תשובתך.

### שאלה 6 (3 סעיפים)

1. הניחי שגרעין מסוג RBF נבחר על מנת ליישם מסווג SVM למידע ב  $\mathbb{R}^2$ . כלומר גרעין המוגדר כך:

$$K(x, y) = \varphi(x) \cdot \varphi(y) = \exp\left(-\frac{1}{2}\|x - y\|^2\right)$$

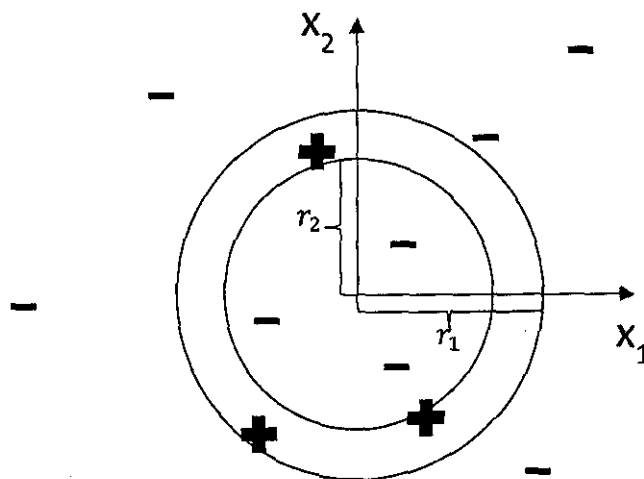
הוכחי שלכל שני נקודות  $x, y \in \mathbb{R}^2$  מתקיים:

$$\|\varphi(x) - \varphi(y)\|^2 = 2 - 2\exp\left(-\frac{1}{2}\|x - y\|^2\right)$$

2. נכון \ לא נכון

אחרי מיפוי לממד גבוה יותר באמצעות גרעין RBF יכול להיות שמסווג מסוג k-NN אשר מבוסס על מרחק אוקלידי ללא משקלים ישיג תוצאות סיווג טובות יותר מאשר התוצאות בממד המקורי.  
רמז: השתמשי בלמה שהוכחת ב1.

3. יהי  $X = \mathbb{R}^2$  ויהי  $C = H$  קבוצת כל המעגלים הקונצנטריים (מעגלים בעלי מרכז משותף). עבור קונספט  $c \in C$  יהיו  $r_1$  ו  $r_2$  הרדיוסים של קונספט הטבעת אשר מגדיר את  $c$  כאשר  $r_1 \geq r_2$  (כמתואר בתמונה). כל נקודת אימון מוגרלת מהתפלגות לא ידועה  $\pi$  ומכילה את מיקום הנקודה  $(x_1, x_2)$  ואת ערך המטרה של  $(1 + \text{אם היא בתוך הטבעת ו } 1 - \text{אחרת})$ .  
תני אלגוריתם  $L$  בעל מורכבות מדגם (sample complexity) פולינומיאלית אשר לומד את  $C$  באמצעות  $H$ . צייני את זמן הריצה ומורכבות המדגם של האלגוריתם  $L$ .  
הוכיחי כל שלב בפתרון.



בהצלחה!