

## זיהוי תבניות ולמידה ממוחשבת

### תרגיל כיתה

1. שלוש תיבות צבעוניות (אדומה, כחולה וירוקה) מכילות פירות.

באדומה יש 3 תפוחים, 4 תפוזים ו- 3 לימונים.

בכחולה יש 1 תפוחים, 1 תפוזים ו- 0 לימונים.

בירוקה יש 3 תפוחים, 3 תפוזים ו- 4 לימונים.

התיבות נבחרות באופן אקראי, עם ההסתברויות הבאות: ההסתברות לבחור את התיבה האדומה היא 0.2, ההסתברות לבחור את התיבה הכחולה היא 0.2, וההסתברות לבחור את התיבה הירוקה היא 0.6. בכל פעם לאחר בחירת התיבה בוחרים פרי באופן אקראי, כך שלכל הפרטים סיכוי שווה להיבחר.

א. מה ההסתברות לבחור תפוח?

ב. אם הפרי שנבחר הוא תפוז, מה ההסתברות שמקורו בתיבה הירוקה?

2. חוקרים בחנו בים שני סוגי דגים, לברק וסלמון. ידוע כי בדרך כלל הלברק קטן באורכו מהסלמון, כאשר אורך הלברק מאופיין בהתפלגות גאוסית (התפלגות נורמלית) עם ממוצע 12.5 ושונויות 6.25, בעוד שאורך הסלמון מאופיין בהתפלגות גאוסית עם ממוצע 17.5 ושונויות 6.25. מספר הלברקים והסלמונים הוא בקירוב שווה.

החוקרים בחנו בכל פעם דג אחד וניסו להחליט על-פי ארכו בלבד האם זהו סלמון או לברק. נסחו כלל החלטה למזעור שגיאת הסווג.

3. במפעל מייצרים שני סוגי ברגים. ידוע כי אורך כל בורג  $x$  מתפלג גאוסית עם

ממוצעים 2 ו- 3 בהתאמה, ועם וריאנס של 0.5 עבור על אחד מהם. כמו-כן ידוע כי מספר הברגים מהסוג הראשון זהה למספר הברגים מהסוג השני, וכן ידוע כי מטריצת העלות (cost matrix) היא:

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0.5 \\ 1 & 0 \end{pmatrix}$$

כאשר  $\lambda_{ij} = \lambda(\alpha_i | \omega_j)$  היא העלות של הפעולה  $i$ , כלומר אריזת הבורג כשייך לסוג  $i$ ,

בהינתן שהבורג שייך למחלקה  $\omega_j$ .

- א. נסחו את כלל ההחלטה הבסיסי למיזעור השגיאה (minimal error), כלומר כתבו באופן מפורש מהו הערך הקריטי  $x_0$  שמעליו הסווג יהיה למחלקה של הברגים הארוכים.
- ב. נסחו את כלל ההחלטה הבסיסי למיזעור הסיכון (minimal risk) באותו אופן.
- ג. נניח כי מספר הברגים הארוכים גדול פי 2 ממספר הברגים הקצרים, כיצד תשתנה ההחלטה בסעיף א'?
- ד. כיצד תשתנה ההחלטה למיזעור הסיכון אם השונות של הברגים הארוכים היא 0.25?
- ה. חזרו על סעיפים א' ו-ב', אם ידוע כי אורך כל בורג  $x$  מתפלג גאוסית עם ממוצעים 2 ו-3, ועם וריאנס של 0.5 ו-0.2 בהתאמה, עבור על אחד מהם. כמו-כן ידוע כי מספר הברגים מהסוג הראשון גדול פי 2 ממספר הברגים מהסוג השני, ומטריצת העלות (cost matrix) היא:

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} 1 & 7 \\ 3 & 1 \end{pmatrix}$$