



המכללה האקדמית תל-חי

קורס: אינטלגנציה עסקית סמסטר ב' תשפ"ד – עבודת סוף

תאריך אחרון להגשה 20/07/2025

הוראות: ענו על כל השאלות. הציגו את דרך הפתרון היכן שנדרש. ספקו הסברים לתשובותיכם.

חלק א': סיווג Classification

שאלה 1:

מאגר הנתונים של מחלת הלב "Framingham" כולל למעלה מ-4,240 רשומות, 16 עמודות ו-15 מאפיינים. מטרת המאגר היא לחזות האם למטופל קיים סיכון למחלת לב כלילית (CHD) במהלך עשר השנים הקרובות. מצ"ב את קובץ: **framingham.csv**

1. תצינו את המאפיינים הכי חשובים באמצעות אלגוריתם Random Forest

2. בעזרת מסווג עץ החלטה, ענו על השאלות הבאות:

1. מהי היתרון המרכזי של מסווג עץ החלטה במקרה זה?
2. אימנו מסווג עץ החלטה על מערך הנתונים (השתמשו ב-70% לאימון ו-30% לבדיקה). דווחו את דיוק המודל, מדדי הדיוק (Precision) והשליפה (Recall).
3. כיצד ניתן להתמודד עם תופעת התאמת יתר (Overfitting) בעץ החלטה?

חלק ב': אשכולות Clustering

שאלה 2:

בהתבסס על תסמיני המטופל, יש לזהות את המטופלים הזקוקים להחייאה מיידית; להקצות כל מטופל לאזור טיפול ייעודי מראש, ובכך לתעדף את הטיפול בו; ולהתחיל בבדיקות/צעדים טיפוליים לפי הצורך.

נעשה שימוש בשלושה מאגרי נתונים נפרדים, שכל אחד מהם מתייחס למחלה או פציעה דחופה אחרת. לכל מאגר יש מאפיינים ותסמינים ייחודיים עבור כל מטופל, ואיחדנו אותם למאגר אחד **patient_dataset.csv** במטרה לזהות מהם התסמינים החמורים ביותר לכל מחלה — וכך לקבוע את עדיפות הטיפול.

בעזרת אשכולות K-means :

- 2.1 זהו את מספר האשכולות האופטימלי באמצעות שיטת המרפק (Elbow Method).
- 2.2 בצעו את תהליך האשכולות ופרשו את התוצאות.
- 2.3 דונו במגבלות האפשריות של אשכולות K-means בתרחיש זה.

חלק ג': כללי אסוציאציה Association

שאלה 3:

מאגר הנתונים של סל הקניות במרכול — (Groceries Market Basket Dataset) שניתן למצוא [כאן] — מכיל 9,835 עסקאות שבוצעו על ידי לקוחות שרכשו מצרכים. הנתונים כוללים 169 פריטים ייחודיים.

המאגר מתאים במיוחד ליישומי כריית נתונים (Data Mining) לצורך ניתוח סל קניות, וכולל משתנים מרובים.

Transaction ID	Items
1	Bread, Milk, Butter
2	Bread, Butter
3	Milk, Eggs, Bread
4	Bread, Eggs, Butter, Milk

מצ"ב את קובץ: [GroceryDataset.csv](#)

בעזרת אלגוריתם Apriori :

3.1 מצאו כללי אסוציאציה עם תמיכה מינימלית של 50% וביטחון מינימלי של 70%.

3.2 פרשו אחד מכללי האסוציאציה שנמצאו.

3.4 דונו ביישום מעשי של כללי אסוציאציה בעולם הקמעונאות.