

מערכות המלצה מבוססות בינה מלאכותית

מטלה 1 – מערכות המלצה שיתופיות – Collaborative Filtering

הוראות כלליות:

- יש להגיש את פתרון המטלה עד לתאריך **17.12.2025** בשעה **23:55** – כל אישור בהגשת הפתרון יגרור הורדת ציון.
- יש להגיש את העבודה בזוגות לאתר הקורס באופן אלקטרוני – רק מגיש אחד יעלה את הפתרון לאתר הקורס.
- יש להגיש לאתר הקורס קובץ ZIP (zip) המכיל שני קבצים:
 - Python Notebook (ipython) מתעדת ומוכנה לבדיקה. מומלץ לבדוק כי המחברת עובדת בסביבת [Colab](#) טרם העלה אותה.
 - דוח (קובץ PDF) המכיל מענה לכל הסעיפים במטלה זו. יש לציין במפורש את שמות המציגים ותעודות הזהות של המציגים בראש המסמך.

הוראות התרגילים:

בתרגיל זה תתקשו למשתמש מערכות המלצה המבוססת על Collaborative Filtering בעבודה עם בסיס הנתונים [MovieLens](#). לצורך העבודה יש להוריד את הנתונים לאתר הקורס ml-latest-small.zip (shnansfo) ווסףקו על ידי [groupLens](#). במטלה זו נתמקד בחיזוי הדירוג (rating) על בסיס הקובץ ratings.csv. ניתן אף רצוי להשתמש בחבילות Python על מנת לפתור מטלה זו. לדוגמה שימוש בחבילת [Surprise](#).

1. שימוש מקדים וביצוע EDA:

- יש לבצע EDA ולהציג לפחות שלושה ניתוחים שונים **וממשמעותיים** אשר תרמו לכמ להבנת הבעיה. יש לפרט כל ניתוח באמצעות תרשימים ומיל שיסביר את משמעותו.
- במטלה זו ובעור שאר הסעיפים נתמקד אך ורק בעמודות: userId, movieId, rating. לכן, עלייכם להסיר כל עמודה נוספת מהנתונים.

2. יישום Collaborative Filtering – Memory Based

- יש למשתמש את הגישות Item-Based ו-User-Based Collaborative Filtering באמצעות KNN על בסיס פונקציית המרחק Cosine.
- עליכם להשתמש בטכנית [cross-validation](#) 5-fold כדיות התוצאות.
- עליכם להציג את המדדים **הממוצעים** עבור MSE ו-MAE.
- עליכם להציג את התוצאות עבור ערכי K שונים מסט הערכים [3,5,10] ולהסביר את המשמעות של כל התוצאות שהתקבלו.
- творס סופי עבור סעיף זה – עבור כל גישה (User-Based/Item-Based) יש לרכז את כל התוצאות שקיבלתם בשני אירורים שונים – אירור עבור כל מדד (MAE/MSE). ציר ה-X בכל אירור יהיה ערך ה-K וציר ה-Y יהיה ערך המדד שהתקבל.

3. יישום :Collaborative Filtering – Model Based

- a. יש למש את המודלים SVD, SVD++, NMF
- b. עליים לשימוש בטכניקת cross-validation 5-לצורך דיווח התוצאות.
- c. עליים להציג את המדדים הממוצעים עבור MAE- MSE.
- d. תוצר סופי עבור סעיף זה – יש להציג את התוצאות השונות בטבלה אחת כך שהשורות בטבלה יהיו המודלים השונים בעוד העמודות בטבלה יהיו המדדים השונים. יש להציג את הערך הטוב ביותר ביותר עבור כל מדד. יש לספק הסבר כללי ומסקנות על התוצאות שהתקבלו.

בצלחה!