בס"ד

**אחזור מידע - תרגיל 1**

תרגיל זה יבוצע ע"י כל אחד מהסטודנטים, העבודה יחידנית, לא בזוגות או שלשות.

התרגיל יחולק שני חלקים, אין קשר ביניהם.

1. המטרה: למידת שימוש במטריצות דלילות (sparse matrices).

עליכם לבנות 2 מטריצות דלילות ולבצע מכפלה ביניהן. במילים אחרות לעשות כפל מטריצות כאשר הייצוג של המטריצות הוא במבנה מיוחד של מטריצה דלילה בשפת פייתון.

1. עליכם לחפש באינטרנט data sets. כל סטודנט יחפש באינטרנט 2 data sets ויכתוב את המידע עליהם. על ה- data sets להיות מהימנים\*. ישנה עדיפות ל- data sets בעברית. אחד מה- data sets צריך לשרת בתהליך למידה של המשימה NER (Name Entity Recognition). ה- data set השני ישרת משימה אחרת. על ה- data sets לא להיות שייכים ל- kagel. עליכם לכתוב עמוד אחד לכל data set המכיל את הקישור ל- data set, שם ה- data set, גודל ה- data set, מספר הדוגמאות (לכל אחד מהסיווגים, אם רלוונטי) המופיעות שם, המשימה שהו משרת ועוד כל מידע רלוונטי למי שיהיה מעוניין להשתמש ב- data set.

\*משמעות data sets מהימן היא כי השתמשו בו במספר מאמרים (ולכתוב אילו מאמרים) או ה- data sets שייך לארגון מוכר (כמו IEEE או AAAI) או שייך לחברה מוכרת כמו גוגל, מיקרוסופט, פייסבוק או שייך לאוניברסיטה מסויימת או שייך לכנס מסויים וכדומה.