

עבודת בית מסכמת - למידה חישובית

עבודת סיכום קורס.

לעבודה יש כמה חלקים. החלק הבסיסי הוא כמובן להסיק מסקנות מדאטה מעניין כלשהוא.

שלב 1 – בחירת הנושא והדטה (30 נקודות)

הדאטה יכול להיות

- (א) מהעבודה שלכם, מהחיים.
- (ב) מתחרויות (עדיף שלא יסתימו – תציינו את המיקום שלכם אל תתבאסו ממנו 😊 בעיקר מהאתר: <https://www.kaggle.com>
- (ג) מהאינטרנט (מאמרים בתחום שעבדו על דאטה ועוד)

ככל שהדאטה יהיה יותר מאתגר, יותר מעניין, עם יותר השלכות, ועם עבודה שלכם להוציא ממנו את וקטור הפיצ'רים שיכנס לאלגוריתם הלמידה כך הניקוד יהיה גבוה יותר. אתם יכולים לחפש דאטה מעניין גם באתר מסעיף ב' וגם לראות דוגמאות לפרויקטים מעולים שנעשו בקורס cs229 בלינק: <http://cs229.stanford.edu/projects.html>

שלב 2 הרצת אלגוריתם למידה על הדטה שבחרתם

המטרה שלכם היא שימוש בכמה שיותר כלים שנלמדו בקורס. עליכם להראות:

- 1) השוואה בין לפחות שני אלגוריתמים (יכולים להשתמש גם בכאלה שלא למדנו)
- 2) כוונן אלגוריתמים בהתאם לדברים שנלמדו בקורס
- 3) גרפיים שמראים איך השפעה של פרמטרים משפיעים על הלמידה
- 4) פיצ'רים (תכונות) חשובים/לא חשובים
- 5) התמודדות עם overfitting .

אין צורך לכתוב אלגוריתמים! השתמשו בספריית sklearn אבל בהחלט להשתמש בפרמטרים לכוון וכו

עליכם להגיש קובץ zip שמורכב מ:

1) report- דוח עם כחמישה עמודים שמתאר את הבעיה, הדאטה התוצאות והמסקנות שלכם. בראש הדוח שמותיכם ות.ז.

2) הקוד שכתבתם

3) test- קובץ הרצה שנקרא test.py או לינק למחברת הפיתון

ניפגש ב 30.9 לרבע שעה לזוג (יפורסם קובץ אקסל כל זוג באיזה שעה נבחן) כדי לבדוק מה עשיתם לשאול שאלות ולתת ציון. רק במקרים נדירים אסכים לשינוי היום! נא להגיש עד יום קודם!