

## סיווג חשבונות בוטים ב-Tweeter

### מגשים:

שם: אסף חאייק ברוך  
שם: בן בנוז  
ת.ז: 206783441  
ת.ז: 207570573

### תיאור הבעיה:

כיום הרשתות החברתיות הן חלק משמעותי ובלתי נמנע מחיי היומיום שלנו בעולם המערבי. יש להן השפעה רבה על האופן שבו העולם מתנהל מבחינת פרסומות, ארגון קהילתי, פוליטיקה ודברים חשובים אחרים.

דבר שיוצר בעיות ביחד עם ההשפעה הזאת הוא קיומם של חשבונות מזויפים שעושים פוסטים באופן אוטומטי ומנסים להטות את דעת הקהל. מדברים קטנים כמו יצירת פרסום חינוך למוצר עד לדברים מהותיים ביותר כמו השפעה על בחירות - הבוטים האלה גורמים לערעור היציבות של המרקם החברתי.

המטרה שלנו היא שבהינתן משתמש ב-Tweeter נרצה לסווג אם אותו המשתמש מופעל ע"י אדם אמיתי או ע"י בוט.

### פתרון הבעיה:

הדרך שחשבנו שאיתה אפשר לפתור את הבעיה היא הדרך הבאה:

1. נוריד מאגר נתונים של משתמשים מסווגים (קיימים מספר מאגרים כאלה שהם תוצאתם של מאמרים מדעיים בנושא).
2. ניקח מכל משתמש את האינפורמציה הרלוונטית עליו ועל כל הפוסטים האחרונים שלו.
3. נמיר את אובייקט המשתמש לטנזור פיצ'רים באמצעות Feature Extractor שנגדיר וכנראה נלמד, כולל המרת מילים לוקטורים באמצעות ספריה מסוג Word2Vec והוספת הקשר חיצוני מ-wikidata ומ-GDELT.
4. לאחר עוד כמה שלבי עיבוד מקדים (כולל סטנדרטיזציה ו-feature selection) נעביר את הטנזור ואת הסיווגים לאלגוריתם למידה (כנראה רשת עמוקה).

### תיאור המערכת:

למערכת יהיו את הרכיבים הבאים:

- מאגר נתונים מסווג\מאסף מידע.
- יוצר אובייקטים (מוודא קיום של פורמט אחיד לכל הדוגמאות).
- המרת אובייקט לטנזור כללי (ספציפית מילים ל-embedding באמצעות Word2Vec).
- Feature Extractor.
- Feature Selection.
- אלגוריתם למידה\מודל.

## ניסויים:

- נרצה לבדוק את ההשפעה של הקודקס שאיתו נלמד את מנגנון ה-Word2Vec. לדוגמא האם כדאי ללמד רק על הציוצים במאגר, על ציוצים באופן כללי או שילוב של ציוצים וטקסטים אחרים.
- נרצה לבדוק את ההשפעה של מספר הפיצ'רים שאנו רוצים להשתמש בו מתוך כולם.
- נרצה לבדוק השפעה של מסווגים שונים עם היפר פרמטרים שונים.
- נרצה לבדוק את ההשפעה של סוגי רגולריזציה שונים על המודל.
- נרצה לבדוק את ההשפעה שיש בכלל להקשר החיצוני שאנו מוסיפים לציוצים.

## פרטים טכניים:

אנחנו נכתוב את הפרויקט הזה בשפה **python**, ונשתמש במנהל החבילות **anaconda**.  
בכדי לשלוף מידע מטוויטר נשתמש בחבילה **tweepy** שמאפשרת לנו לעבוד עם הממשק של טוויטר ולהוריד ממנו את המידע הרלוונטי.  
נרצה להשתמש במימוש **Word2Vec** של **gensim** בכדי ליצור קידוד של מילים לוקטורים.  
נשתמש במאגר הנתונים **GDEL** בכדי להשיג הקשר לאירועים וב-**wikidata** וקטגוריות ויקיפדיה בכדי להשיג הקשר כללי.  
נשתמש ב-**scikit-learn** בשביל דברים כלליים של למידה ומסווגים פשוטים וב-**pytorch** בכדי לבנות מסווג מורכב יותר.