

# פרויקט חלק א' – עיבוד שפות טבעיות בטכניקות למידה عمוקה

תאריך הגשה - 11/12/2025

## **מטרת הפרויקט**

בפרויקט זה תבנו מערכת לדיאוי רגשות בטקסט תוך שימוש בטכניקות של למידה عمוקה. המערכת תקבל משפטים קצרים (תגובה מטויויר) ותשוג כל משפט לאחד מששת סוגי הרגשות הבאים:  
0: sadness, 1: joy, 2: love, 3: anger, 4: fear, 5: surprise

המטרה היא לתרגל בניה, אימון והערכה של מודלים לעיבוד שפה טבעית ולהשווות בין ביצועי ארכיטקטורות שונות.

## **Dataset**

ה-dataset\* אליו תעבדו מכיל ציטוטים של תגובות מטויויר מסווגות לאחת מבין שש סוגי הרגשות המציגים לעלה. ישנו שני קבצים, קובץ אחד הוא קובץ האימון אותו תאמנו את הרשת וקובץ נוסף שהוא קובץ הolidzia עליו תראו את התוצאות. הקוד צריך לתרום בהרצה של קובץ test.csv שתיבדק על ידי.

## **דרישות**

בחלק זה תישמו רשת ניורונים מסוג GRU-N-LSTM על מנת לבצע מיפוי Emotion Detection ותדרשו לבנות מערכת יציבה בה ניתן לשנות פרמטרים על מנת לבצע השוואה ולקבל החלטה לגבי המודל המיטב. עבור כל רשת יש לשים את אחת מושיות Embeddings שנלמדו בקורס (למשל GloVe או Word2Vec) בונוס – תוכלו לנסות לשים שיטת Embedding משלכם.

עבור כל רשת יש לבדוק את השפעת שינוי-hyperparameters השונים ולהציג את ההבדלים בתוצאות (אם ישנו) על גבי קובץ叫做 validation.

## **קריטריונים להערכתה**

- הקוד צריך למלא את דרישות המטלה באופן מלא וכן להיות כתוב באופן קרייא וברור
- עליוכם להציג באופן ברור את ההשוואה בין רשות הניורונים השונות וכן השפעת שינוי-hyperparameters .

\*בחלק זה לא תימדדו על תוצאות המודלים אלא על אופן היישום אך שימוש בהרשות כן צריכה להתכנס ולקבל תוצאות סבירות, השתדלו להגיע לדיווק של 75% ומעלה.

## **קבצים להגשתה**

- קובץ/מחברת Python המכילים את הקוד המלא והערות ולונטיות.
- דוח באורך של עד 3 עמודים המסכם את עיקרי התהילה, הניסויים והתוצאות.

\*Reference : CARER:Contextualized Affect Representations for Emotion Recognition