

תרגיל 4 רשת LSTM.

- את התרגיל הבא נעשה בסביבת keras.
נקודת ההתחלה היא הקוד הנתון ב[git הזה](#).
נעבוד עם התוכנית הנקראת words.py.
בתוכנית תמצאו רשת LSTM ממומשת בkeras, הרשת לומדת מודל שפה מתוך סט נתונים נתון (ויקיפדיה כברירת מחדל) ולאחר מכן מחוללת משפטים באמצעות המודל הנלמד.
- (1) בשלב הראשון יש לחלק את סט הנתונים ל 3 קבוצות: train, validation, test, ביחסים של 80-10-10.
 - (2) יש לייצר מדד perplexity אותו נפעיל לאחר האימון על כל אחד משלושת סטי הנתונים שלנו.
 - (3) יש לקרוא בתיעוד של keras ולמצוא דרך להפעיל את הרשת על ה-datan בכיוון ההפוך. כלומר, הרשת תוכל לבצע למידה רגילה שבה סדר המילים כמו הסדר שבהן נכתבו, או לחילופין הרשת תוכל ללמוד את הסדר ההפוך שלה מילים (מהסוף להתחלה).
 - (4) יש לאמן 4 רשתות LSTM שונות, שבהן יש 1 או 2 שכבות של נוירוני LSTM, ולכל אחת מהרשתות פעם לאמן כרגיל ופעם בכיוון ההפוך.
 - (5) יש לכתוב פונקציה המקבלת משפט ומודל מאומן ומחזירה את ההסתברות למשפט על פי המודל.
 - (6) יש לבצע חילול של משפט באורך 7 המתחיל במילים "I love" ובטמפרטורה של 0.1, 1 ו-10.
 - (7) יש לכתוב פונקציה שנקראת לאחר האימון ומאפשרת UI, שבו ניתן להכניס מילה ולקבל את המילה הבאה החזויה על פי המודל המאומן.
 - (8) יש לאמן את ארבעת המודלים ולתעד את תוצאות ה perplexity עבור כל אחד מהם לכל אחד משלושת הסטים, סה"כ 12 תוצאות.
 - (9) יש לחשב את ההסתברות לכל משפט שהתקבל בסעיף 6 ובנוסף לחשב את ההסתברות למשפט "i love cupcakes"

לכל סעיף יש לתעד את הפלטים והמדדים הרלוונטים, במידה ויש כאלה, ולהגישם בצירוף הקוד.