

## Deep-learning-assignment 3, part2:

בחלק זה נתבקשנו לייצר 3 שפות שהחננו לא יצליח ללמוד.

- **השפה הראשונה - palindrome:**

מילה שהיא פולינדרום-pos, כל מילה אחרת היא neg.

- **למה חשבנו שהיא תהיה קשה ללמוד?**

פולינדרום זה שפה לא רגילה, זאת אומרת שפה סימטרית וכדי להצליח לזהות את השפה, הוא יצטרך לעבוד קשה יותר על מנת למצוא את נקודת הסימטריה המדויקת, ובנוסף יצטרך כמות משתנה של זיכרון. יכול להיות שאם היינו משתמשים ב-LSTM-bidirectional הוא היה עובד פחות קשה, מכיוון שיש לו גישה בזמנית גם לכיוון הראשון וגם לכיוון השני ויותר אפשרות למצוא את נק' הסימטריה.

- **למה השפה נפלה ב-LSTM\_acceptor?**

כמו שצינתי בשאלה הקודמת המודל שלנו עבד עם LSTM פשוט. ובשביל לבדוק אם הקלט הוא פולינדרום המודל יידרש לדעת גם את הכיוון השני של המילה ולהשוות כל הזמן.

- אנחנו הרצנו את המודל שלנו על 4000 דוגמאות של positive ו-4000 דוגמאות של negative בשביל הtrain ולקחנו 1500 דוגמאות בשביל הdev\_set. הצלחנו להגיע אחרי 10 אפוקים ל-58%.

- **השפה השנייה-הסכום של ערכי הASCII האותיות מצד ימין שווה לסכום של ערכי הASCII של האותיות מצד שמאל:**

- **למה חשבנו שהיא תהיה קשה ללמידה?**

חשבנו שלמודל יהיה קשה ללמוד את השפה מכיוון שיצטרך כמות אקספוננציאלית של זיכרון כדי לשמור את הסכומים של האסקי על כל sub-set של המילים.

- **למה השפה נפלה ב-LSTM\_acceptor?**

באופן כללי מלבד העובדה ש-LSTM חסר יכולות מתמטיות וקשה לו כל פעם מחדש להשוות ולחלק ל-2 תתי קבוצות. העבודה שאנחנו עובדים עם LSTM פשוט ולא ב-LSTM-bidirectional לא מסייעת ומוסיפה קושי להצליח ללמוד את השפה.

- **השפה השלישית - ww, double word (המילה מופיעה פעמיים) :**

- **למה חשבנו שהיא תהיה קשה ללמידה?**

חשבנו שלמודל יהיה קשה ללמוד את השפה מכיוון שהמודל כל פעם רואה רק חלקים מהמילה ולא את המילה כולה ויהיה לו קשה לזהות מתי יהיה צריך לעצור ולהגיע לאמצע ולהתחיל את המילה מחדש.

**למה השפה נפלה ב-LSTM\_acceptor?** כמו שאר הדוגמאות ה-LSTM\_acceptor עובר על המילה רק בכיוון אחד ובגודל חלון מצומצם כך שלא יכול לנחש את העתיד של הרצף או לזכור את הרצף עד כה. אולי ב-LSTM-bidirectional בגלל שעובר שוב על הרצף וראה את כולו בשלמותו יוכל לעבוד פחות קשה.

- אנחנו הרצנו את המודל שלנו על 5000 דוגמאות של positive ו-5000 דוגמאות של negative בשביל הtrain ולקחנו 1000 דוגמאות בשביל הdev\_set. הצלחנו להגיע אחרי

אסנת אקרמן 315747204  
שירה יאיר 315389759

בס"ד

10 אפוקים ל-51%. וכמובן שניסינו לשחק עם ההיפרפרמטרים ועדיין הaccuracy נשאר  
באיזור ה-49%-51%.