**NLP – assignment 3**

Koren Abitbul - 318796448, Pan Eyal - 208722058

Question 1

Q1.1

עבור מילים במשפט הממוספרות לפי סדר הופעתן במשפט, נסמן:

* - גודל ה-vocabulary.
* בתור ווקטור בגודל , כך שכל ערכיו הם 0 מלבד הערך במיקום שבו נמצאת המילה .
* בתור פלט של המודל, ווקטור בגודל , כך שערכיו הם ערך פונקציית ההסתברות של המודל לחיזוי המילה בהינתן המילים הקודמות,

היינו רוצים לבטא את פונקציית ההפסד בצורה הבאה, בעזרת cross entropy:

וכך היינו גם יכולים לאמן בצורה unsupervised, אבל אין לנו את ההסתברויות האמיתיות של כל מילה.

מה שכן אנחנו יכולים לעשות, זה להשתמש בsupervision וב-one-hot vector שהגדרנו מקודם וכך ניתן לבטא את פונקציית ההפסד כ:

Q1.2

המודל מנבא את המילה הבאה רק על סמך חמש המילים הראשונות, בחיזוי המילה השישית אין בעיה לקבוע הערכה על ביצועי המודל. אך למעשה, עבור משפטים ארוכים יותר מ-6 מילים, כדי שהמודל ייבחר את המילה השביעית עד למילה ה-n-ית, הוא צריך להתבסס על מילים אשר גם כן, הוא בחר בעצמו. כאן ייתכנו אי דיוקים להערכת המודל. מכיוון שאם המודל לא היה מדויק בבחירת המילה במקום השישי לדוגמא, אין לנו בהכרח דרך לוודא שהמודל בחר את המילה השביעית כראוי בהינתן המילים הקודמות (כולל המילה הלא נכונה במקום השישי).

לכן שגיאות של המודל בהתחלה יכולות להיות "מפועפעות" להמשך המילים שהוא צריך לחזות.

Question 2

Q2.1

לא הבנתי

Q2.2

סדר הפעולות לפיarc-standard יהיה:

* SHIFT, SHIFT, SHIFT, L(case), L(det), SHIFT, SHIFT, L(aux), L(nsubj),  
  SHIFT, R(amod), SHIFT, R(punct), R(root)

Q2.3

לאחר הפעולה השביעית (SHIFT) המצב המלא של המנתח הינו:

|  |  |
| --- | --- |
| קומיים | . |

|  |
| --- |
| מערכונים |
| היו |
| \_הם |
| ROOT |

|  |
| --- |
| \_הם -> case -> שניים |
| \_הם -> det -> מ |

Buffer: Stack: Added arcs:

Q2.4

נדרשות אפס פעולות נוספות?