

NLP – assignment 3

Koren Abitbul - 318796448, Pan Eyal - 208722058

Question 1

Q1.1

עבור w_1, \dots, w_{20} מילים במשפט הממוספרות לפי סדר הופעתן במשפט, נסמן:

- $|V|$ - גודל ה-vocabulary.
- y_k בתור ווקטור בגודל $|V|$, כך שכל ערכיו הם 0 מלבד הערך במיקום שבו נמצאת המילה w_k .
- \hat{y}_k בתור פלט של המודל, ווקטור בגודל $|V|$, כך שערכיו הם ערך פונקציית ההסתברות של המודל לחיזוי המילה w_k בהינתן המילים הקודמות, $P_{model}(w_k | w_1, \dots, w_{k-1})$.
- $n = 20$

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

$$L_{CE} = - \sum_{k=1}^n P(w_k | w_1, \dots, w_{k-1}) \cdot \log(\hat{y}_k)$$

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

מה שאנחנו יכולים לעשות, זה להשתמש ב-supervision וב-one-hot vector שהגדרנו מקודם וכך ניתן לבטא את פונקציית ההפסד כ:

$$L_{CE} = - \sum_{k=1}^n y_k \cdot \log(\hat{y}_k)$$

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 יבחר 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).

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

Question 2

Q2.1

התרגום האנגלי של המשפט הוא: Two of them were comedy skits.

סוג הקשת שתחליף כיוון הוא: **amod**, זאת כיוון שסדר המילים "מערכונים קומיים" מתחלף באנגלית ל-"comedy skits". לכן גם הכיוון של הקשת תתחלף.

Q2.2

סדר הפעולות לפי arc-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:

Buffer:	<table><tr><td>קומיים</td><td>.</td></tr></table>	קומיים	.	Stack:	<table><tr><td>מערכונים</td></tr><tr><td>היו</td></tr><tr><td>הם</td></tr><tr><td>ROOT</td></tr></table>	מערכונים	היו	הם	ROOT	Added arcs:	<table><tr><td>הם -> case -> שניים</td></tr><tr><td>הם -> det -> מ</td></tr></table>	הם -> case -> שניים	הם -> det -> מ
קומיים	.												
מערכונים													
היו													
הם													
ROOT													
הם -> case -> שניים													
הם -> det -> מ													

Q2.4

סדר הפעולות לפי arc-eager יהיה:

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

ניתן לראות שכאן נדרשו **17 פעולות**. 3 פעולות יותר מהשיטה הקודמת, בדיוק כמספר הקשתות קדימה בגרף (RightArc) אשר להן נצטרך לבצע פעולת REDUCE נוספת.