

# NLP – assignment 2

Koren Abitbul - 318796448, Pan Eyal - 208722058

## Question 1

$$\hat{\theta}_R = \operatorname{argmax} \left[ \sum_{i=1}^m \log p[y^{(i)} | x^{(i)}] - \gamma \sum_{j=1}^n \theta_j^2 \right] = \operatorname{argmax} \operatorname{func}_R(\theta)$$

$$\hat{\theta}_L = \operatorname{argmax} \sum_{i=1}^m \log p[y^{(i)} | x^{(i)}] = \operatorname{argmax} \operatorname{func}_L(\theta)$$

נקרא לפונקציית המטרה עם הרגולריזציה  $\operatorname{func}_R(\theta)$  ולפונקציית המטרה ללא הרגולריזציה  $\operatorname{func}_L(\theta)$ .

נניח בשלילה  $\|\hat{\theta}_R\|_2^2 > \|\hat{\theta}_L\|_2^2$ : נבחין כי  $\operatorname{func}_R(\theta) = \operatorname{func}_L(\theta) - \gamma \|\theta\|_2^2$ .

לכן:  $\operatorname{func}_R(\hat{\theta}_L) = \operatorname{func}_L(\hat{\theta}_L) - \gamma \|\hat{\theta}_L\|_2^2$  וגם  $\operatorname{func}_R(\hat{\theta}_R) = \operatorname{func}_L(\hat{\theta}_R) - \gamma \|\hat{\theta}_R\|_2^2$ .

כמובן ש:  $\operatorname{func}_L(\hat{\theta}_L) \geq \operatorname{func}_L(\hat{\theta}_R)$  כי  $\hat{\theta}_L$  ממקסמת את  $\operatorname{func}_L(\theta)$  וגם  $\|\hat{\theta}_R\|_2^2 > \|\hat{\theta}_L\|_2^2$  מההנחה בשלילה, לכן, עבור  $\gamma > 0$  אנחנו מקבלים  $\operatorname{func}_R(\hat{\theta}_L) > \operatorname{func}_R(\hat{\theta}_R)$  בסתירה לנתון ש  $\hat{\theta}_R$  ממקסם את  $\operatorname{func}_R(\theta)$ .

## Question 2

### Q2.1

המידע הזה תומך בטיעון מכיוון שהמילה "gonna" היא סלנג של "going to" (שמתאים רק כשלאחר מכן מגיע פועל), שמכיל את המילה "to". במשפט השני gonna לא מתאים למשפט, אך going to כן מתאים, לכן אפשר להגיד של to יש 2 תפקידים. בתפקיד אחד מגיע פועל במשפט אחרי ה-to (וניתן להשתמש בgonna) ובמקרה כזה נשתמש בתג TO. בתפקיד השני שלא מגיע פועל אחרי (כמו going to the beach), נצטרך למצוא חלק דיבר חלופי.

במקרה השני, תג עבור המילה to יכול להיות מילת יחס (preposition). במשפט השני אם נחליף את gonna עם going to המילה to משמשת כמילת יחס בדיוק כמו המילה in במשפט she is going in the water.

### Q2.2

נוכיח שלא ניתן לבצע סכמה במסגרת תיוג רצפים שתאפשר טיפול בישויות מקוננות:

תחילה נראה למה סכמת תיוג לא עובדת. עבור המשפט "הועד הפועל של רשות השידור" ישנם 2 ישויות, כל המשפט ורק החלק "רשות השידור". שתי הישויות נגמרות במילה "השידור" ולכן צריך תג מיוחד שיבשר ש-2 הישויות נגמרות במילה הזאת ותג 1 לא מספיק. אפשר להבין אינטואיטיבית שכדי לפתור את הבעיה צריך שימוש בסוגריים, למשל: "{הועד הפועל של {רשות השידור}}".

בשביל 2 הסוגריים בסוף, צריך תג מיוחד שיסמן על שני סוגרים. אותו דבר אם 2 ישויות שונות היו גם מתחילות באותה המילה.

בחזרה להוכחה, נניח בשלילה שאכן יש סכמה כזאת. כמות התגים בסכמה הזאת צריכה להיות סופית, נסמנה ב-M. ניצור טקסט\משפט באורך גדול מזה שמכיל N ישויות מקוננות (כך ש:  $N > M$ ) שנגמרות באותה המילה. מאחר ו- $N > M$ , אין תג שיבשר את סיום N הישויות המקוננות, לכן הסכמה לא פותרת את הבעיה.

### Q3

מוצג בעמוד הבא כהדפסה של colab notebook.