**NLP – assignment 5**

Koren Abitbul - 318796448, Pan Eyal - 208722058

שאלה 1

1.1

המרכיבים השונים הם השאילתה (Q), המפתח (K), ערך (V).

מנגנון הצומי העצמי מחשב את הscore בין השאילתה (Q) לכל מפתח (K).

על ידי: כאשר: .

ציון זה מציין עד כמה כל מילה אחרת (Key) רלוונטית למילה הנוכחית (Query).

כך, בהינתן קשת התלות הרלוונטית , נסתכל פשוט על הscore: .

1.2

ייתכן שנמצא קשרי צומי חזקים יותר בשכבות שונות של הרובוטריק.

תכונה אפשרית ראשונה: מרחק התלות.  
בשכבות המוקדמות יותר של הרובוטריק המודל עשוי ללכוד תלויות של מילים שקרובות יותר זו לזו במשפט.  
ולכן תלויות המילים הרחוקות זו מזו לא יהיו חזקות בשכבות הראשונות כמו בעמוקות יותר.

תכונה אפשרית שניה: מורכבות התלות.  
תלות מורכבת יותר עשויה לחייב את המודל לשלב מידע מחלקים שונים של המשפט. קשרים אלה עשויים להופיע בשכבות מאוחרות יותר, בהן הרובוטריק עיבד מידע דרך שלבי צומי עצמי מרובים ויכול להתמודד עם תלות בטווח ארוך יותר.

1.3

בעזרת האלגוריתם החמדן לניתוח תלויות במעברים, לאחר שנבצע forward pass על הרובוטריק, נשיג את הscore עבור הצומי העצמי בין כל צמדי המילים עבור כל השכבות. נוכל לבדוק את הscore עם שאר המילים לפי שכבות שונות ולאט לאט לבנות את התלויות התחביריות.

שאלה 2

2.1

נוכיח שהפונקציה הזאת היא ההופכית של פונקציית הסיגמואיד.

עבור פונקציית הסיגמואיד נסתכל על כאל הפלט של הסיגמואיד בהינתן הקלט ונחפש פונקציה ש הוא הפלט שלה בהינתן :

*2.2*

נעשה את אותו הדבר עבור פונקציית ה;softmax ,

קיבלנו תלות ב-, לכן אין לפונקציה הופכי. וזה הגיוני כי לדוגמה אם בsoftmax היינו מקבלים התפלגות אחידה, אז הלוגיטים יכולים להיות כל וקטור אשר הערכים שלו הם זהים.   
בשביל לחשב את הלוגיטים יש צורך בלדעת את הערך:

2.3

בשביל לחשב את הלוגיטים יש צורך בלדעת את הערך:

2.4

אם ניקח את k האסימונים המובילים ישירות מהלוג'יטים, ולאחר מכן ננרמל בעזרת הsoftmax הדבר יחזיר לנו תוצאה זהה ללקיחת k האסימונים מהsoftmax ולאחר מכן לנרמל אותם.

מכיוון שפעולת הsoftmax היא מונוטונית, בשתי הדרכים נקבל את אותם רכיבים מובילים.  
כלומר, עבור לוג'יטים , אם אז גם .

כעת, עבור הערך בהתפלגות שלהם נראה:  
*עבור אוצר מילים ולקיחת האסימונים מהsoftmax:*

נסמן כך ש *ו- שייכים ל.  
ההסתברויות שנקבל לאחר נרמול הן:*

*כאשר הוא הווקטור המכיל את הערכים הגבוהים של ווקטור .*

*כלומר, קיבלנו את מה שרצינו להראות שמתקיים שווין.*

3

ישנם טעויות בהם כתוב פעמיים “the” כאן:

תמונה שמכילה טקסט, צילום מסך, גופן, מספר

התיאור נוצר באופן אוטומטי

וכאן:

תמונה שמכילה טקסט, צילום מסך, גופן, מספר

התיאור נוצר באופן אוטומטי

עוד קצת כפילויות של מילים:

תמונה שמכילה טקסט, צילום מסך, גופן, מספר

התיאור נוצר באופן אוטומטי

כאן אנחנו חושבים שזה צריך להיות במקום usual, usually (?)

תמונה שמכילה טקסט, צילום מסך, גופן, מספר

התיאור נוצר באופן אוטומטי

וכאן יש היפר-לינק כושל ביותר:

תמונה שמכילה טקסט, גופן, צילום מסך, לבן

התיאור נוצר באופן אוטומטי