עיבוד שפות טבעיות

תרגיל בית 4

שם: רנים אבראהים (212920896) , אסיל נחאס (212245096)

חלק א:

במשימה זו בנינו מודל Word2Vec ליצירת וקטורי word embeddings מתוך הקורפוס של הכנסת של תרגיל בית 1. בבנית המודל השתמשנו בפרמטרים הבאים:

1. **vector_size=100: הגדרנו את גודל הווקטורים ל-100 הסיבה מכך הינה כי לאפשר למודל לייצג בצורה טובה את המשמעות של המילים בהקשר שלהן. לא בחרנו גודל יותר גודל לחשש לבעיות בביצועים ו overfitting .**
2. **window=5: גודל חלון ההקשר נבחר כ-5, משמעות דבר זה הוא שהמודל מתחשב ב 5 מילים לפני ואחרי המילה הנוכחית. ערך זה הוא טוב על מנת לעשות איזון בין הקשר רחב לצר כך שהמודל יכול לכלול מידע הקשרי רחב מספיק בלי לאבד את הפוקוס על הקשרים המקומיים.**
3. **min_count=1: השתמשנו בערך זה כדי להבטיח שגם מילים נדירות ייכללו במודל. הסיבה מכך היא שהקורפוס מכיל מידע ממוקד (כמו דיונים פרלמנטריים) לכן חשוב לכלול גם מילים שמופיעות רק פעם אחת כדי שלא נאבד מידע משמעותי.**

1. **מה המשמעות, היתרונות והחסרונות של הגדלת והקטנת גודל הוקטור- size_vector ?**

גודל הווקטור מגדיר את מספר הממדים שבהם מיוצגות המילים במודל. הגדלת גודל הווקטור מאפשרת למודל לייצג משמעויות ותכונות מורכבות יותר של מילים, וזה מועיל במיוחד בקורפוסים גדולים ומגוונים. עם זאת, גודל וקטור גדול מדי ידרוש יותר זיכרון וזמן חישוב ועלול לגרום למודל להיות פחות יעיל. מצד שני, הקטנת גודל הווקטור תקטין את הדרישות החישוביות ותהיה יעילה יותר בקורפוסים קטנים, אך עלולה לפגוע ביכולת המודל לייצג משמעויות מורכבות של מילים.

2. **מה הבעיות שיכולות לעלות משימוש במודל הנ"ל , שאומן על הקורפוס שלנו. התייחסו בתשובתכם לאופן שבו יצרנו את הקורפוס, לגודל שלו ולשימושים פוטנציאלים של המודל?**

א. **הבנת הקשר מוגבלת : המודל אומן רק על קורפוס הכנסת ולכן משקף את השפה, ההקשרים הייחודיים לו. הדבר מגביל את השימושיות שלו בתחומים אחרים או במשימות שיחה כלליות.**

ב. **גודל הקורפוס: בקורפוס יש כ-110,415 משפטים וכ-80,105 מילים, אך זהו גודל קטן יחסית למודלים מתקדמים שמאומנים על מיליוני משפטים. גודל כזה עלול לא ללכוד את כל מגוון ההקשרים בשפה העברית ולהטעות את המודל.**

ג. פיזור המילים בקורפוס :כמות המילים הייחודיות בקורפוס היא כ-80,105, יחסית ל-110,415 המשפטים שבו. יחס זה מצביע על כך שחלק גדול מהמילים מופיעות רק לעיתים רחוקות מאוד. מצב זה מקשה על המודל ללמוד את ההקשרים והמשמעויות של מילים נדירות, מכיוון שאין מספיק דוגמאות ללמידה מהן. התוצאה היא ייצוג וקטורי פחות מדויק עבור מילים אלו וזה יוביל לפגיעה בביצועים לשימושים במילים אלה בעתיד.

חלק ב:

סיכום השלבים למציאת מילים חלופיות עבור המילים המסומנות:

1. דקות: המטרה הייתה למצוא מילה שמתארת זמן כגון שניות או ימים, בהתחלה קיבלנו מילים כמו חמש, שלוש והדקות לכן החלטנו לשים מילות חיוביות שמתארות זמן כגון "דקה", "רגעים", ו"זמן". אחרי ההוספה הזו קיבלנו את המילים שניות, ספורות, הדקות ובחרנו את שניות להיות המילה המחליפה מכיוון שהיא מתארת זמן קצר בצורה ברורה ושומרת על המשמעות המקורית של המשפט. ולכן הצלחנו במשימה במילה הזו.

2. הדיון: המטרה היא למצוא מילה שמתארת דו שיח. בהתחלה קיבלנו מילים כמו "הנושא","לדיון" ו"בדיון" אשר לא נכונות לכן נתנו למודל מילים חיוביות כמו "דיבור", "מפגש", "מועצה", "שיחה", שמייצגות את הרעיון של אינטראקציה ושיתוף פעולה. הוספנו כמו כן את המילים קונפליקט", "ויכוח", "סכסוך", למילים שליליות כדי למנוע תוצאות עם משמעויות שליליות. יש כמה מילים שגם הוספנו אותם בצד השלילי והחיובי אך דחינו אותם אחרי שניתחנו את השפעתם, בסוף קיבלנו את התוצאות, "השרה", "הישיבה" ו"מכתבו" ובחרנו את "הישיבה" להיות המילה המחליפה מכיוון התאמתה להקשר של משפט פורמלי ופגישות מקצועיות במילים אחרות שמרה את המשמעות והתחביר המקוריים, ולכן הצלחנו במשימה במילה הזו.

3. הוועדה: המטרה למצוא מילה ששומרת על הקשר של גוף מקצועי בהתחלה קיבלנו מילים כמו "הקואליציה", "הנשיאות" השתמשנו במילים כמו "ועידה", "אסיפה", "רשות", שמייצגות קבוצות מאורגנות עם משימה דומה לזו של ועדה. כמו כן לא השתמשנו במילים שליליות כדי למנוע הגבלת חיפוש תוצאות מגוננות, המילים שהתקבלו היו: ההסתדרות(0.9076) , העירייה(0.8857) , המועצה(0.8761) , ומתוך תוצאות אלה בחרנו ב" המועצה" משום שהיא מתארת גוף דומה מאוד לוועדה מבחינת התפקיד וההקשר, ולכן הצלחנו במשימה במילה הזו.

4. אני: המילה השניה הכי קשה לפי דעתי, המטרה הייתה למצוא מילה שמתארת זהות אישית או מתאימה מבחינה תחבירית למשפט, כמו "אנוכי" או מילים דומות. בהתחלה נסינו למצוא את אנוכי קיבלנו תוצאות כמו "לשמוע", "דברי", "שוב", אשר לא התאימו כלל למשמעות או להקשר. לאחר מכן ניסינו להוסיף מילים חיוביות כמו "עצמי", הצלחנו לשפר את ההקשר למילים כמו דברי (0.8439), "הייתי" (0.8512) ו-"לשמוע" (0.8525). לבסוף בחרנו ב-"הייתי" להיות המילה המחליפה, משום שהיא שמרה על ההקשר התחבירי של המשפט. התוצאה הייתה סבירה, ולכן הצלחנו במשימה באופן חלקי במילה זו.

5. ההסכם: המטרה היא למצוא מילה חלופית שמתארת מסמך רשמי או פורמלי בהתחלה, קיבלנו מילים כמו "הפשע", "הסכסוך" ו"הניסיון", כאשר נתנו מילים חיוביות כמו

"עסקה" שלא בהכרח התאימו להקשר. לכן השתמשנו בגישה רחבה יותר ללא מילים חיוביות או שליליות, כדי לתת למודל גמישות בחיפוש וקיבלנו את התוצאות הבאות: "הדוח" (0.9117), "הוויכוח" (0.9045) ו-"המסמך(0.8981) ". מתוך התוצאות הבאת בחרנו ב "דוח" להיות המילה המחליפה מכיוון שהיא שומרת על ההקשר המקורי של משפט פורמלי, פורמליות ותיעוד רשמי. המילה השתלבה היטב במשפט מבחינה תחבירית ומשמעותית, ולכן הצלחנו במשימה במילה זו.

6. בוקר: המטרה הייתה למצוא מילה חלופית שמתארת זמן ביום או חלק מהיממה שתתאים לברכה "בוקר טוב" כמו "יום טוב". בתחילה, קיבלנו תוצאות כמו "לכולם", "ממני " ו"מפורסם", שהיו לא רלוונטיות או שינו את משמעות המשפט באופן לא מתאים. ניסינו להכוון את המודל על ידי הוספת מילים חיוביות כמו "צהריים", "יום" ו"אור", כדי לשפר את הכיוון ולמקד אותו בזמני יום. עם זאת, למרות ההכוונה, התוצאות שקיבלנו הם: "לילה" (0.9236), "וערב" (0.9127) ו "בסבלנות" (0.9107) התוצאה הקרובה ביותר שהתקבלה הייתה "לילה". למרות שהיא לא מתאימה לחלוטין להקשר של "בוקר טוב", היא שומרת על התחביר התקין של המשפט, ולכן הצלחנו במשימה במילה חלקית.

7. פותח: המטרה הייתה למצוא מילה חלופית שמתארת פתיחה של פעולה באופן שמותאם להקשר של פתיחת או תחילת ישיבה במשפט הנתון, בהתחלה קיבלנו "אוציא" ו "בודק" שהיו לא מתיאמות לניסוח של המשפט, לכן התחלנו עם כיוונים חיוביים למילים כמו "עוצר" כדי לתאר התחלה או פעולה, אך המודל החזיר תוצאות כמו "מפנה (0.9566) ", "אתחיל (0.9374) ", ו"שומע (0.9357)". מתוך התוצאות הללו, בחרנו את "אתחיל" כמילה המחליפה, מכיוון שהיא מתארת פעולה של התחלה ושומרת על התחביר התקין והמשמעות הפורמלית של המשפט. התוצאה הייתה מתאימה להקשר הכללי של המשפט, ולכן הצלחנו במשימה זו.

8. שלום: המילה הכי קשה לפי דעתי הסיבה מכך נובעת שהמילה שלום יכולה תפורש לשתי משמעויות משמעות ראשונה היא ברכה, משמעות שניה היא (peace), קורפוס הכנסת מתיחס במשפטים שלו למשמעות השניה יותר ולכן היה קשה לכוון את המודל למשמעות של ברכה, כפי שהייתה הכוונה במשפט. בהתחלה קיבלנו תוצאות כמו "ירושלמי", "לשאלות", ו"ערפאת", שלא היו מדויקות למשמעות של ברכה ניסינו להוסיף למודל מילים חיוביות כמו "נעים", "ברוכים", "וברכה", ו"וסהלן", וכן מילים שליליות כמו "מלחמה", אך לא הצלחנו להגיע לתוצאה שמתארת ברכה באופן מדויק. התקבלו התוצאות הבאות "בשמו (0.9254)", "חביבי (0.9214)", ו"לפניו (0.9199)" בסופו של דבר בחרנו במילה "חביבי", שהיא מתארת אמנם גישה נעימה אך אינה ברכה מלאה. המילה השתלבה בתחביר של המשפט, אבל לא שמרה על המשמעות המקורית באופן מלא. הצלחנו במשימה באופן חלקי.

9. שמחים: המטרה הייתה למצוא מילה שמתארת מצב רוח חיובי שמתאים להקשר של המשפט. בשלב הראשון קיבלנו תוצאות כמו " משתדלים (0.9513) ", "נשמח (0.9189) ", ו"ממשיכים (0.9120) ", שבהן רק "נשמח" התקרבה מעט למשמעות הרצויה. ולכן בחרנו את "נשמח" כמילה המחליפה. המילה השתלבה היטב בתחביר המשפט וגם השאירה את אותה אותה משמעות של המשפט הראשון. ולכן הצחלנו במשימה במילה הזו.

10. היקר: המטרה הייתה למצוא מילה שמתארת תחושה של הערכה או אהבה, כמו "מוערך" או "אהוב."בהתחלה קיבלנו תוצאות כמו "בחדר", "בלשונו", ו"התיאור" אשר הן לא תואמות להיותם מחליפות למילה היקר. כדי לשפר את התוצאות, הוספנו מילות כיוון חיוביות כמו "מוערך", ומילות שלילה כמו "זול". התקבלו התוצאות הבאות " לנגר (0.8187)", " ואהובינו (0.7917)", ו" מאגף (0.7878)" בסופו של דבר בחרנו במילה " ואהובינו ", משום שהיא מצליחה לשמור על הקשר רגשי של המשפט. הצלחנו במשימה זו ברמה טובה.

11. קידום: כאשר עבדנו על המילה הזאת היה קשה לנו למצוא מילה מחליפה, המטרה הייתה למצוא מילה שמייצגת התקדמות, הצלחה או שיפור, כמו "עלייה" או "שיפור." בשלב הראשון קיבלנו תוצאות כמו "כיבוי ", "האוויר ", ו"לכלכלה", שלא התאימו למשמעות. לאחר מכן, ניסינו לכוון את המודל עם מילים חיוביות כמו "עצום", "עלייה", ו"שיפור." וקיבלנו את והתוצאות הבאות: "צמצום (0.9789) ", " תשתיות (0.9767) ", ו" לכלכלה (0.9746) " בחרנו ב"צמצום" להיות המילה המחליפה למרות שהיא נותנת את המשמעות הנגדית, הסיבה מכך נובעת כי אחרי הרבה נסיונות (נסינו גם הרבה מילים שונות חיוביות כמו שכר,חברה..., וגם הרבה מילים שלילות כמו ירידה) לא יכולתי למצוא מילה בעלת משמעות עליה/ירידה אשר היא דקוקית נכונה, לכן הצלחנו במשימה הזאת דקוקית.

12. מניעה: המטרה הייתה למצוא מילה שמשתלבת בהקשר של "אין מניעה להמשיך לעסוק בנושא", ומשמעותה רשות או אישור. בהתחלה קיבלנו תוצאות כמו "אומץ ", "שקיפות", ו"תחרות ", שלא תאמו את המשמעות הנדרשת. לכן, הוספנו מילות כיוון חיוביות כמו "הרשאה", "אישורים", ו"היתר", כדי להנחות את המודל לתשובות מתאימות יותר. בסופו של דבר, הצלחנו לקבל תוצאות כמו " בחירה (0.9597) ", " שקיפות (0.9514) ", ו"הגבלה (0.9501) ". בחרנו את "הגבלה" כמילה המחליפה, כי היא משתלבת היטב במשפט ומשמרת את התחביר אך משנה את המשמעות להיפך. אחרי הרבה נסינות של כיוון ע"י מילים חיובים ושלילים מצאנו מילים כמו מנוס שתאומות דקוקית אך משנות את המשמעות ולכן לא מצאנו מילה משאירה את אותה משמעות אך הצלחנו במשימה הזאת דקוקית.

1. האם המילים הכי קרובות שקיבלתם בסעיף א' תואמות את הציפיות שלכם? הסבירו. גם אם תאמו לציפיות וגם אם לא, נסו להסביר מדוע זה עבד או לא עבד טוב.

המילים שהתקבלו תואמות חלקית את הציפיות. למשל, עבור "ישראל" קיבלנו מילים כמו "פלשתין" ו"מדינתי" שמשקפות הקשרים מדיניים וביטחוניים, מה שמראה שהמודל למד על נושאים רלוונטיים מהקורפוס. עבור "גברת", קיבלנו שמות פרטיים כמו "דליה" ו"איילת", שמדגישים הקשר פורמלי. לעומת זאת, עבור מילים כמו "מים" ו"רשות", קיבלנו תוצאות פחות קשורות, כמו "הטבות" ו"תפילין", כנראה בגלל מיעוט הקשרים או שימושים מגוונים בקורפוס. לכן, כאשר המילה נפוצה ובעלת הקשר ברור בקורפוס, המודל מצליח, אך במילים כלליות או נדירות יש פערים.

2. אם ניקח שתי מילים שנחשבות להפכים, האם היינו מצפים שהמרחק בין שני וקטורי המילים שלהן יהיה קצר או ארוך? הסבירו

המרחק בין וקטורי המילים שלהן יהיה קצר יחסית, למרות שהן הפכים במשמעות. הסיבה לכך היא שהמודל מבוסס על ההקשרים שבהם מילים מופיעות בקורפוס, ולא על המשמעות המילולית שלהן.

מילים הפוכות נוטות להופיע באותם סוגי הקשרים, כמו משפטים שמדברים על רגשות או משקל. כך המודל ילמד לשקף אותן למרחב דומה. עם זאת, המשמעות השונה ביניהן יכולה לבוא לידי ביטוי בכיווני הוקטורים, אך לא באורכם.

3. מצאו שלושה זוגות של מילים שנחשבות להפכים הקיימות בקורפוס שלנו ובדקו את המרחק ביניהן. האם הציפייה שלכם מסעיף 2 מתקיימת עבורן עם המודל שבניתם?

לקחנו את שלושת ההפכים עם הדימון הבאים: "אהבה ושנאה- (0.9353) ", " גדול וקטן- (0.7016)"

ו" קל וכבד- (0.8418) " ולכן המרחק בין כל שתי הפכים הינו: 0.2984,0.0647 ו 0.1582 בהתאמה. תוצאות אלה מאשרות שהמודל לוכד את הדמיון ההקשרי בין מילים הפוכות, כפי שציפינו, למרות המשמעות הניגודית שלהן, הן ממוקמות קרוב במרחב הווקטורי.

4. האם המשפטים הכי קרובים בסעיף ג' תאמו לציפיות שלכם? הסבירו. גם אם תאמו לציפיות וגם אם לא, נסו להסביר מדוע זה עבד או לא עבד טוב.

המשפטים שנמצאו כקרובים ביותר בסעיף ג' תואמים לציפיות באופן חלקי. בחלק מהמקרים, לדוגמה, "מה עמדתך בנוגע ל41א ?" והמשפט הדומה לו "מה העלתה החקירה ?", המודל הצליח למצוא משפטים דומים מבחינת מבנה, מה שמדגיש את הבנתו את הקשר התחבירי. עם זאת, במשפטים כמו "אנחנו חיים במדינה שבה צריכים להיות קשובים ורגישים האחד לשני, ולא ייקוב הדין את ההר" והמשפט שנמצא כדומה לו, "אנחנו לא יכולים לתת יד לשכתוב ההיסטוריה ולהלבנה של סייעני הנאצים במדינות," הדמיון נראה חלש יותר, משום שהמשפטים נוגעים לנושאים שונים למרות הדמיון התחבירי. ההצלחות של המודל נובעות מהיכולת שלו לזהות הקשרים תחביריים ומבניים דומים. כמו כן כישלונתיו נובעות ממגבליות של הקורפוס שהזכרנו אותם למעלה כמו כן האופי של משפטים בקורפוס הכנסת, המשלב שפה פורמלית ונושאים מגוונים, מקשה על המודל לזהות דמיונות.

חלק ג:

קודם כל נכין טקסט בעברית ומנקים אותו כך שיישארו רק מילים באותיות עבריות.

נטעין או נאמן (בגרסה אחרת) מודל Word2Vec כדי לתרגם מילים לווקטורים.

מחשבים לכל משפט וקטור ממוצע של ווקטורים (Embeddings) של המילים שלו.

ממפים דוברים שונים לאותם שמות (כדי לאחד על פי דובר).

מורידים דגימה (Downsampling) של המחלקה שיש לה הרבה יותר דוגמאות, כדי ליצור מאגר נתונים מאוזן.

משתמשים ב־ KNN וב־ 5-Fold Cross Validation כדי לסווג מי הדובר של כל משפט, ומדפיסים דו"ח סיווג מפורט שמציג את תוצאות האלגוריתם ('F1, Precision, Recall) .

התמונה שצירפת מכילה את דו"ח הביצועים של מודל הסיווג, כפי שהתקבל מהפונקציה classification_report של ספריית scikit-learn. הדו"ח מציג מדדים עיקריים עבור כל אחת מהכיתות (הדוברים), וכן מדדים ממוצעים עבור כלל המודל.

1. דיוק כללי:(Accuracy)
המודל הצליח לסווג 79% מהמשפטים בצורה נכונה.

2. תוצאות לכל כיתה:
המודל מצליח באופן די דומה לזהות את שני הדוברים, עם יתרון קל לזיהוי הדובר ריבלין ובואר (Recall)של 82%.)

3. מדדים ממוצעים:
המדדים הממוצעים מעידים על איזון טוב בין הביצועים של הכיתות. אין עדות לחוסר איזון משמעותי בין הדוברים.

גרוב 'א	0.81	0.76	0.78	2159
ויליביר ובואר	0.77	0.82	0.80	2159
accuracy			0.79	4318
macro avg	0.79	0.79	0.79	4318
weighted avg	0.79	0.79	0.79	4318

התוצאה הכוללת היא ביצוע טוב עם דיוק של 79%, מה שמעיד על כך שהמודל למד לזהות בצורה אפקטיבית את שני הדוברים בהתבסס על הנתונים שסופקו.

תשובה לשאלה 1:

המודל הנוכחי עם Word2Vec ו־ KNN הצליח להגיע לתוצאות קרובות מאוד לאלה של תרגיל 3, אך התוצאות הקודמות היו מעט טובות יותר. הסיבה העיקרית לכך היא שהשימוש ב TF-IDF- בתרגיל 3 אפשר למודל להתרכז בתכונות משמעותיות יותר עבור סיווג הדוברים, בעוד ש- Word2Vecעשוי היה להוסיף רעש לתהליך הסיווג.

מדוע הביצועים בתרגיל 3 היו מעט טובים יותר?

1. ייצוג המאפיינים:(Feature Representation)

- בתרגיל 3 השתמשנו בוקטורי TF-IDF עם בחירת תכונות (Feature Selection), מה שאפשר למודל להתמקד במילים ייחודיות לדוברים ולהתעלם ממידע פחות חשוב.

- במודל הנוכחי, המאפיינים מבוססים על ממוצע וקטורי .Word2Vec שיטה זו עשויה לטשטש את ההבדלים בין דוברים, כיוון שהיא משקללת את כל המילים במשפט באופן שווה.

2. מרחק קוסינוס:

- במודל הנוכחי השתמשנו ב־KNN עם מרחק קוסינוס .בעוד שמרחק זה יעיל לעבודה עם ,Word2Vec ייתכן שבמקרה זה הוא לא מבדיל מספיק טוב בין המשפטים של הדוברים השונים.

חלק ד:

טוענים קודם את המודל HeBERT ואת המשפטים מהקובץ

מחליפים את המילים החסרות בטוקן .MASK

משתמשים במודל לניבוי המילים החסרות על בסיס הקשרים בטקסט.

משלימים את המשפטים עם המילים שניבאנו ושומרים את התוצאות בקובץ פלט.

התהליך מאפשר לנו לנצל את HeBERT להשלמת טקסטים בעברית באופן יעיל ומדויק.

תשובה לשאלה 1:

במקרים רבים, המשפטים שהמודל השלים נראים הגיוניים מבחינת ההקשר הכללי.
דוגמאות:

.1 קוהרנטיות תוכן

- משפט 1:
  - משפט מקורי" :אתה התעקשת [MASK] לייצר הצבעה על כלום".
  - משפט שהושלם" :אתה התעקשת [לא] לייצר הצבעה על כלום".
  - הערכה :התוצאה "לא" מתאימה להקשר ומשלימה את המשפט באופן הגיוני.
- משפט 6:
  - משפט מקורי" :זה שוב , [MASK] זה עובר כחוט השני בין כל הנקודות הללו"...
  - משפט שהושלם" :זה שוב [קורה] , זה עובר כחוט השני"...
  - הערכה :המילה "קורה" משתלבת היטב במשמעות ובזרימה של המשפט.

.2 קוהרנטיות תחבירית

המשפטים שהושלמו נשארים נכונים מבחינה תחבירית ברוב המקרים.
דוגמאות:

- **משפט 5:**
  - **משפט מקורי" :אנחנו לא נאפשר שילד... הוא לא [MASK] בפעילות שלנו"...**
  - **משפט שהושלם" :אנחנו לא נאפשר שילד... הוא לא [יהיה] בפעילות שלנו".**
  - **הערכה :המילה "יהיה" מתאימה באופן תחבירי להמשך המשפט.**
- **משפט 10:**
  - **משפט מקורי" :זה חתום [MASK] ידי אגף התקציבים"...**
  - **משפט שהושלם" :זה חתום [על] ידי אגף התקציבים"...**
  - **הערכה :המילה "על" נכונה תחבירית ועונה על הציפיות של מבנה המשפט.**

3. דוגמאות של השלמות פחות מדויקות או כלליות

במקרים מסוימים, המודל נוטה להשלים מילים פשוטות או כלליות, שמספקות תחביר אך ייתכן שלא מעשירות את התוכן. דוגמאות:

- **משפט 2:**
  - **משפט מקורי" :אדוני [MASK] אתה שאלת, אני אענה".**
  - **משפט שהושלם" :אדוני [,] אתה שאלת, אני אענה".**
  - **הערכה :המודל בחר בפסיק במקום מילה. למרות שזה נכון תחבירית, ייתכן שציפינו למילה עם משמעות רבה יותר.**
- **משפט 3:**
  - **משפט מקורי" :הוראת השעה התייחסה לתקינה של מספר הילדים".[MASK]**
  - **משפט שהושלם" :הוראת השעה התייחסה לתקינה של מספר הילדים."[.]**
  - **הערכה :השלמת נקודה היא נכונה תחבירית אך אינה מוסיפה משמעות תוכנית.**

4. התאמה להקשר

במקרים שבהם יש הקשר ברור במשפט, המודל מצליח להציע מילים שמתאימות באופן משמעותי:

- **משפט 8:**
  - **משפט מקורי [MASK]" :תעשה להם את הריכוז"?**
  - **משפט שהושלם" :[איך] תעשה להם את הריכוז"?**
  - **הערכה :המילה "איך" תואמת להקשר ומובנת היטב.**
- **משפט 7:**

- משפט מקורי ,[MASK]" :אנחנו רוצים לשמוע את מוטי ששון".
- משפט שהושלם" :[לכן], אנחנו רוצים לשמוע את מוטי ששון".
- הערכה :השלמת המילה "לכן" תורמת למשמעות המשפט ומחזקת את הקוהרנטיות

באופן כללי, המשפטים שהושלמו על ידי המודל הגיוניים מבחינה תוכנית ותחבירית ברוב המקרים:

1. השלמות המילים לרוב מתאימות להקשר ולמשמעות של המשפט.
2. התחביר נשמר בצורה נכונה בכל המשפטים.
3. במקרים מסוימים (כמו פסיק או נקודה), המודל מספק השלמות נכונות אך לא בהכרח "עשירות תוכן."

לסיכום :המודל ביצע עבודה טובה במתן השלמות הגיוניות מבחינת התוכן והתחביר, עם מספר מקרים שבהם ניתן לשפר את התאמת התוכן המלא.

תשובה לשאלה 2:

המודל סיפק השלמות קרובות למילים החסרות האמיתיות בחלק מהמקרים, במיוחד כאשר הייתה זיקה חזקה להקשר התחבירי או המילולי במשפט (למשל משפטים עם מילים כמו "לא", "על", ו-"יהיה"). עם זאת, במקרים מסוימים, ההשלמות של המודל לא תאמו את המילים האמיתיות מבחינת משמעות (כגון החלפת "בכוח" ב-"לא" או "מראש" ב-"קורה"), למרות שעדיין יצרו משפטים קוהרנטיים והגיוניים. בסך הכול, המודל הפגין יכולת טובה בזיהוי מילים חסרות בהקשרים פשוטים וברורים, אך התקשה במשפטים מורכבים יותר שבהם נדרשת הבנה עמוקה יותר של הקשר ותוכן.

תשובה לשאלה 3:

השוואה בין התוצאות בתרגיל בית 2 לתוצאות הנוכחיות, ניתן לראות שחל שיפור משמעותי בתוצאות הנוכחיות מבחינת דיוק המילים החסרות והתאמתן להקשר. בתרגיל 2, המודל נטה להשלים בעיקר סימני פיסוק (כגון פסיקים ונקודות) או מילים כלליות כמו "אני", שהן פחות משמעותיות להשלמת המשפטים מבחינת תוכן. לעומת זאת, בתרגיל הנוכחי, השימוש במודל HeBERT הוביל להשלמות מדויקות וקוהרנטיות יותר מבחינה סמנטית, כמו "לא", "איך" ו-"לכן", שהשתלבו היטב במשפטים ושיפרו את משמעותם ואת קוהרנטיות התוכן. בעוד שבתרגיל 2 ההשלמות נראו שטחיות ומבוססות על תדירות מילים פשוטות, התרגיל הנוכחי מצביע על הבנה עמוקה יותר של ההקשר הלשוני במשפטים.

תשובה לשאלה 4:

כן, ישנם משפטים שבהם המודל עבד פחות טוב, בעיקר במשפטים בעלי הקשרים מורכבים או במקומות שבהם המילה החסרה דרשה הבנה עמוקה יותר של הקשר הטקסט. לדוגמה, במשפט "זה שוב מראש..." המודל השלים "קורה", מה שנכון חלקית מבחינת המשמעות אך אינו תואם לחלוטין למילה המקורית "מראש". בנוסף, במשפטים שבהם המילה החסרה לא הייתה ברורה באופן חד-משמעי מההקשר, כמו "אדוני [MASK] אתה שאלת...", המודל השלים פסיק, שהיה נכון תחבירית אך חסר משמעות תוכנית עמוקה יותר. הסיבה לכך היא שהמודל HeBERT מתבסס על דפוסים לשוניים כלליים שנלמדו מתוך קורפוסים גדולים, אך ייתכן שאין לו די מידע ספציפי כדי לזהות הקשרים דקים או מילים נדירות המופיעות לעיתים רחוקות.

תשובה לשאלה 5:

למודל יש סיכוי גבוה יותר להיכשל או להחזיר תוצאה פחות טובה במקרים שבהם המילה החסרה תלויה בהבנה עמוקה של ההקשר הסמנטי או בתחביר מסובך, כמו משפטים עם ריבוי משמעויות או שימוש בשפה מטפורית. בנוסף, המודל עשוי להיכשל כאשר המילה החסרה נדירה יחסית או כאשר יש מספר אפשרויות הגיוניות להשלמה, אך ההבדלים ביניהן תלויים בקשרים תרבותיים, היסטוריים או הקשריים שאינם קיימים במאגר הנתונים שעליו המודל אומן. כמו כן, במשפטים שבהם יש רמזים מעורפלים או חלקיים בלבד למילה החסרה, המודל עשוי לבחור מילים נפוצות או גנריות שאינן תואמות במדויק למשמעות הרצויה. הסיבה לכך נעוצה בכך שהמודל מתבסס על תבניות לשוניות שנלמדו מתוך קורפוסים גדולים, ולא תמיד יש לו יכולת להתמודד עם חריגות או יוצאי דופן.