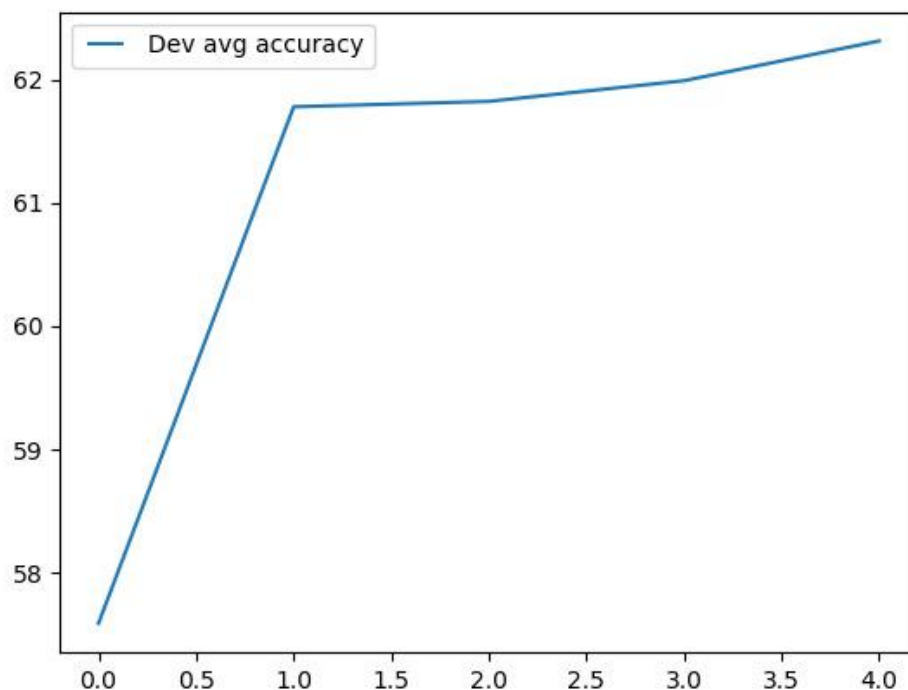
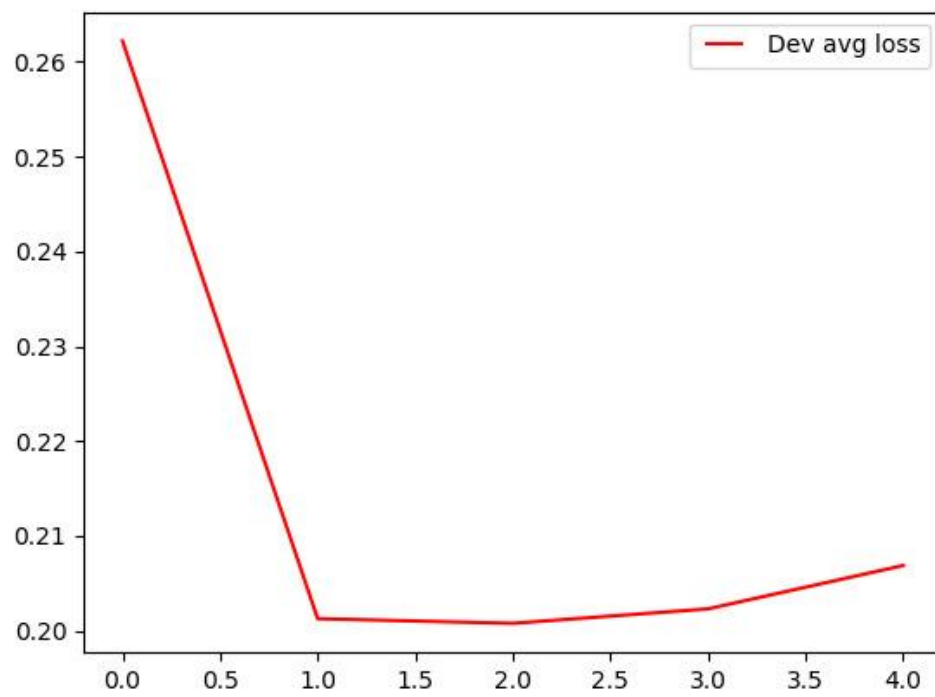


Pos without pre-trained embeddings

היפר פרמטרים (רלוונטים למודל שלנו):

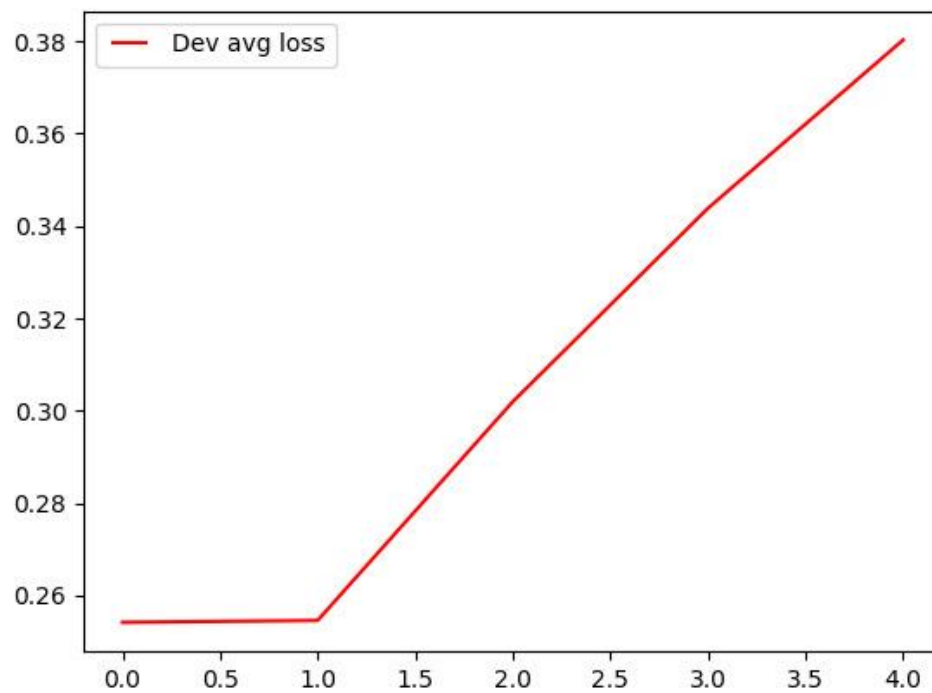
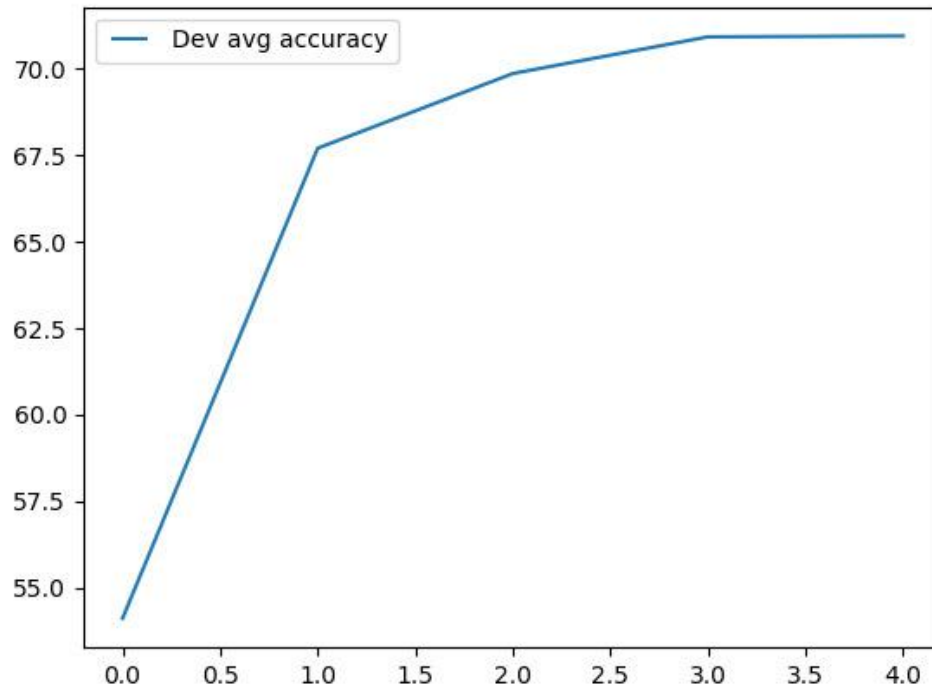
1. Layer size – בחרנו ב100 לאחר שראינו שאנו מקבלים את התוצאות הטובות ביותר עם שכבה זו.
2. Learning rate – על מנת לאמן את המודל בצורה טובה בחרנו במקדם למידה איטי יחסית על מנת להגיע לתוצאות טובות על חשבון זמן הריצה (כי בגלל מקדם הלמידה הנמוך אנו עושים יותר אפוקים). לכן בחרנו במקדם למידה 0.01.
3. Epochs – אנו מבצעים 3 אפוקים על מנת לא להגיע over fitting, מכיוון שהמודל לומד את דוגמאות האימון מהר מאוד.
4. Batch size – בחרנו batch בגודל 1024 על מנת שנעדכן את הדוגמאות עבור מספר סביר של דוגמאות, ולא לאחר כל דוגמה.





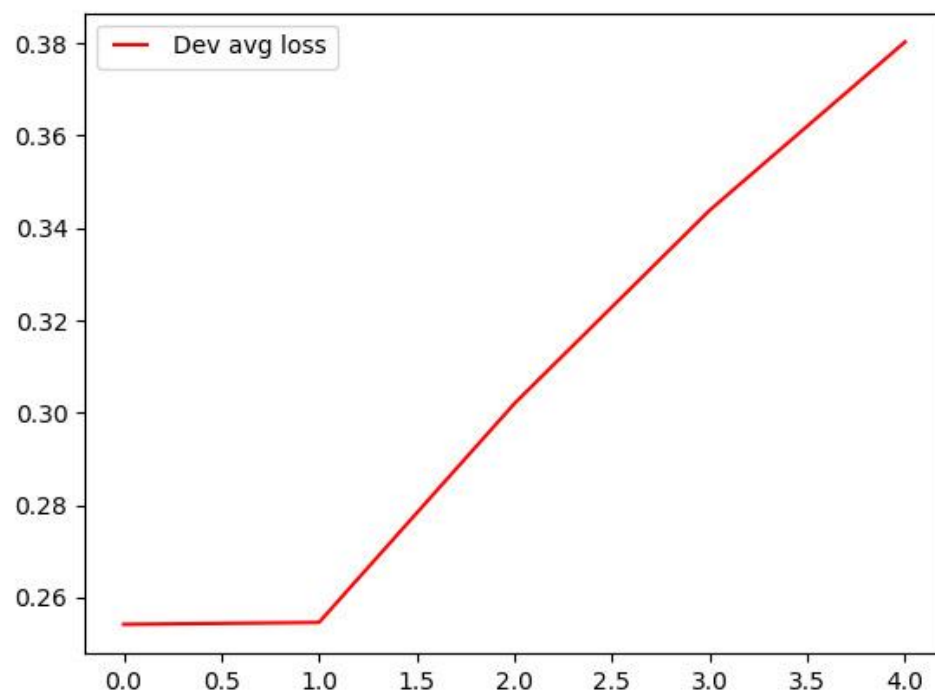
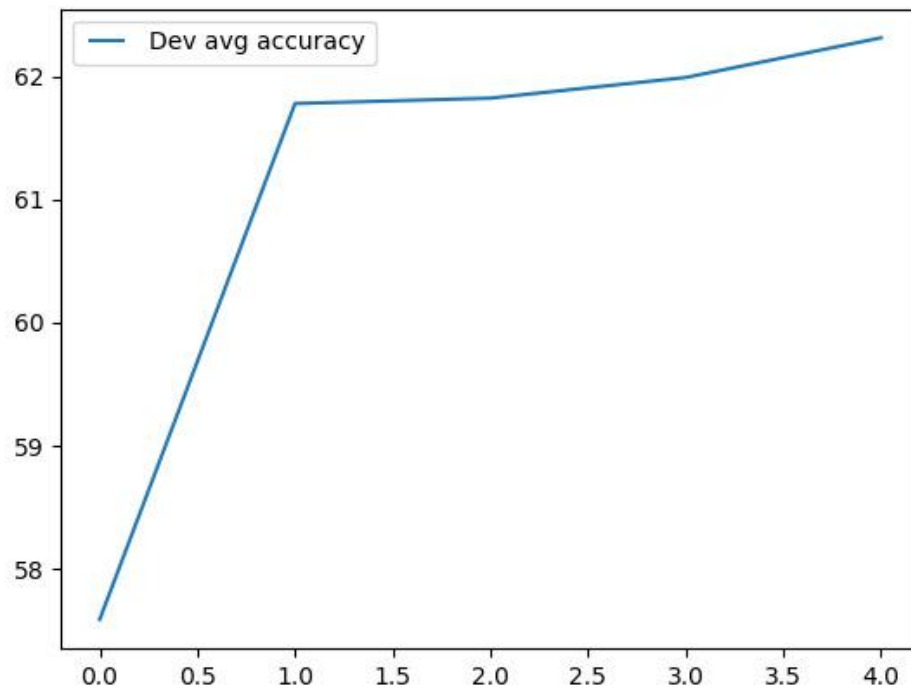
:Pos with pre-trained embeddings

בחרנו באותם פרמטרים מאותם הסיבות.



:Ner without pre-trained embeddings

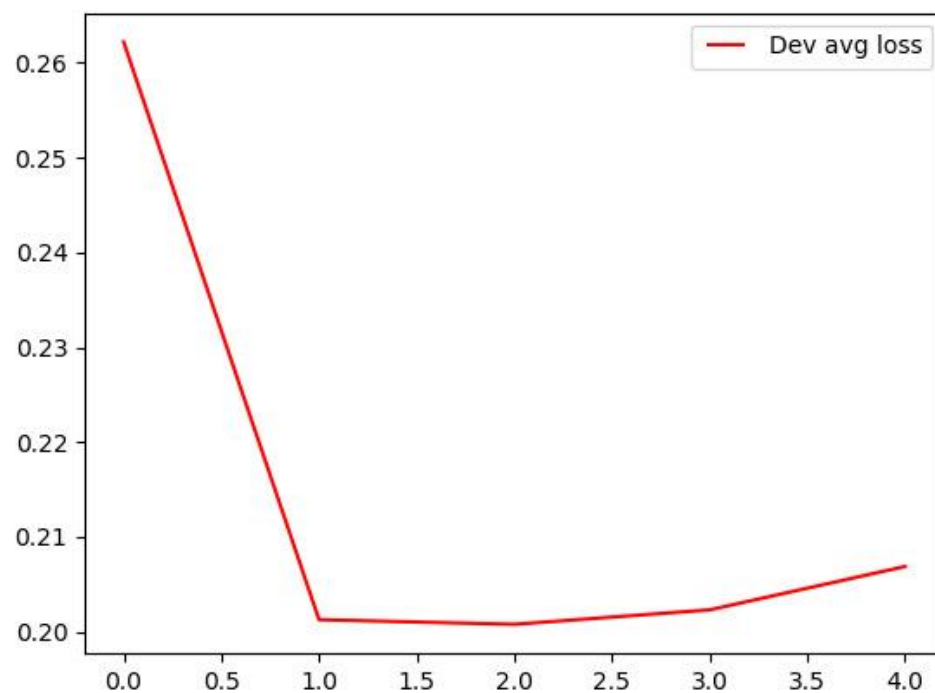
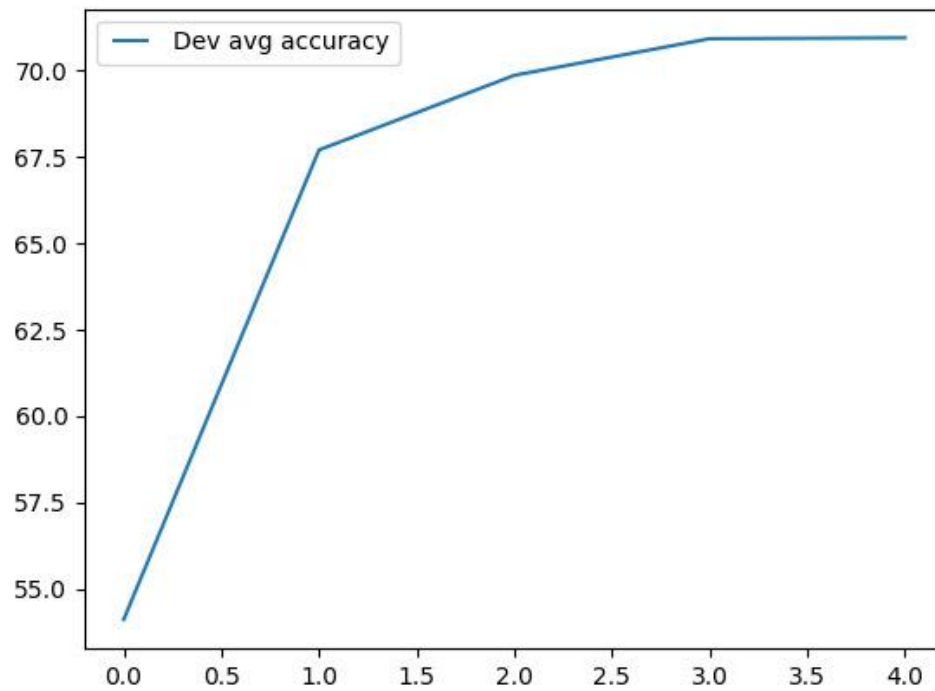
בחרנו באותם פרמטרים מאותם הסיבות.



:Ner with pre-trained embeddings

בחרנו באותם פרמטרים מאותם הסיבות.

להוסיף גרפים



יצרנו 2 מטריצות עבור suffix כן שכל שורה תייצג רישא או סיפא. ניתן לראות כי לאחר שקיבלנו מידע על המילה, בכל זאת קיבלנו תוצאות פחות טובות. ניתן להסביר זאת על ידי כך שהמודל שלנו לא מספיק "חזק" ומעמיק על מנת לנצל את המידע. בנוסף ה sub-words pre-trained יחדיו לא עזרו, כי למרות השימוש בהם התוצאות עדיין לא היו גבוהות יותר. במקרה של pos לא היה שיפור כאשר נעשה שימוש ב sub words pre trained vectors, ואף גרע. לסיכום, כפי שאמרנו, המודל שלנו לא בנוי בצורה טובה ועמוקה מספיק על מנת להשתמש במידע המסופק לו כדי לנצל אותו.