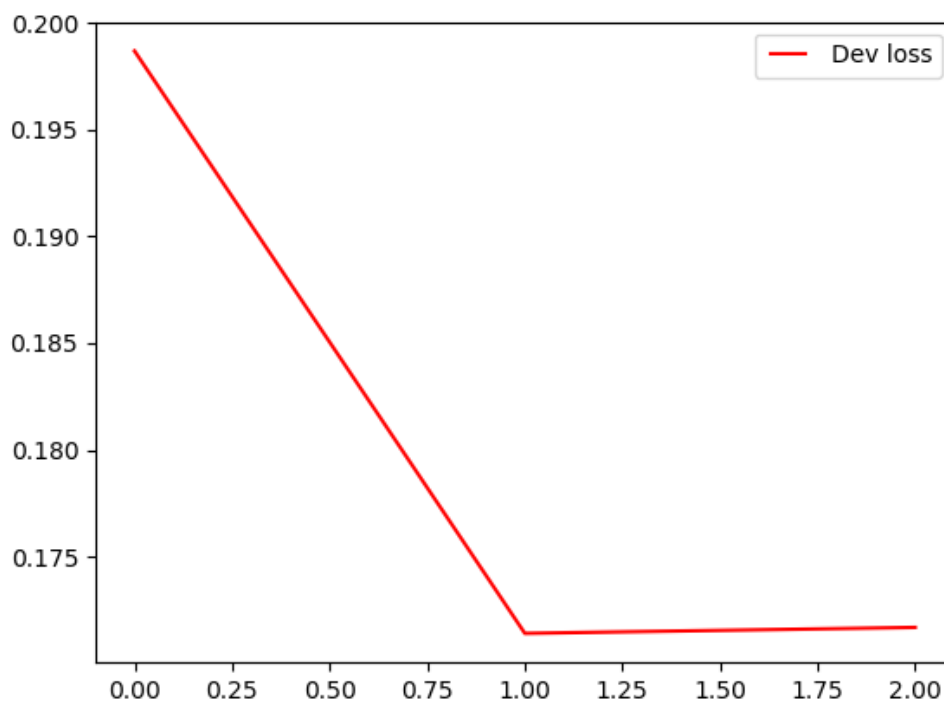
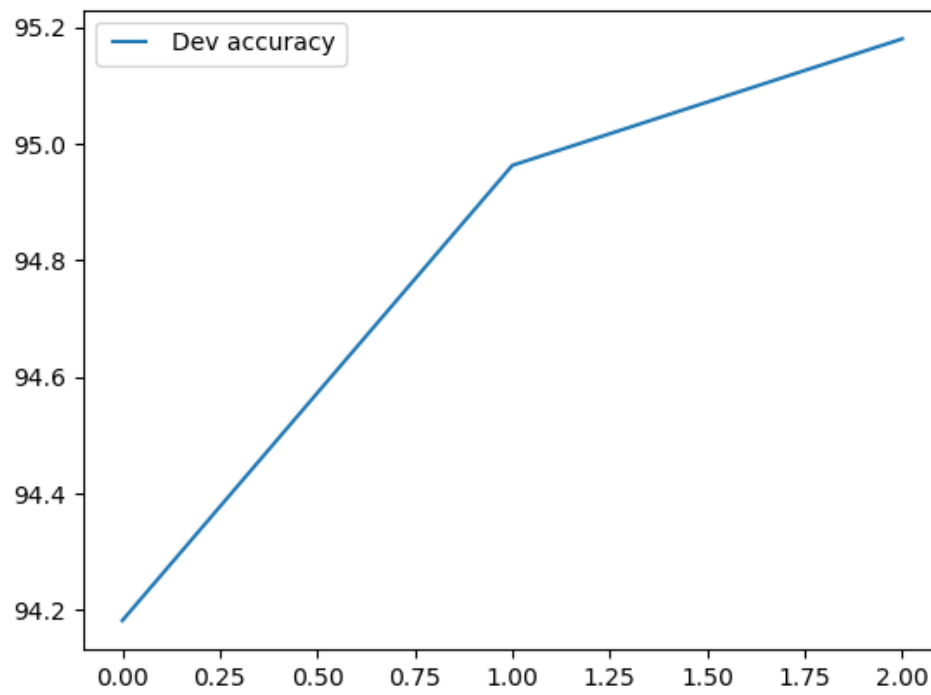


POS:

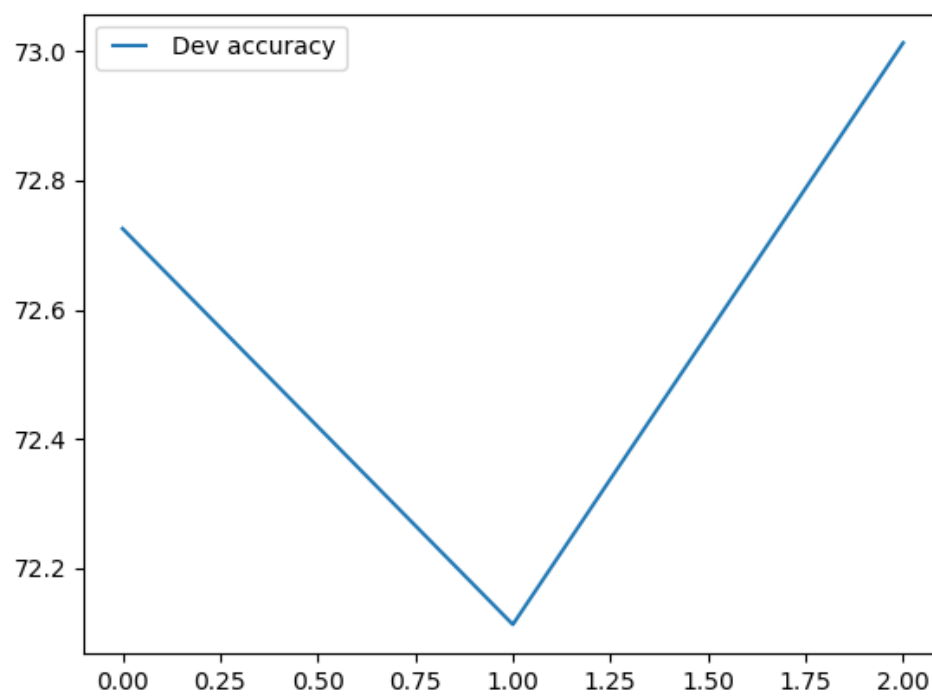
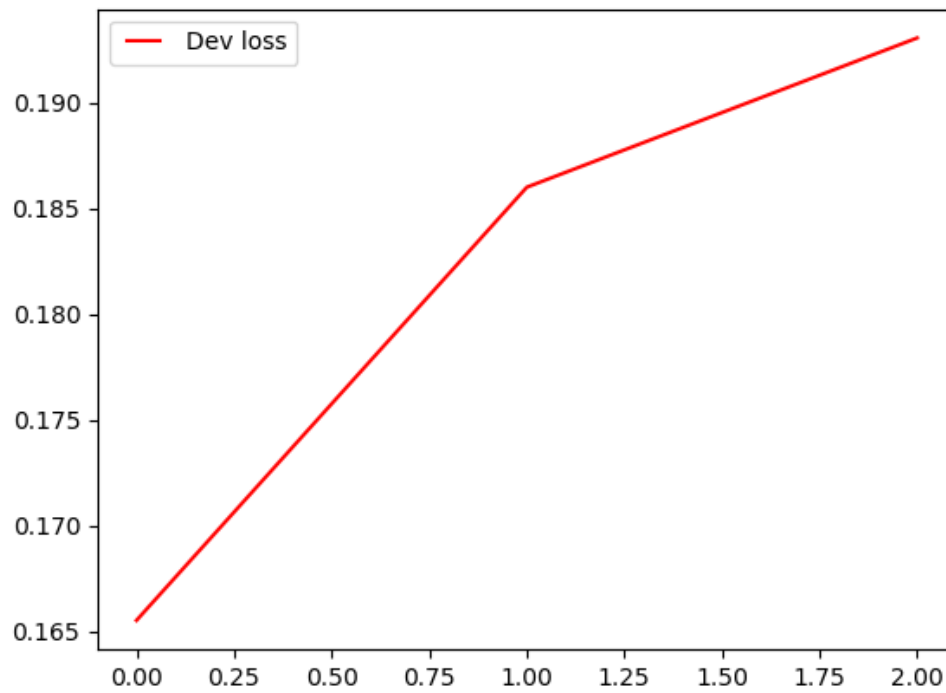
- Layer size – בחרנו ב100 לאחר שראינו שאנו מקבלים את התוצאות הטובות ביותר עם שכבה זו.
- Learning rate – על מנת לאמן את המודל בצורה טובה בחרנו במקדם למידה איטי יחסית על מנת להגיע לתוצאות טובות על חשבון זמן הריצה (כי בגלל מקדם הלמידה הנמוך אנו עושים יותר אפוקים). לכן בחרנו במקדם למידה 0.01.
- Epochs –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





:NER

הפרמטרים דומים, מלבד מקדם הלמידה שהוא 0.05 על מנת שהמודל ילמד את הדוגמאות בצורה מהירה יותר.



Considerations:

1. במודל הקדשנו מילה ייעודית למילים לא מוכרות, וכל מילה לא מוכרת כזאת נחשבת כאילו ראינו את המילה הייעודית. למילה זו כמובן יש התייחסות כמו לכל מילה רגילה והיא מופיעה ב.E.
2. מכיוון שעבור כל מילה אנו לוקחים את 2 המילים לפניה ו2 המילים שאחריה, נוצרת בעיה עבור ה-2 הראשונות וה-2 האחרונות. לכן על מנת לפתור את הבעיה השתמשנו בריפוד, כלומר הוספנו עוד 2 מילים להתחלה ו2 מילים לסוף.