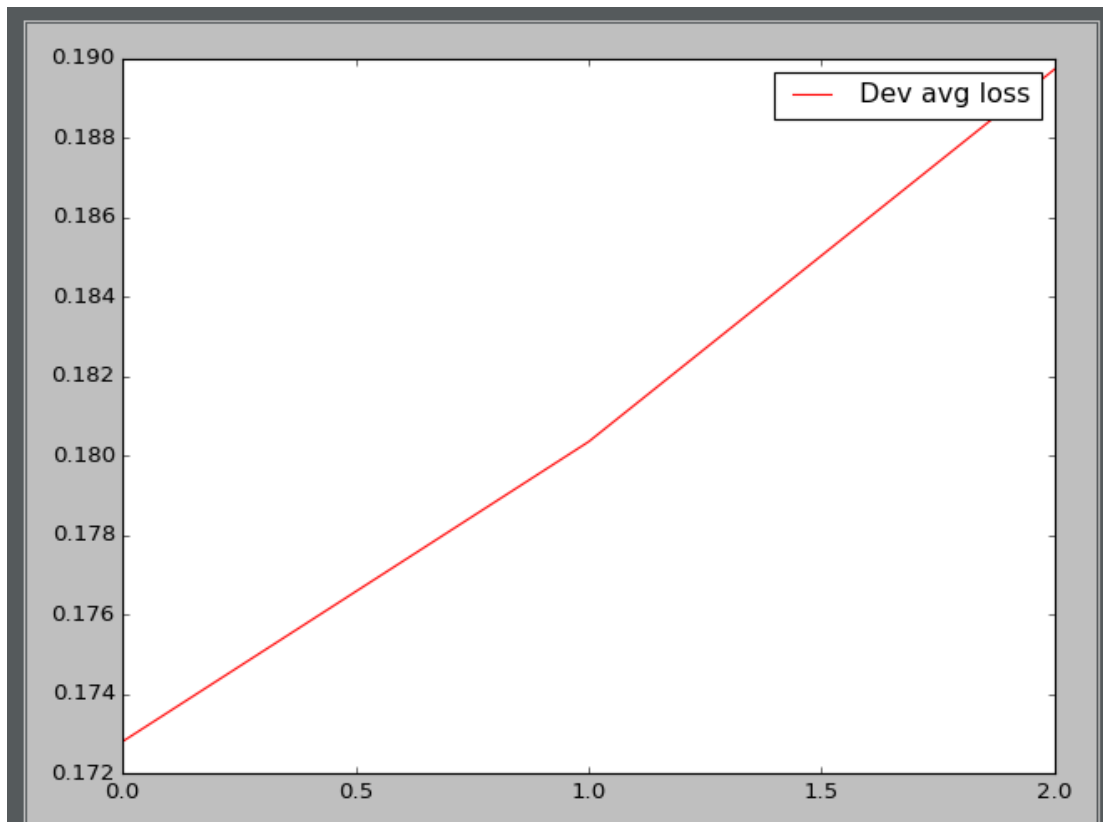
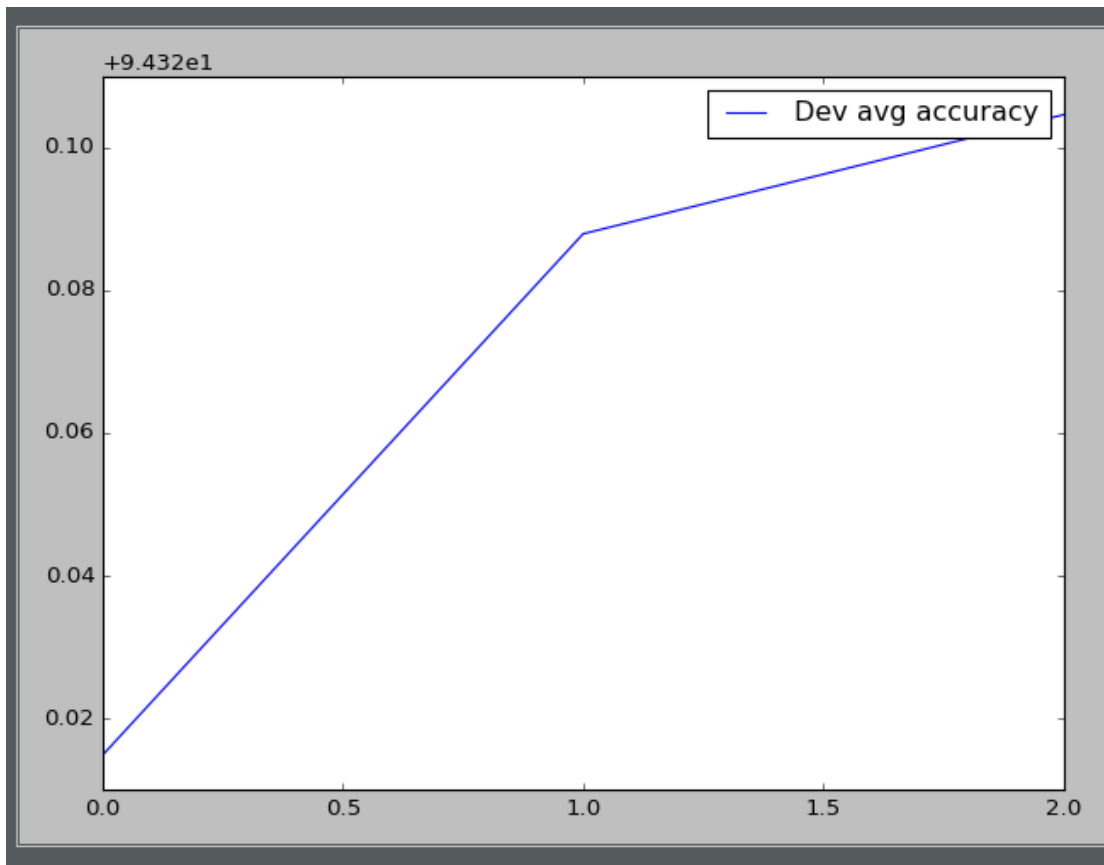


:POS

כ

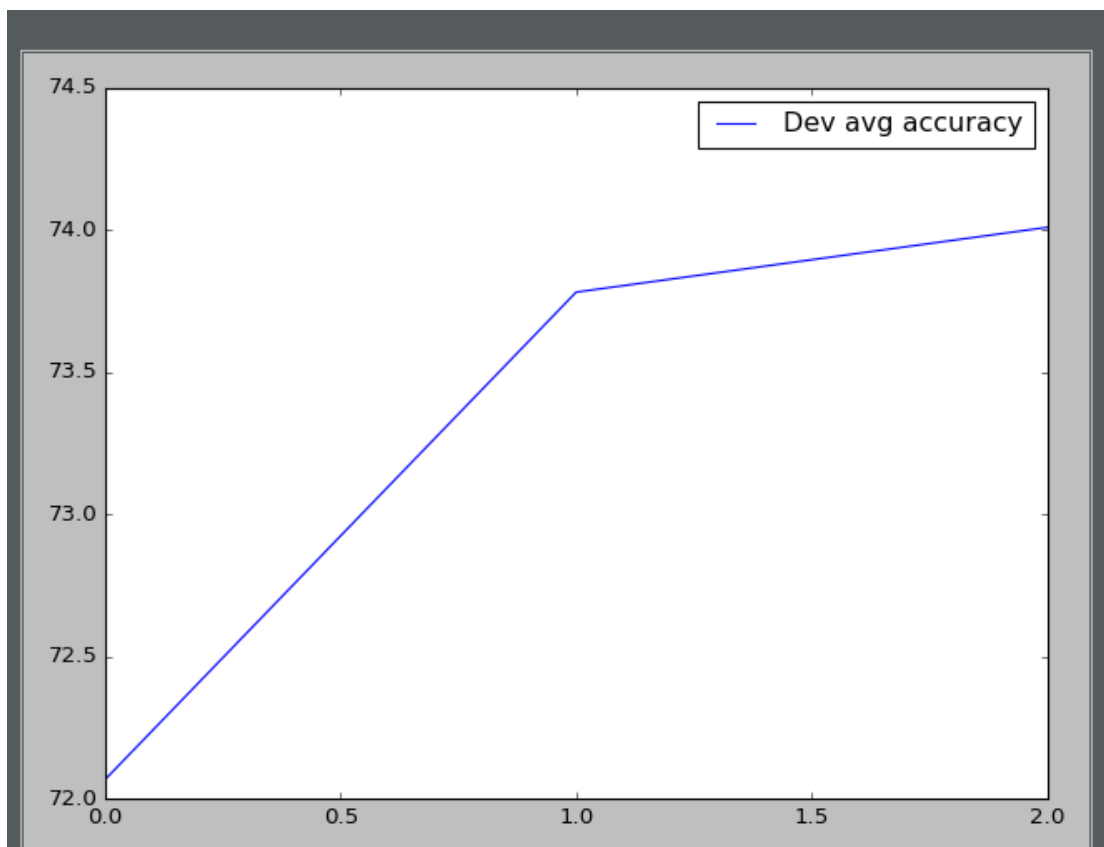
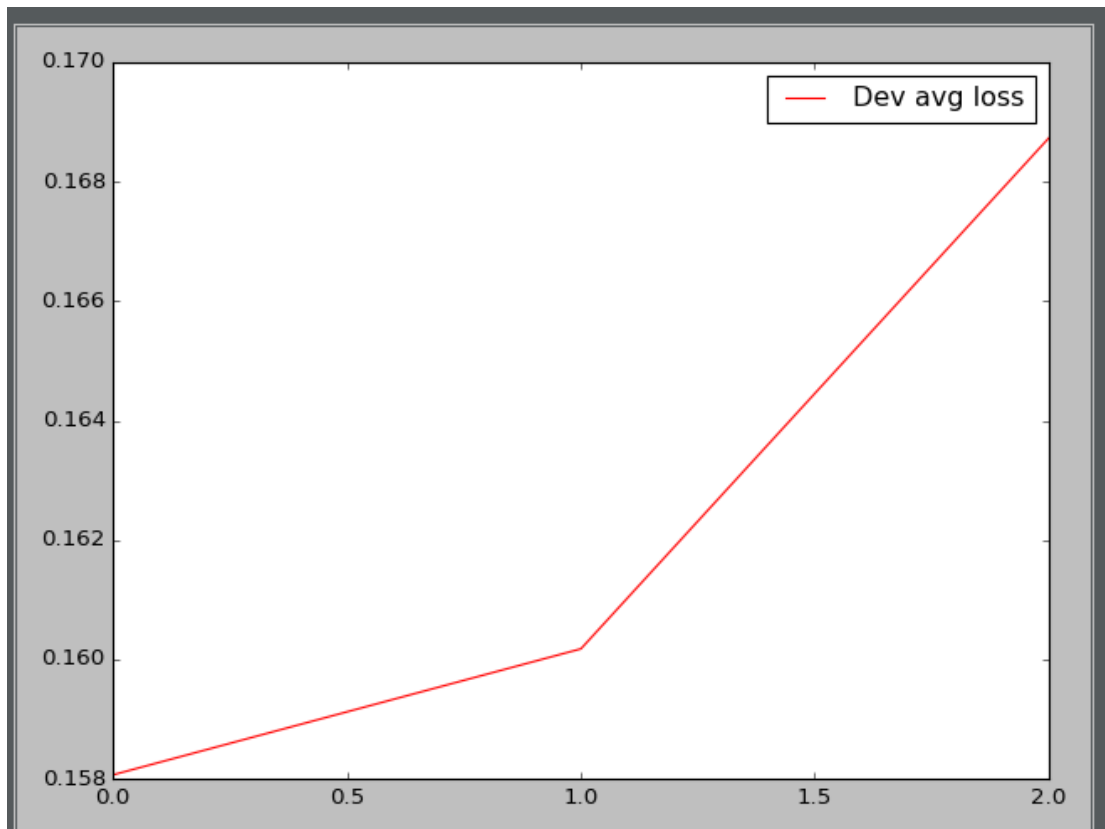
- פונקציית אקטיבציה – השתמשנו בפונקציית אקטיבציה \tanh כפי שהתבקשנו לפי דרישות התרגיל. מטרת השימוש בפונקצייה זו בין השכבות היא לאפשר למודל להיות אי לינארי על מנת שיוכל להתמודד טוב יותר עם בעיות שלא ניתנות להפרדה לינארית.
- Learning rate – על מנת לאמן את המודל בצורה טובה בחרנו במקדם למידה איטי יחסית על 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 – זהו תהליך שמטרתו הינה להגביר את היציבות של הרשת ולכן עלינו לנרמל את הנתונים הנכנסים לכל שכבה על ידי הפחתת הממוצע וחלוקה בסטיית התקן. בחרנו batch בגודל 1024 על מנת שנעדכן את הדוגמאות עבור מספר סביר של דוגמאות, ולא לאחר כל דוגמה.
- גודל השכבה הנסתרת – הגדרנו אותה להיות ממימד 100 כפי שהתבקשנו בתרגיל.





:Ner

הפרמטרים במקרה זה יהיו זהים למודל הקודם



תשובות לשאלות:

- בתרגיל זה היה עלינו להתמודד עם המצב בו אוצר המילים שלנו היה רשום בצורה של אותיות קטנות באנגלית. העיקר היה להתמודד עם מצב בו יהיו 2 מילים או יותר שהן זהות אך נכתבו אופן שונה מבחינת אותיות גדולות וקטנות. לשם כך נעזרנו במתודה בפייטו שנקראת lower שבעזרתה שינינו את הייצוג של כל מילה שהוכנה למילון שלנו לייצוג של מילים מאותיות קטנות וכך שמרנו על ייצוג אחיד ונמנענו מהכנסה של מילים עם ייצוג שונה אבל זהות בערכן למילון שלנו.
- בתרגיל זה היה עלינו להתמודד עם המצב בו אוצר המילים שלנו היה רשום בצורה של אותיות קטנות באנגלית. העיקר היה להתמודד עם מצב בו יהיו 2 מילים או יותר שהן זהות אך נכתבו אופן שונה מבחינת אותיות גדולות וקטנות. לשם כך נעזרנו במתודה בפייטו שנקראת lower שבעזרתה שינינו את הייצוג של כל מילה שהוכנה למילון שלנו לייצוג של מילים מאותיות קטנות וכך שמרנו על ייצוג אחיד ונמנענו מהכנסה של מילים עם ייצוג שונה אבל זהות בערכן למילון שלנו.
- לאחר שאימנו את שני המודלים שלנו על שתי הסטי אימון השונים ולאחר שבחרנו את התוצאה הטובה ביותר עבור כל מודל נוכחנו לדעת שלא היה שיפור משמעותי בין המודל הקודם למודל הנוכחי.