**חלק 3:**

שלום בודק יקר,

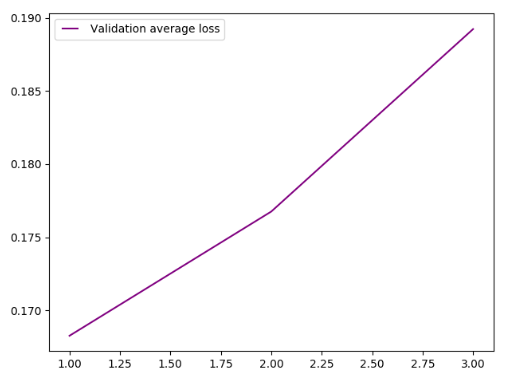
לצורך הכנת הtest.pred בחרנו להפעיל את המודל יחד עם ההיפר פרמטרים הבאים:

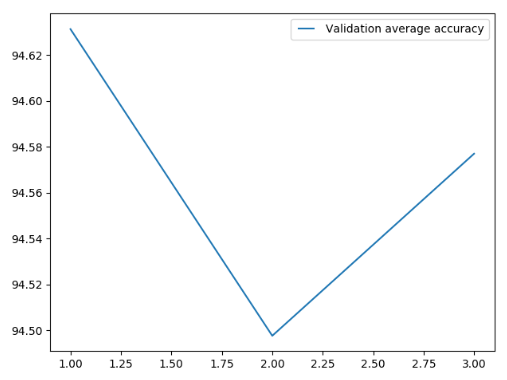
**מודל pos:**

**היפר פרמטרים (רלוונטים למודל שלנו):**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Layer size** | **Learning rate** | **Epochs** | **גודל batch** | **פונקציית אקטיבציה** | **oprimizer** |
| 100 | 0.01 | 3 | 1024 | tanh | adam |

**גרפים רלוונטים למודל זה:**





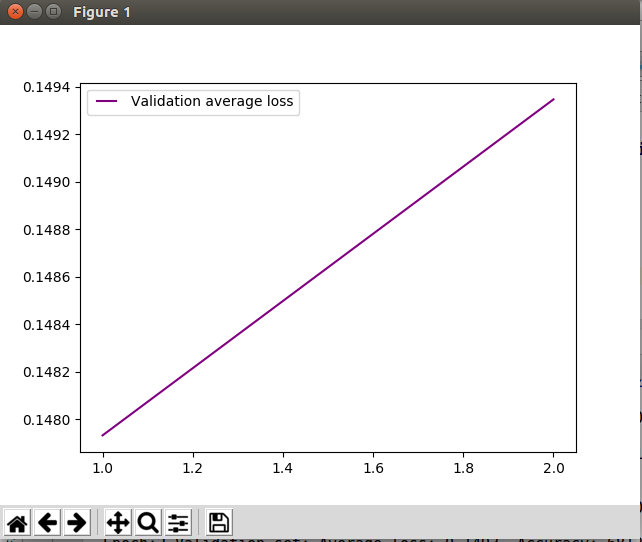
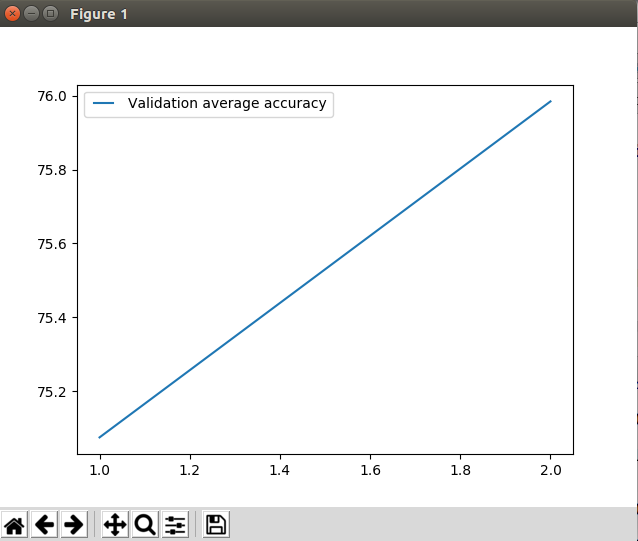
**מודל ner:**

**היפר פרמטרים (רלוונטים למודל שלנו):**

נציג הפעם את הנתונים בצורת טבלה כדי להימנע מחזרות מיותרות:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Layer size** | **Learning rate** | **Epochs** | **גודל batch** | **פונקציית אקטיבציה** | **oprimizer** |
| 100 | 0.01 | 3 | 1024 | tanh | adam |

**גרפים רלוונטים למודל זה:**

**שאלות:**

1. יש  להסביר את הלוגיקה בשימוש בpre trained vectors:

1. איך התמודדנו עם העובדה שהvocabulary עבור הword embeddings הגיע lower case.
2. מה עשינו עם מילים שלא נמצאות בpre trained vocabulary- ייתכנו מילים שנראה בtrain ולא בembedding.
3. האם התוצאות השתפרו בהשוואה למודל במחלק הראשון? אם כן בכמה?

**תשובות:**

1. א. כדי להתמודד עם העובדה שאוצר המילים הגיע בצורה lower case, הפעלנו פונקציית lower על כל מילה שנתקלנו בה במהלך האימון, הוולידציה והטסט, כך שנוכל להתאים אותו לword embedding המתאים. עשינו זאת מתוך כוונה למצוא את ההתאמה הטובה ביותר עבור המילים ולא לפספס word embeddings שנכתבו בצורה שונה מבחינת lower case או upper case.

ב. במידה דומה למה שעשינו בחלק הראשון, מיפינו את כל המילים הלא מוכרות למילה unk, מילה זו קיימת באוצר המילים ויש עבורה word embedding.

1. בחלק הראשון אחוזי ההצלחה על הvalidation עבור מודל הpos היו 94% ועבור מודל הner היו 77%.

בחלק השני, אחוזי ההצלחה על הvalidation עבור מודל הpos היו 95% ועבור מודל הner היו 76%.

אם כן, אנו רואים שהpre-trained word embeddings לא כל כך השפיעו על התוצאות במודלים אלו.