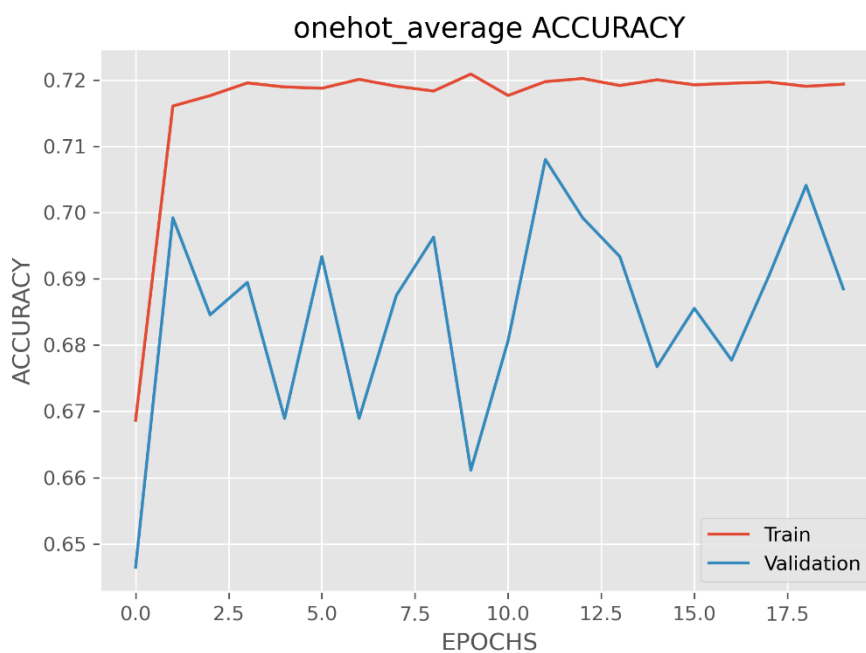
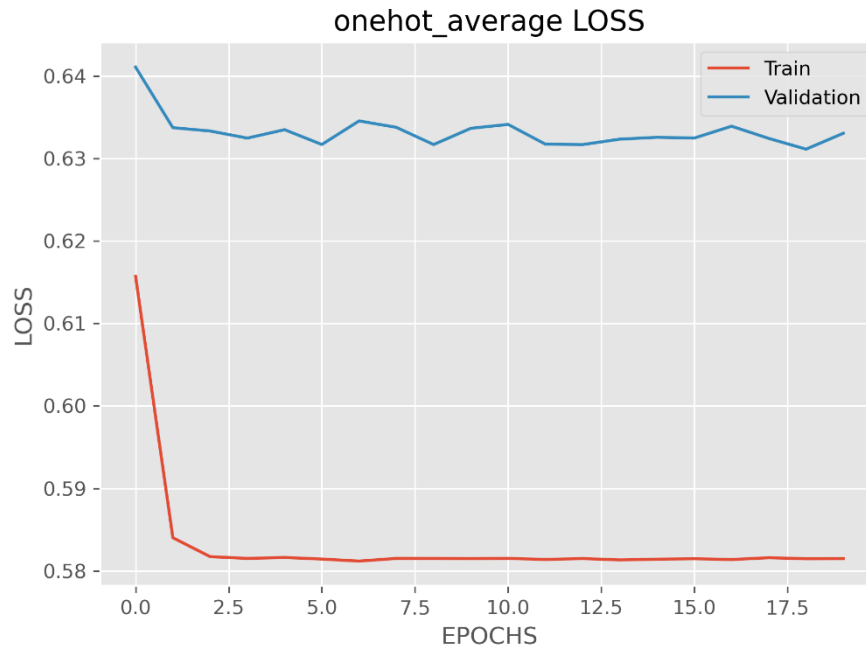


עיבוד שפה טבעית תרגיל 4

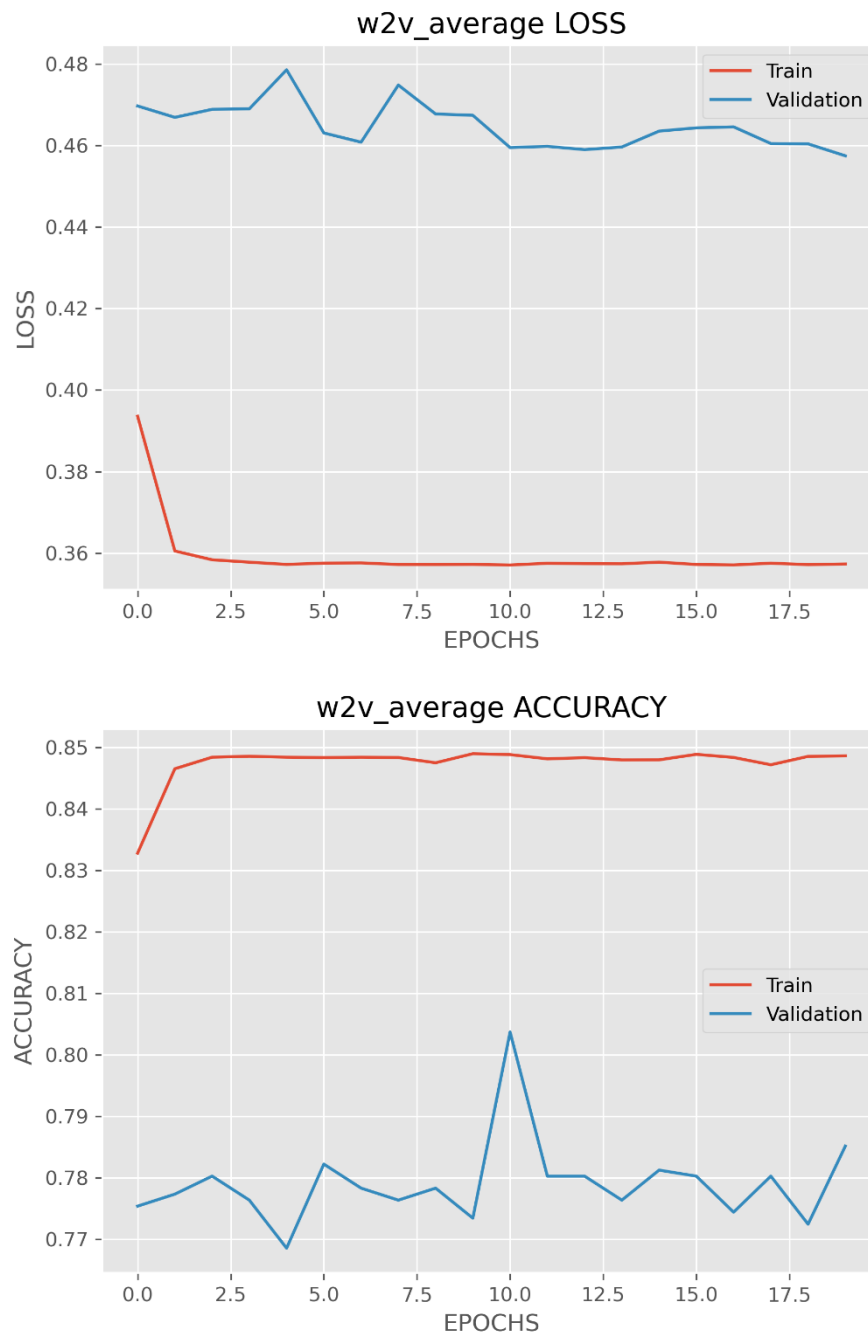
תוצאות אימון המודל הלוג לינארי עם קידוד One Hot :



תוצאות על ה-Test Set :

- דיוק עבור Test Set - 0.6808731808731808
- Loss עבור Test Set - 0.6263510063290596
- דיוק עבור המשפטים המנוגדים - 0.5806451612903226
- דיוק עבור המשפטים על המילים הנדירות - 0.42

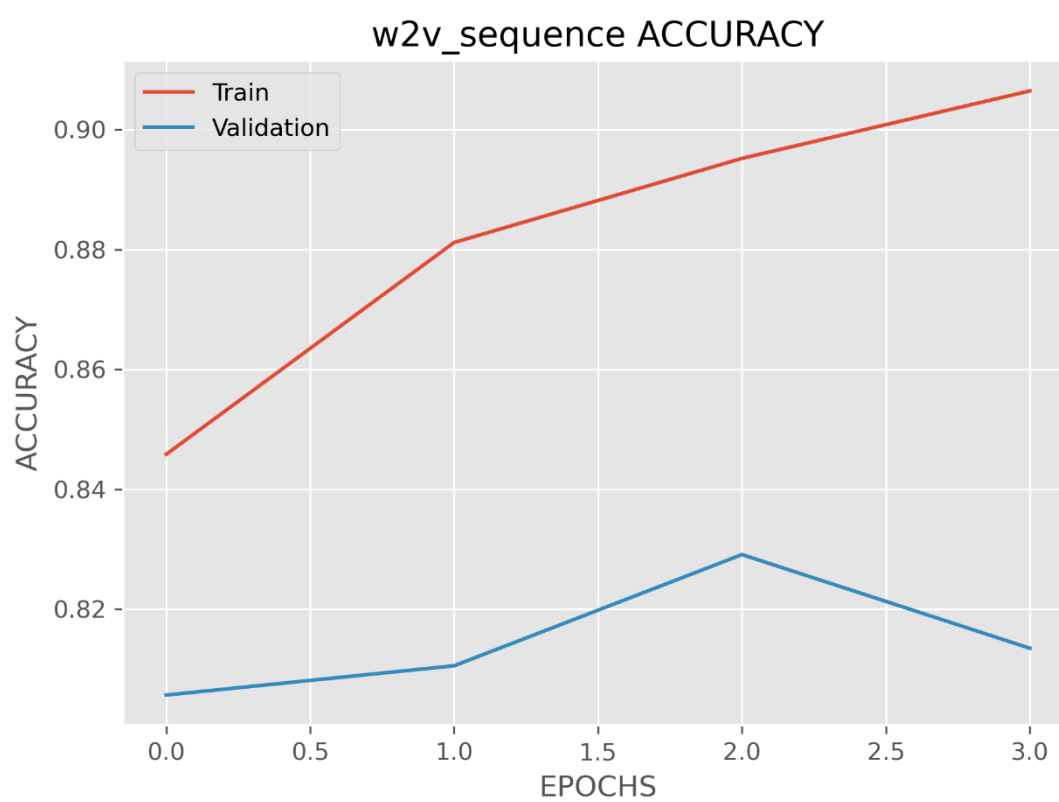
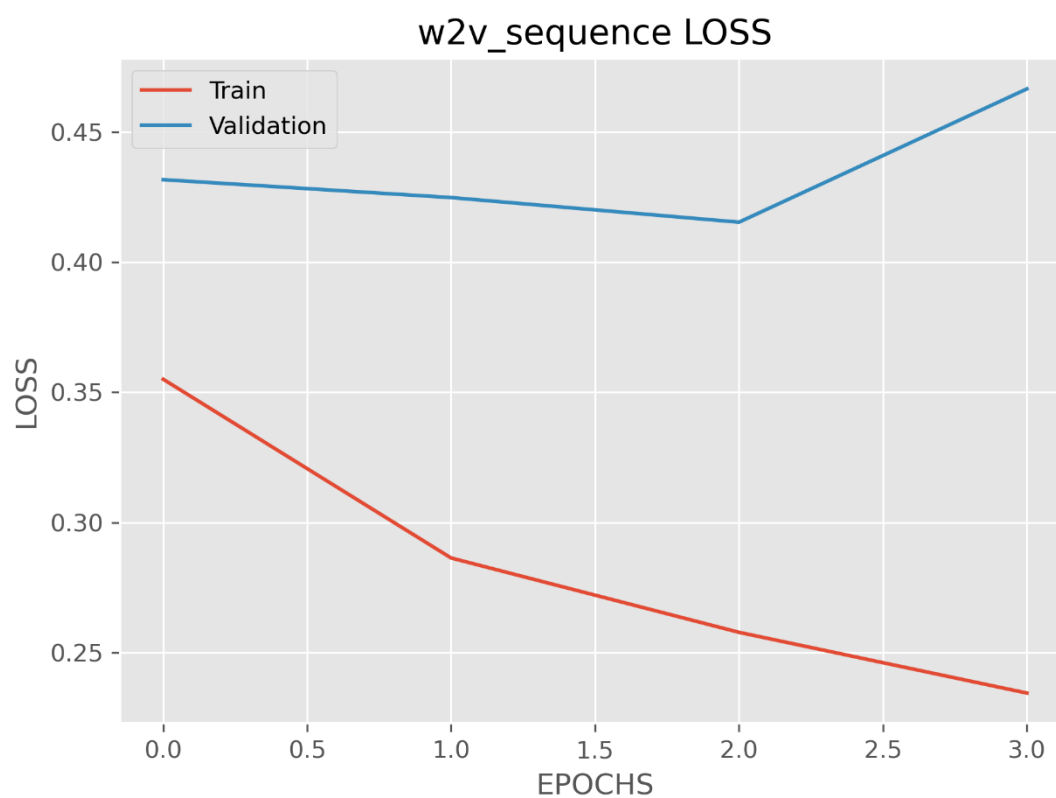
תוצאות אימון המודל הלוג לינארי עם קידוד Word To Vec :



תוצאות על ה-Test Set :

- דיוק עבור Test Set - 0.8284823284823285
- Loss עבור Test Set - 0.3946915939450264
- דיוק עבור המשפטים המנוגדים - 0.6129032258064516
- דיוק עבור המשפטים על המילים הנדירות - 0.82

תוצאות אימון מודל ה-LSTM :



תוצאות על ה-Test Set :

• דיוק עבור Test Set - 0.8513513513513513

- Loss עבור Test Set – 0.3418325331003871
- דיוק עבור המשפטים המנוגדים - 0.7419354838709677
- דיוק עבור המשפטים על המילים הנדירות – 0.76

השוואת המודלים השונים:

1.

| | לוג לינארי פשוט | embedding עם |
|------------|-----------------|--------------|
| Validation | 0.69 | 0.82 |
| Test | 0.68 | 0.78 |

המודל עם ה-embedding נתן תוצאות טובות בהרבה מהמודל הפשוט, לדעתנו זה קרה בגלל שהמודל עם ה-embedding מכיל יותר מידע שקל יותר ללמוד, כמו כן הקידוד בוצע באמצעות מודל שאומן לשם כך, העובדה שהמודל הפשוט מקודד כל מילה ל- one hot vector משרה sparsity משמעותי והופכת את הלמידה לקשה בהרבה.

2.

| | לוג לינארי פשוט | לוג לינארי עם embedding | LSTM |
|------------|-----------------|-------------------------|------|
| Validation | 0.69 | 0.82 | 0.82 |
| Test | 0.68 | 0.78 | 0.85 |

מודל ה-LSTM סיפק את התוצאות הכי טובות על ה-Test Data, האופי של המודל שיועד להסתכל על חלקים רלוונטיים שהוא ראה במשפט במהלך ריצתו מאפשרים לו לקבל החלטות מדויקות יותר, כמו כן במשפטים יש חשיבות גדולה להיסטוריה ולכן הפערים בולטים. נשים לב כי הוספת שכבת ה-dropout גרמה לכך שהתוצאות על ה-test set יהיו אפילו טובות יותר מהתוצאות של ה-Validation Set.

3.

| | לוג לינארי פשוט | לוג לינארי עם embedding | LSTM |
|---------|-----------------|-------------------------|------|
| Rare | 0.42 | 0.82 | 0.76 |
| Negated | 0.58 | 0.61 | 0.74 |

ניתן לראות עבור המשפטים המנוגדים מודל ה-LSTM סיפק תוצאות טובות יותר באופן משמעותי מהמודלים הלוג לינאריים, העובדה הזאת כנראה מוסברת באמצעות האופי של המודל, העובדה שהמודל זוכר את המילים שהוא קרא עד כה במשפט מסייעת לו ככל הנראה להתמודד עם ניגודים במשפט.

מצד שני עבור המילים הנדירות, המודל הלוג לינארי עם ה-embedding הציג תוצאות טובות יותר ממודל ה-LSTM, יכול להיות שהמיוצע של המילים במשפט גרמו למילים הנדירות "להיטמע" במילים מוכרות יותר באופן שאפשר למודל להצליח להתגבר על העובדה שהיא במשפט מילים שהוא לא הכיר.

נבחין כי המודל הלוג לינארי הפשוט סיפק תוצאות גרועות עבור שני ה-subsets ככה"נ מאותן סיבות שגרמו לו להיות פחות טוב כבר בסעיף 1, כפי שהוסבר שם.