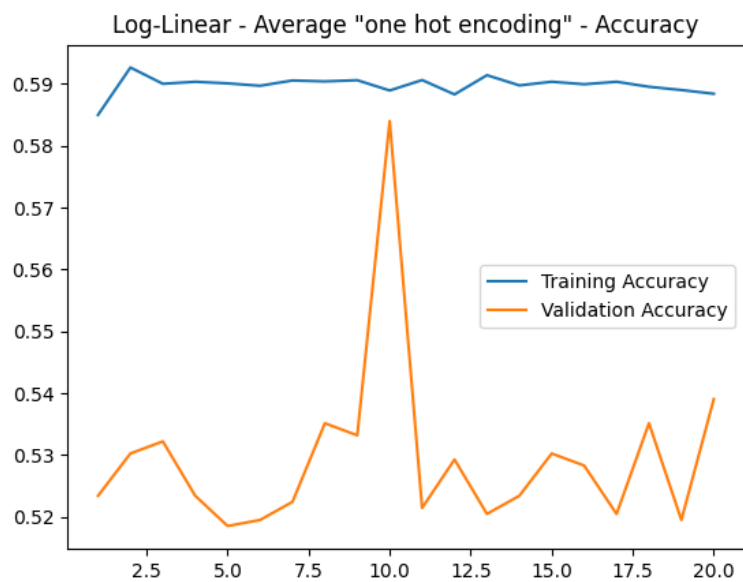
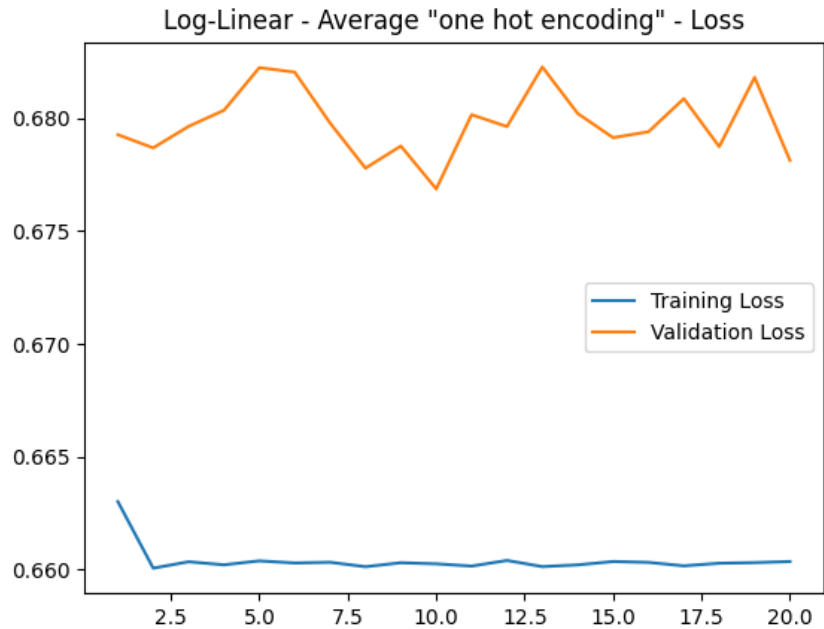
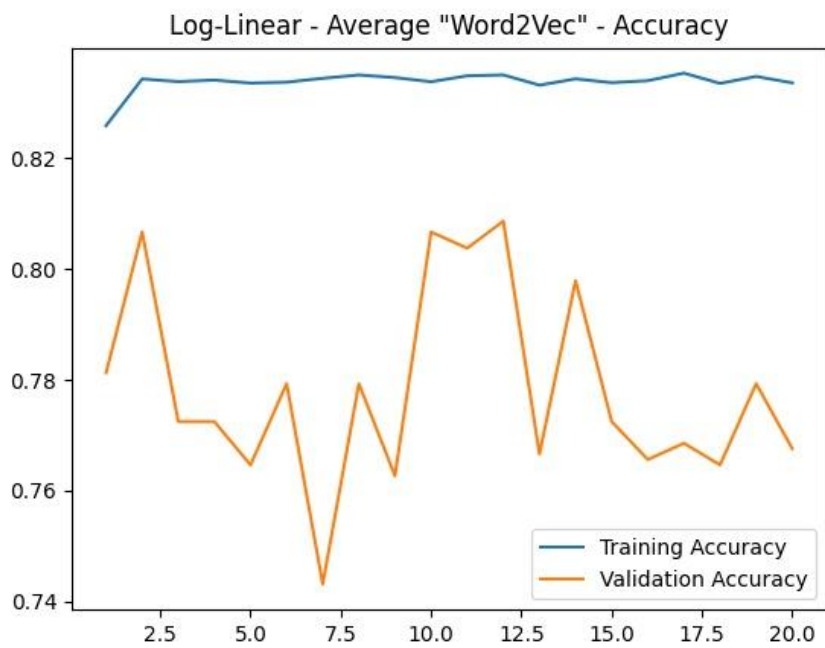
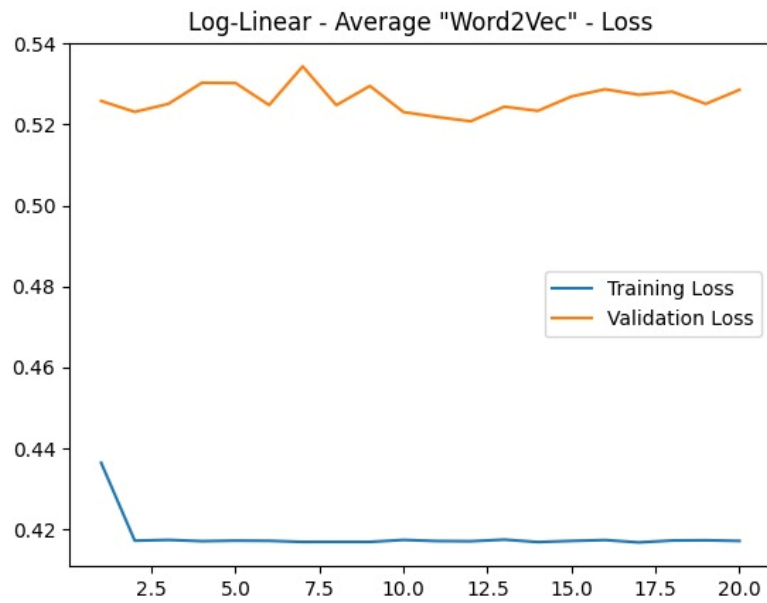


NLP Exercise 3

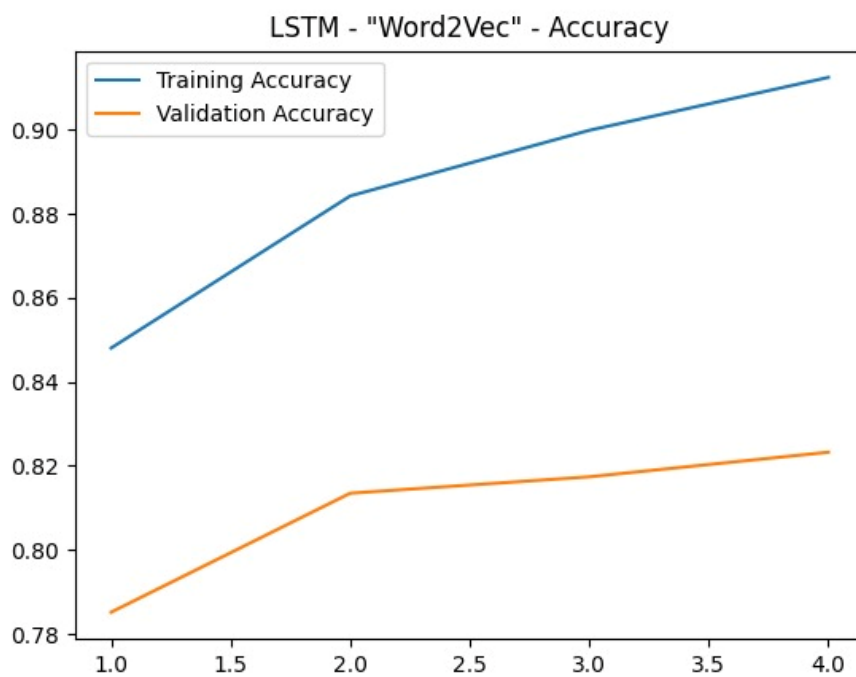
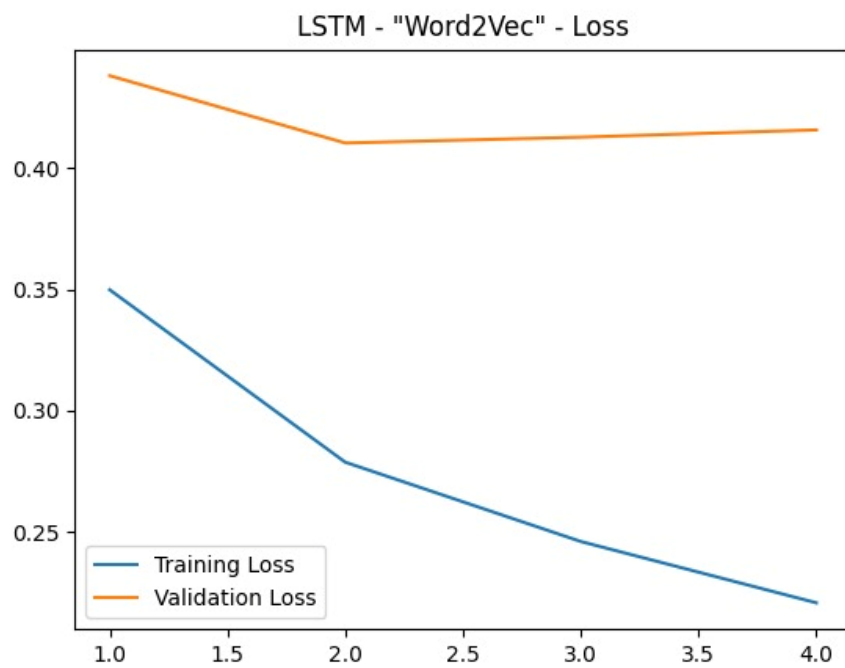
Comparing the different models



```
Log-Linear - Average "one hot encoding" - Test accuracy: 0.569
Log-Linear - Average "one hot encoding" - Test Loss: 0.674
Log-Linear - Average "one hot encoding" - Negated accuracy: 0.484
Log-Linear - Average "one hot encoding" - Rare accuracy: 0.300
```



```
Log-Linear - Average "Word2Vec" - Test accuracy: 0.803
Log-Linear - Average "Word2Vec" - Test Loss: 0.487
Log-Linear - Average "Word2Vec" - Negated accuracy: 0.597
Log-Linear - Average "Word2Vec" - Rare accuracy: 0.640
```



```
LSTM - "Word2Vec" - Test accuracy: 0.886
LSTM - "Word2Vec" "Word2Vec" - Test Loss: 0.291
LSTM - "Word2Vec" - Negated accuracy: 0.677
LSTM - "Word2Vec" - Rare accuracy: 0.840
```

1. כפי שניתן לראות על פי הגרפים והנתונים הנומריים, מודל ה-Log linear שקיבל כ- input את ממוצע קידוד ה-Word2Vec השיג דיוק גבוהה משמעותית מהמודל שקיבל את ממוצע one hot encoding – בטסט: 0.803 מול 0.569, ובולידציה: 0.77 מול 0.54.

תוצאות אלו אינן מפתיעות, שכן הקידוד של Word2Vec אינפורמטיבי ומכיל מידע רב על כל מילה לעומת וקטור יחידה פשוט. ניתן להסתכל על קידוד איכותי בתור נקודת פתיחה טובה יותר למודל, דבר המקל משמעותית על משימתו.

במילים אחרות: Word2Vec מספק ייצוגים וקטוריים צפופים עבור מילים התופסים קווי דמיון סמנטיים בהתבסס על ההקשר שלהן בקורפוס גדול, לעומת ייצוג באמצעות וקטורי היחידה, שם כל מילה רחוקה מכל מילה אחרת, בצורה זהה.

2. כפי שניתן לראות על פי הגרפים והנתונים הנומריים, מודל ה-LSTM שקיבל כ- input את קידוד ה-Word2Vec השיג דיוק גבוהה משמעותית מהמודל הלוג ליניארי שקיבל את ממוצע קידודי Word2Vec. בטסט: 0.886 מול 0.803, ובולידציה: 0.825 מול 0.77.

גם במקרה הנ"ל, ניתן היה לצפות לתוצאות אלו, שכן:

- א. מודל LSTM מכיל מס' רב יותר של פרמטרים, על כן הוא בעל יכולת אקספרסיביות גבוהה יותר. באופן כללי למודל זה, בדומה ל-RNN, יש יכולת לנתח משפטים ארוכים. לכן המודל מתאים יותר למשימה של ניתוח סמנטי של משפטים.
 - ב. מודל ה-LSTM רץ בצורה דו-כיוונית על כל משפט, דבר המעניק לו הבנה עמוקה ושלמה יותר של המשפט בכללותו.
 - ג. כל תא ב-LSTM מקבל ייצוג וקטורי של W2V לעומת המיצוע שמקבל המודל הלוג ליניארי.
- נציין כי תוצאות אלו עדיין מרשימות שכן מודל ה-LSTM ביצע רק 4 מעברים על סט האימון לעומת 20 מעברים שביצע המודל הלוג ליניארי.

3.

א. **Negated polarity**

המודל עם הדיוק הגבוהה ביותר על תת קבוצה זו של סט הבדיקה הינו **LSTM** עם דיוק של 0.677, לעומת זאת המודל עם הדיוק הנמוך ביותר הינו מודל לוג ליניארי עם one hot encoding עם דיוק של 0.484.

ניתן ליחס את הדיוק הגבוהה יותר שקיבל מודל ה- LSTM ליכולתו להבין משפטים ארוכים ולשמור באמצעות hidden state את משמעות הסמנטית של המשפט לאורכו. לעומת זאת, המודל הלוג ליניארי הפשוט, מתקשה במשימה זו.

ב. **Rare words** –

המודל עם הדיוק הגבוהה ביותר על תת קבוצה זו של סט הבדיקה הינו **LSTM** עם דיוק של 0.84, לעומת זאת המודל עם הדיוק הנמוך ביותר הינו מודל לוג ליניארי עם one hot encoding עם דיוק של 0.3.

ניתן ליחס הבדל גדול זה, לקידוד ה- Word2Vec, שכן מילים נדירות, שאולי ולא ראינו במהלך האימון, מקבלות קידוד דומה למילים שאותן ראינו, לכן התוצאות על משפטים עם מילים אלו, יישארו גבוהות. תכונה זו, הינה תוצאה ומטרה ישירה של קידוד ה- Word2Vec, אשר ניסה לקדם מילים עם משמעות דומה, כך שוקטורי הייצוג יהיו דומים במרחב והזווית ביניהן תהיה קטנה בכל שניתן.

קידוד ה- one hot encoding לעומת זאת, אינו אינפורמטיבי בכלל, דבר המקשה על המודל הלוג ליניארי להצליח במשימה זו, ולהכליל למילים שלא ראה.

הערה כללית:

בתמונות המכילות את תוצאות המודלים על סט האימון, התוצאות מעוגלות עד 3 ספרות אחרי הנק' העשרונית.