Genre Classification and Sentiment Analysis of Game Reviews

ASAD KOTHAWALA

B00825551

Introduction

▶ The video game industry is one of the fastest growing industries in the world.

▶ Lots of games are reviewed and advertised online. This makes lots of data available.

▶ Goal is to perform Sentiment Analysis and Multi-Label Classification on Steam Review Data.

Datasets

- Steam Python API no longer working

- The first dataset contains over 6.4 million reviews however it has only 5 columns (including whether others found a review helpful).

- Supplementary dataset containing Steam games and genre information was used to label the reviews.

|	app_id	app_name	review_text	review_score	review_votes
0	10	Counter-Strike	Ruined my life.	1	0
1	10	Counter-Strike	This will be more of a "my experience with th...	1	1
2	10	Counter-Strike	This game saved my virginity.	1	0
3	10	Counter-Strike	• Do you like original games? • Do you like ga...	1	0
4	10	Counter-Strike	Easy to learn, hard to master.	1	1
...	...	...	...	...	...
6417101	99910	Puzzle Pirates	I really ove this game but it needs somethings...	-1	0
6417102	99910	Puzzle Pirates	Used to play Puzzel Pirates 'way back when', b...	-1	0
6417103	99910	Puzzle Pirates	This game was aright, though a bit annoying. W...	-1	0
6417104	99910	Puzzle Pirates	I had a nice review to recommend this game, bu...	-1	0
6417105	99910	Puzzle Pirates	The puzzles in this game are fun, but you have...	-1	0

Data Preparation

- Merged review dataset and games list to create a labelled dataset.

- Removed null fields.

- For sentiment analysis I took the top ten games of the genres I am interested in.

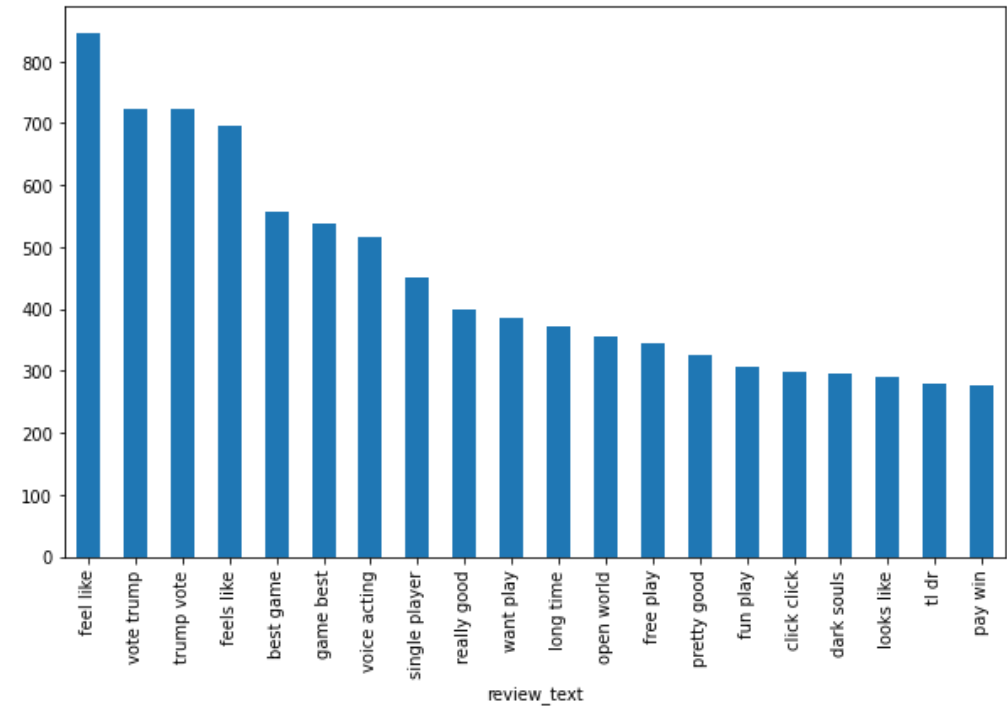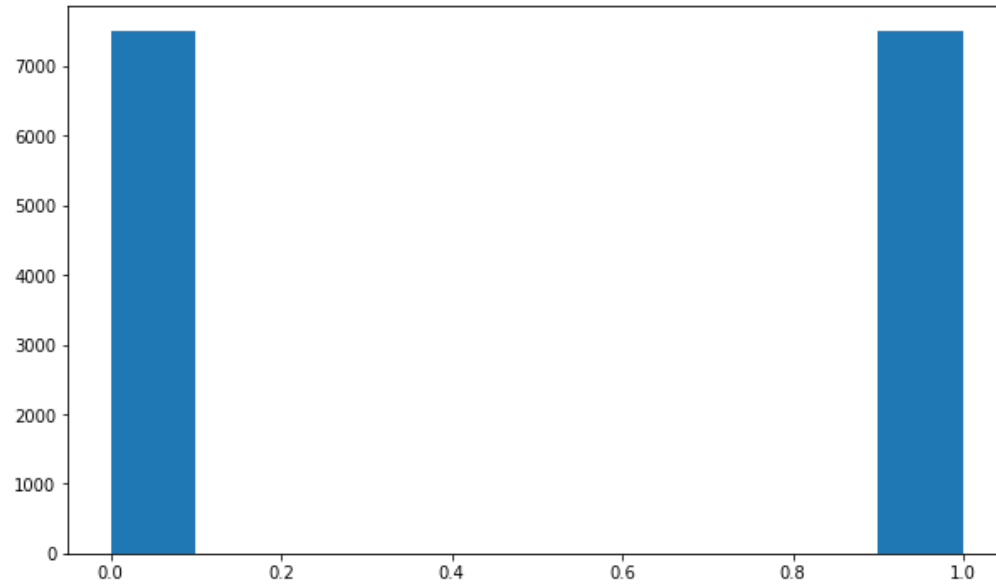- For Multilabel Classification I took random sample from the more helpful reviews.

Data Exploration
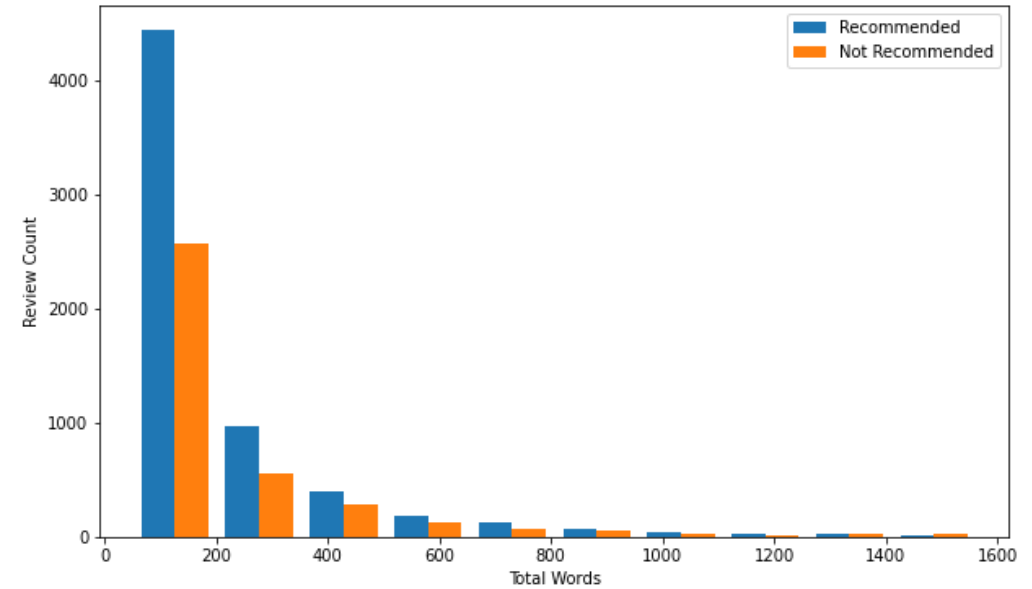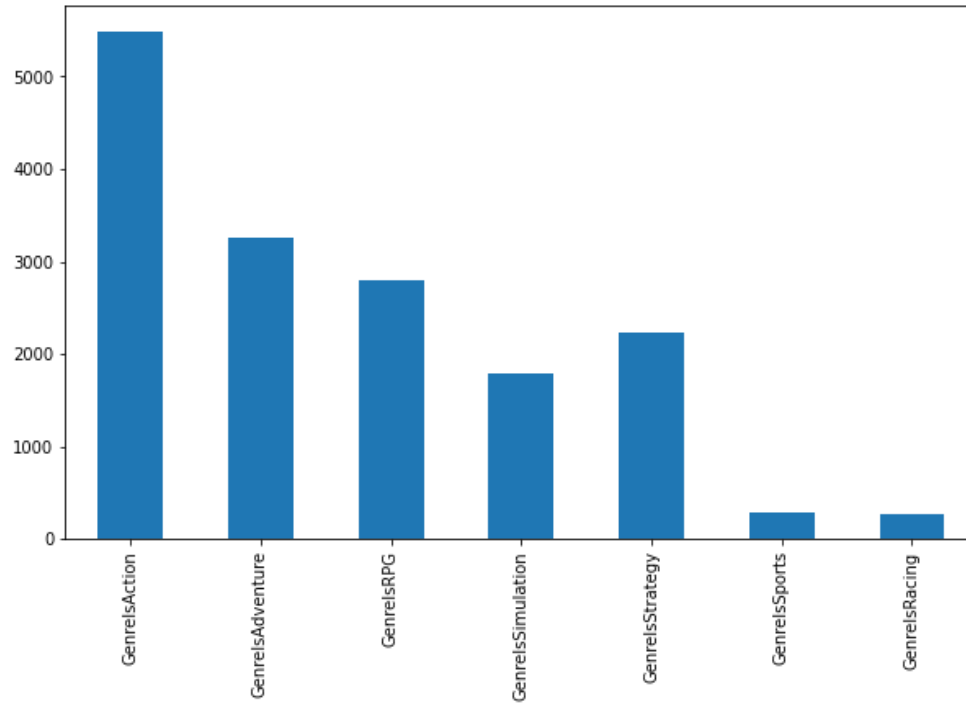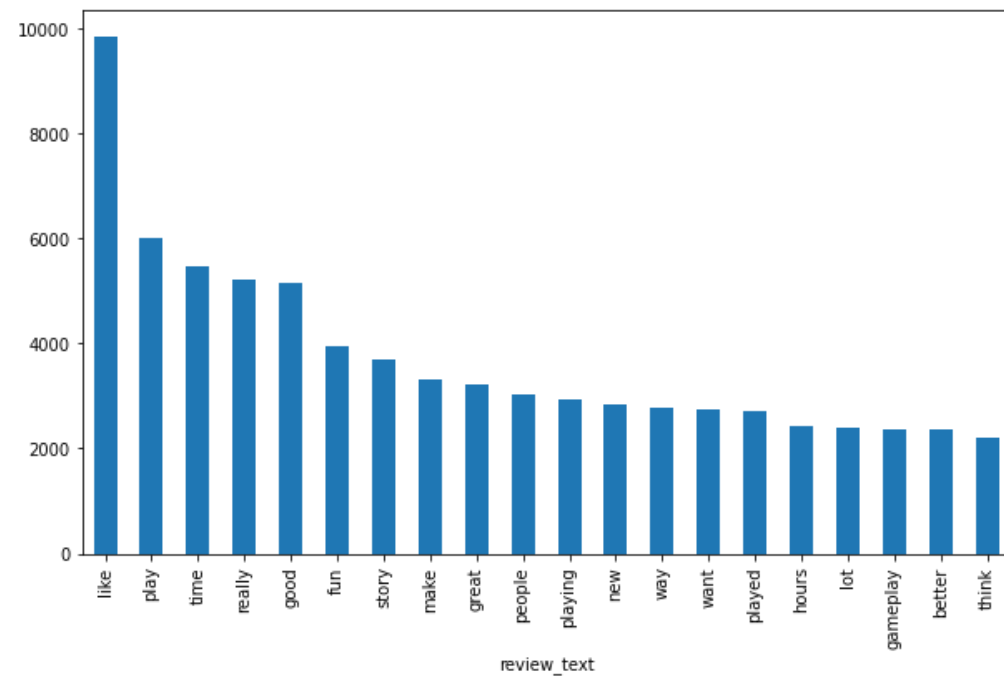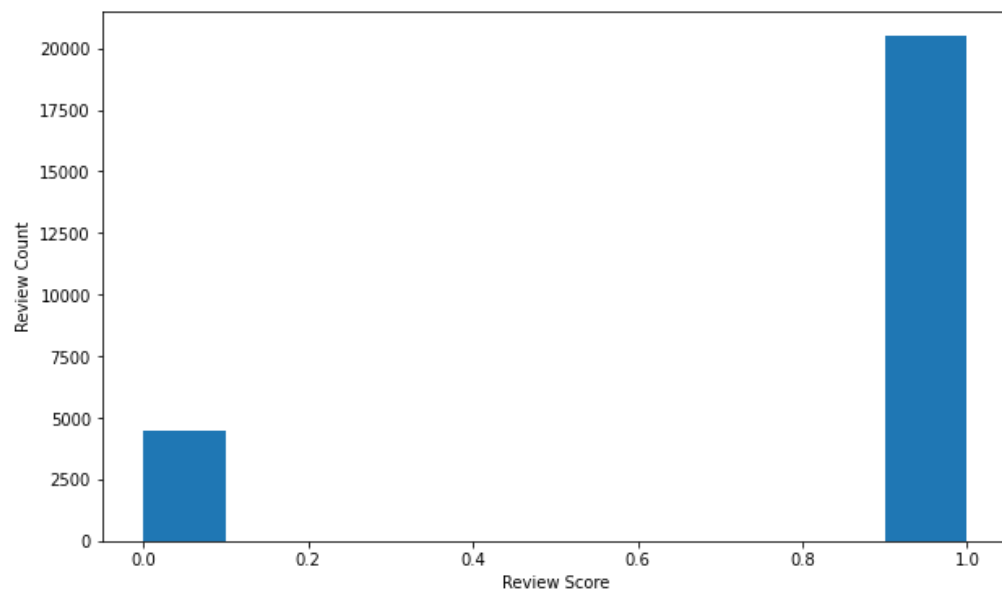
Data Exploration

Data Exploration

Proposed Techniques

- LSTM
- SVM
- Word Embeddings
  - GloVe
  - Word2Vec
  - TF-IDF

Evaluation Metrics

- Accuracy
- Hamming Loss
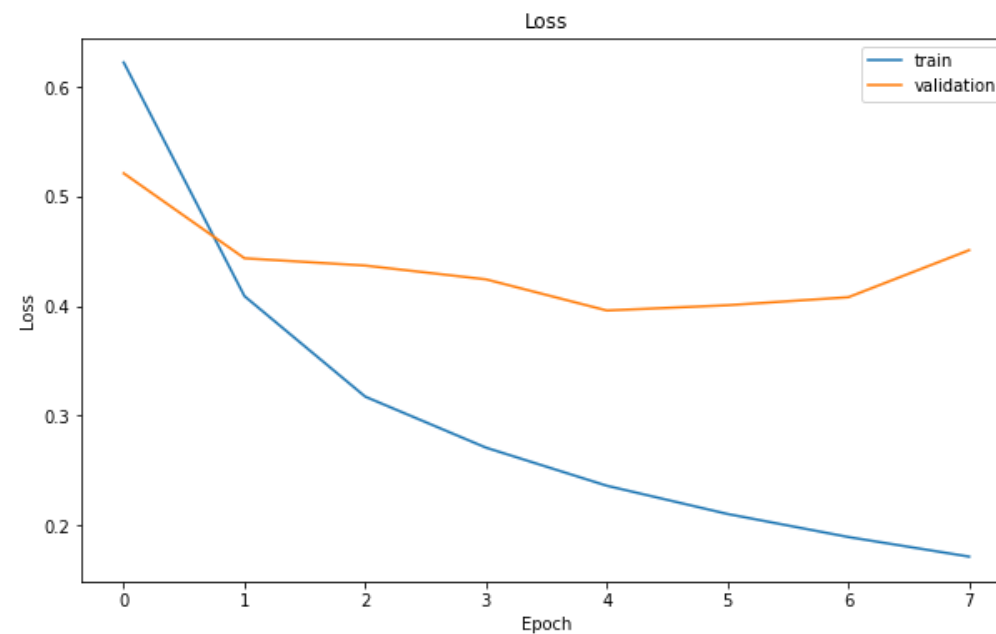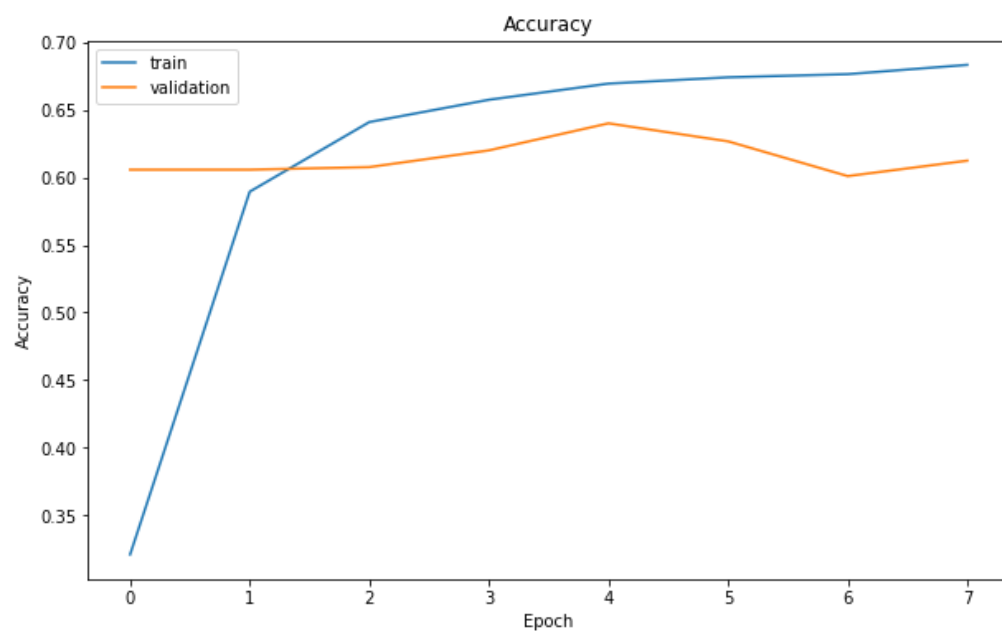- Precision, Recall, F1-score

Part A: Multi-Label Genre Classification

Bidirectional LSTM (Base)

- Applied preprocessing pipeline and sequenced data for LSTM.

- Bidirectional

- SpatialDropout1D Layer

```
Model: "sequential_7"

Layer (type)                    Output Shape           Param #
=================================================================
embedding_6 (Embedding)         (None, 250, 100)       3642500

spatial_dropout1d_6 (Spatial    (None, 250, 100)       0

bidirectional_1 (Bidirection    (None, 128)            84480

batch_normalization_3 (Batch    (None, 128)            512

dropout_3 (Dropout)             (None, 128)            0

dense_6 (Dense)                 (None, 7)              903
=================================================================
Total params: 3,728,395
Trainable params: 3,728,139
Non-trainable params: 256
```

Bidirectional LSTM (Base)

Bidirectional LSTM (Base)

```
Hamming loss is  0.1678095238095238
              precision    recall  f1-score   support

           0       0.71      0.83      0.76      2537
           1       0.64      0.37      0.47      1467
           2       0.65      0.37      0.47      1266
           3       0.75      0.16      0.26       802
           4       0.70      0.37      0.48       983
           5       0.94      0.13      0.23       130
           6       1.00      0.14      0.24       117

   micro avg       0.69      0.50      0.58      7302
   macro avg       0.77      0.34      0.42      7302
weighted avg       0.70      0.50      0.54      7302
 samples avg       0.60      0.52      0.53      7302
```
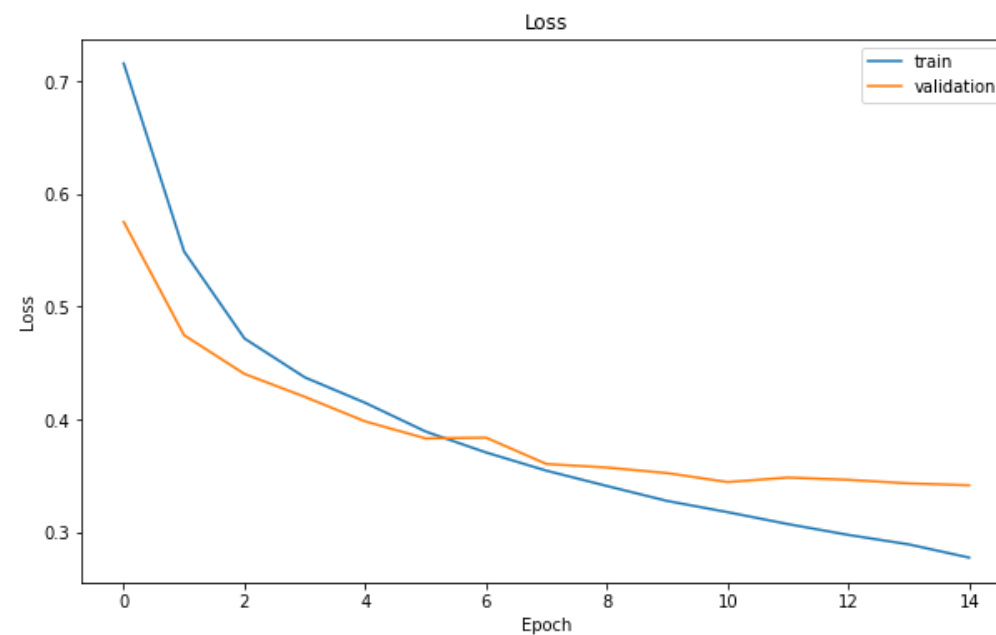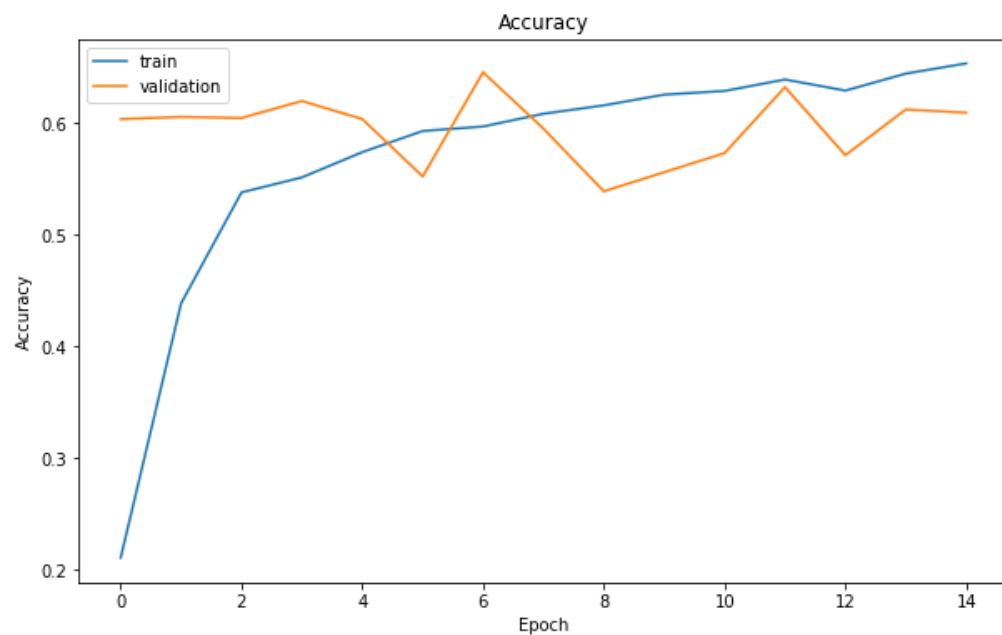
Bidirectional LSTM (GloVe)

Bidirectional LSTM (GloVe)

```
Hamming loss is  0.14765079365079364
              precision    recall  f1-score   support

           0       0.77      0.82      0.79      2537
           1       0.66      0.47      0.55      1467
           2       0.68      0.57      0.62      1266
           3       0.69      0.31      0.43       802
           4       0.72      0.46      0.56       983
           5       0.70      0.45      0.54       130
           6       0.75      0.58      0.65       117


   micro avg       0.72      0.59      0.65      7302
   macro avg       0.71      0.52      0.59      7302
weighted avg       0.71      0.59      0.64      7302
 samples avg       0.70      0.62      0.62      7302
```
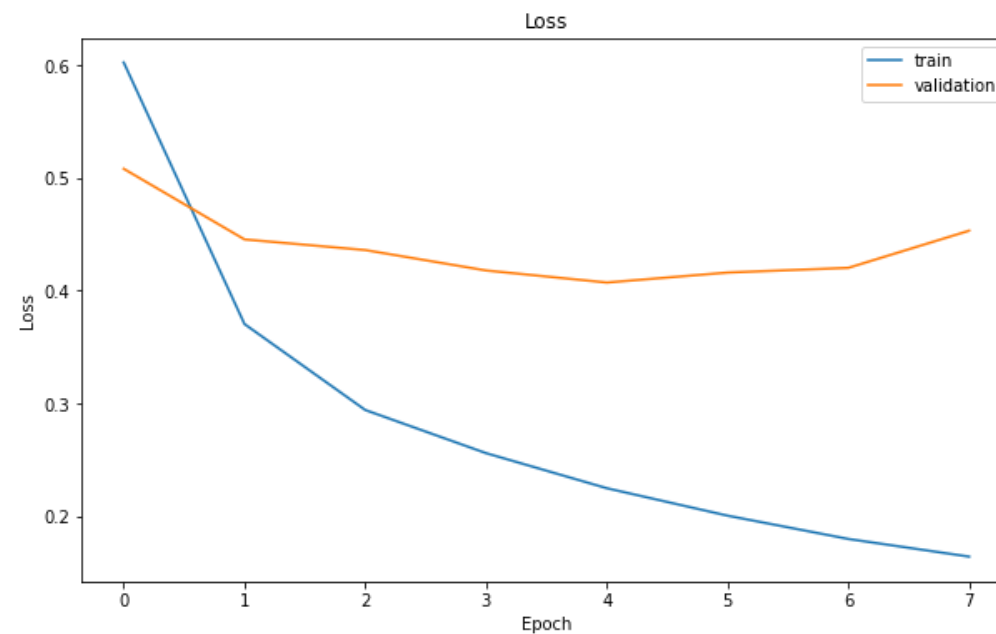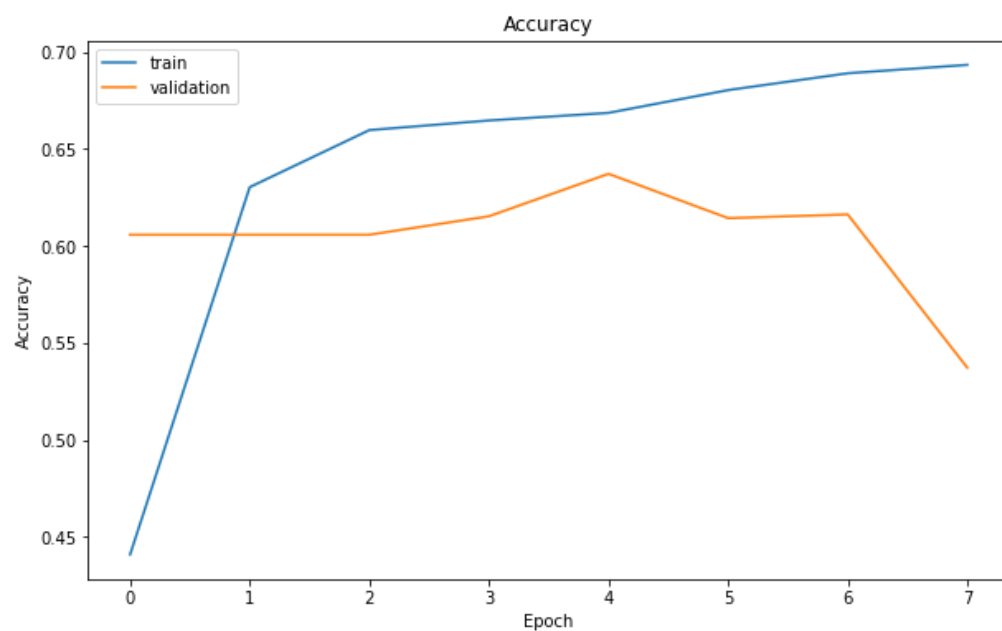
Bidirectional LSTM (Word2Vec)

Bidirectional LSTM (Word2Vec)

```
Hamming loss is  0.17266666666666666
           precision    recall  f1-score   support

         0       0.75      0.68      0.72      2537
         1       0.66      0.34      0.45      1467
         2       0.60      0.47      0.52      1266
         3       0.57      0.29      0.39       802
         4       0.62      0.45      0.52       983
         5       1.00      0.14      0.24       130
         6       0.90      0.16      0.28       117

 micro avg       0.68      0.48      0.56      7302
 macro avg       0.73      0.36      0.44      7302
weighted avg     0.68      0.48      0.55      7302
samples avg      0.57      0.50      0.50      7302
```

Support Vector Machine (TF-IDF)

```
Hamming loss is  0.14609523809523808
              precision    recall  f1-score   support

           0       0.77      0.80      0.79      2537
           1       0.62      0.49      0.55      1467
           2       0.72      0.53      0.61      1266
           3       0.69      0.38      0.49       802
           4       0.76      0.48      0.59       983
           5       0.93      0.28      0.44       130
           6       0.96      0.42      0.58       117


   micro avg       0.73      0.59      0.65      7302
   macro avg       0.78      0.49      0.58      7302
weighted avg       0.73      0.59      0.64      7302
 samples avg       0.66      0.60      0.60      7302
```
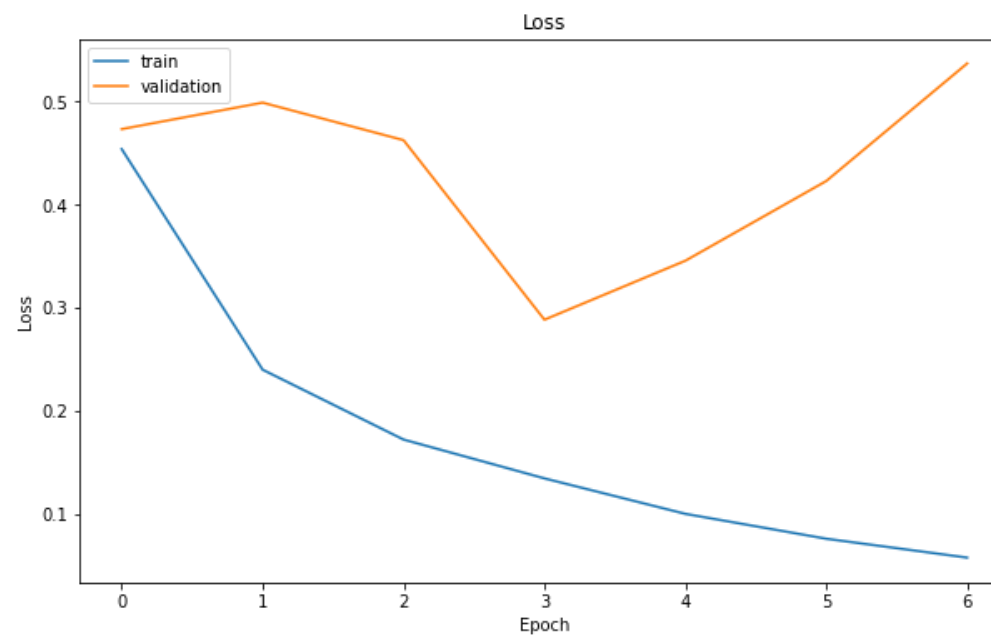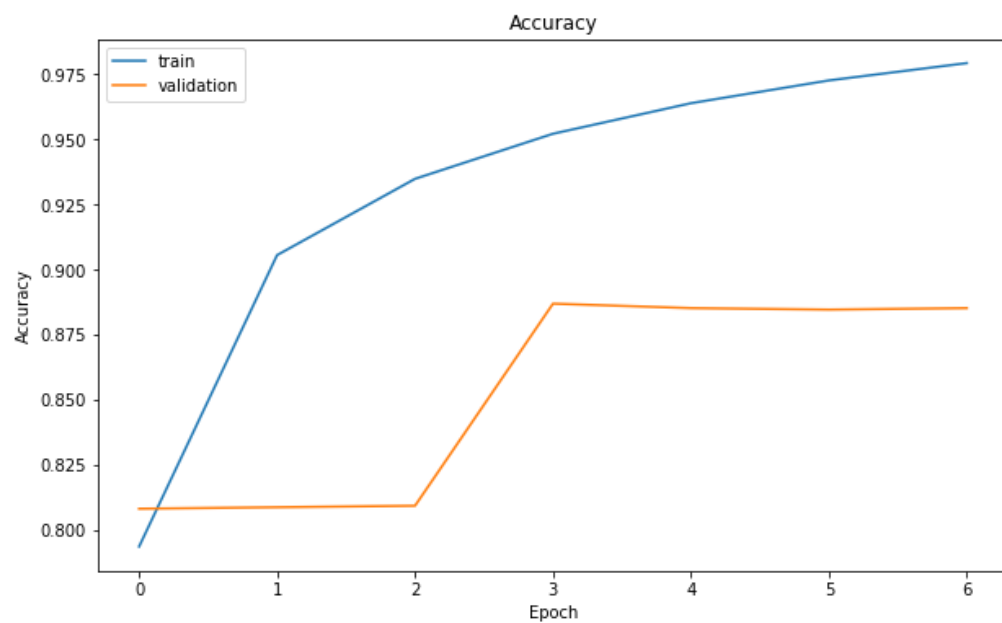
- ▶ LinearSVC
- ▶ OneVsRestClassifier

Part B: Sentiment Analysis

Bidirectional LSTM (Base)

Bidirectional LSTM (Base)
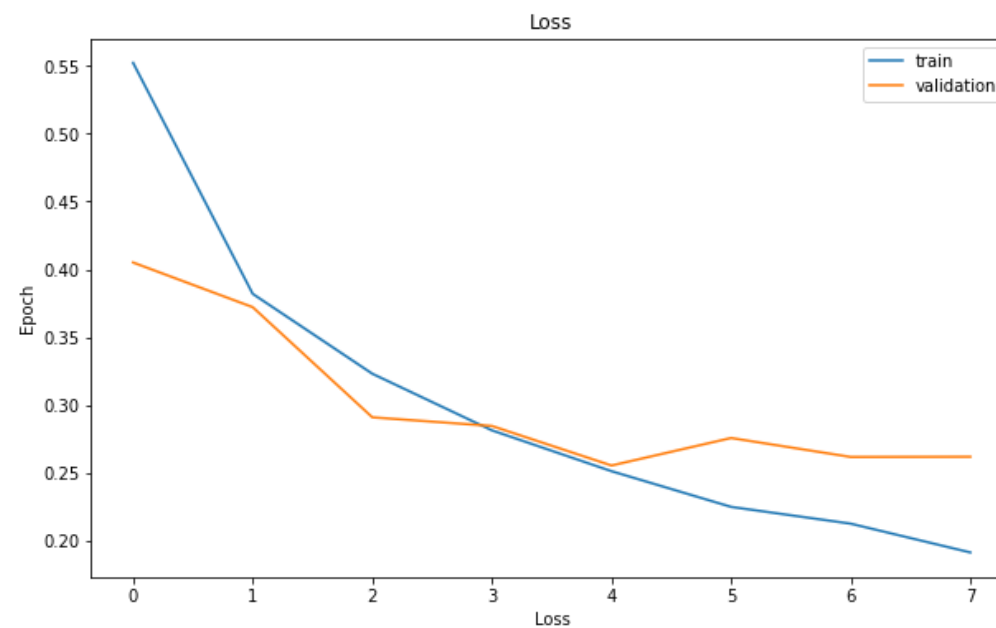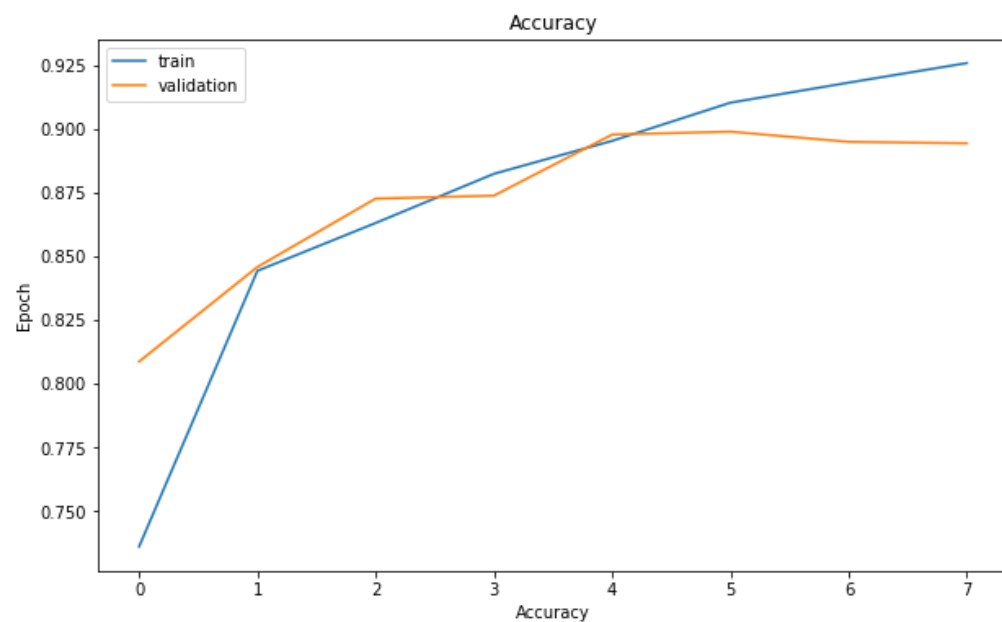
```
Test accuracy is  0.884
              precision    recall  f1-score   support

           0       0.70      0.61      0.66      1351
           1       0.92      0.94      0.93      6149

    accuracy                           0.88      7500
   macro avg       0.81      0.78      0.79      7500
weighted avg       0.88      0.88      0.88      7500
```

Bidirectional LSTM (GloVe)

Bidirectional LSTM (GloVe)

```
Test accuracy is  0.8886666666666667
              precision    recall  f1-score   support

           0       0.66      0.78      0.72      1351
           1       0.95      0.91      0.93      6149


    accuracy                           0.89      7500
   macro avg       0.81      0.84      0.82      7500
weighted avg       0.90      0.89      0.89      7500
```
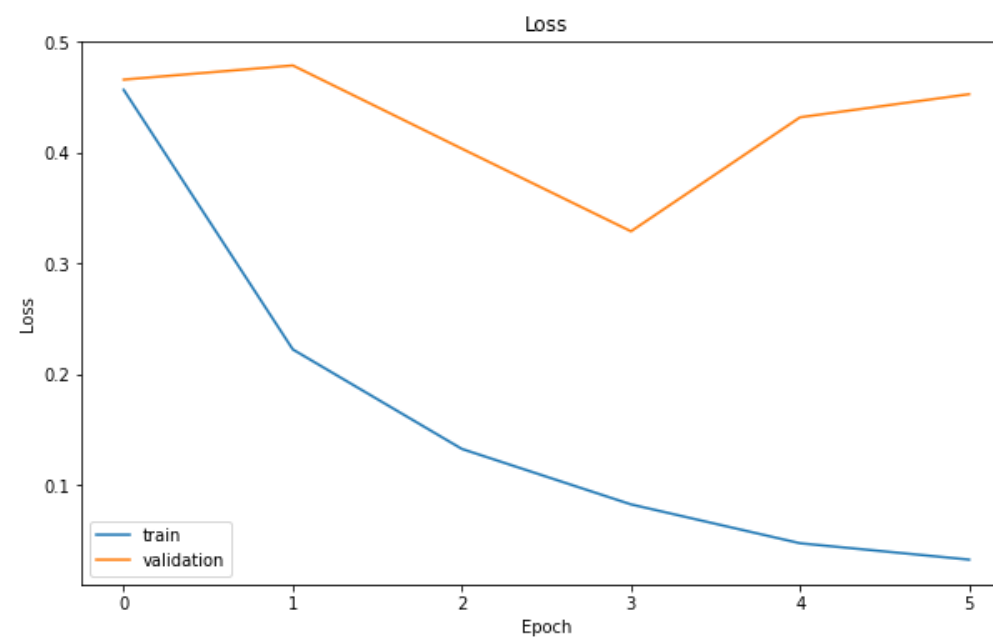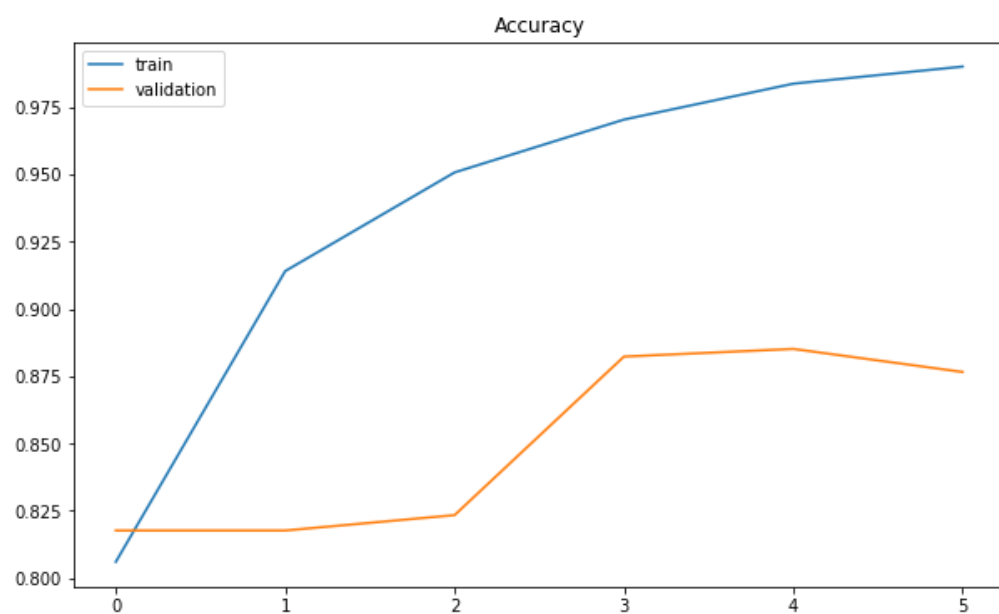
Bidirectional LSTM (Word2Vec)

Bidirectional LSTM (Word2Vec)

```
Test accuracy is  0.88
              precision    recall  f1-score   support

           0       0.74      0.53      0.62      1365
           1       0.90      0.96      0.93      6135

    accuracy                           0.88      7500
   macro avg       0.82      0.74      0.77      7500
weighted avg       0.87      0.88      0.87      7500
```

Support Vector Machine (TF-IDF)


```
Test accuracy is  0.9082666666666667
              precision    recall  f1-score   support

           0       0.83      0.62      0.71      1365
           1       0.92      0.97      0.95      6135


    accuracy                           0.91      7500
   macro avg       0.87      0.80      0.83      7500
weighted avg       0.90      0.91      0.90      7500
```

Conclusion

▶ Although limited in scope, Multilabel model shows similar performance to similar work. This is, however, not very good.

▶ On the other hand, all variants performed sentiment classification of reviews quite well.