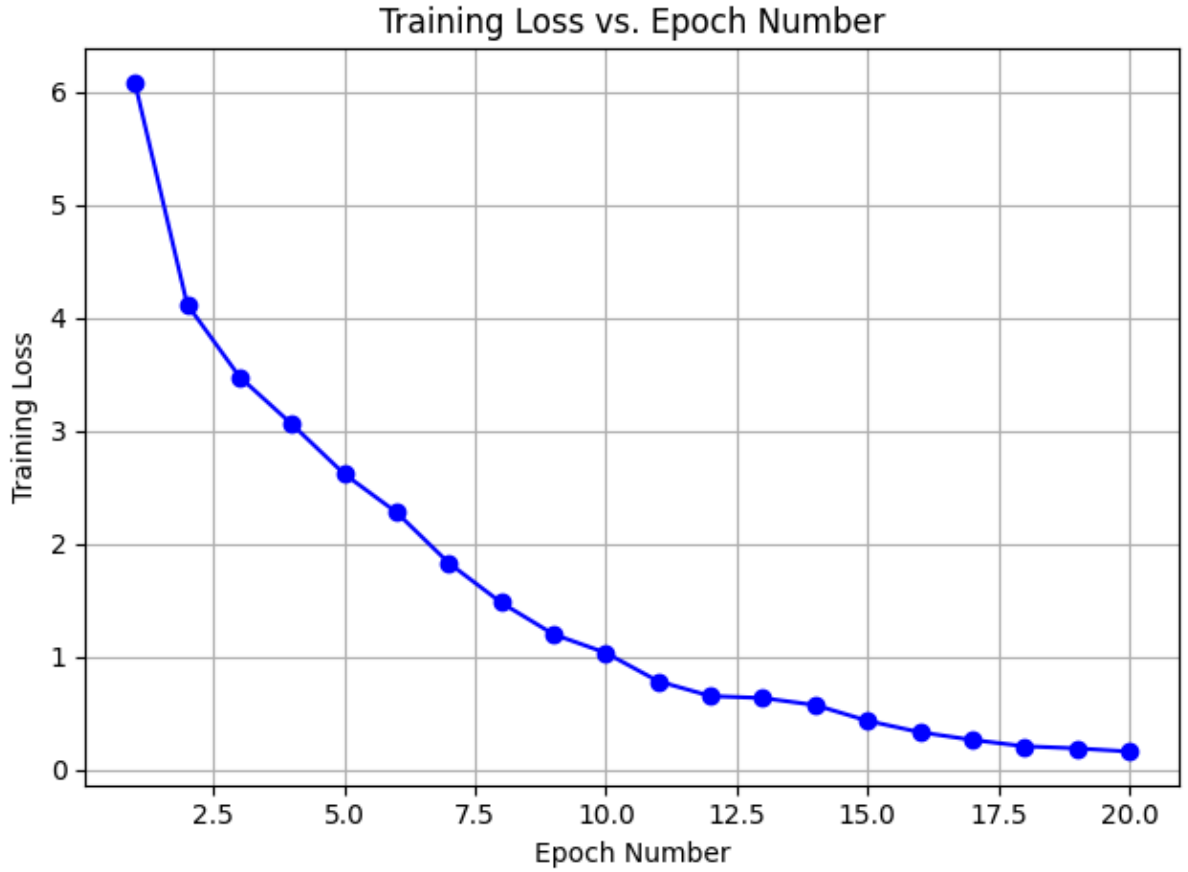
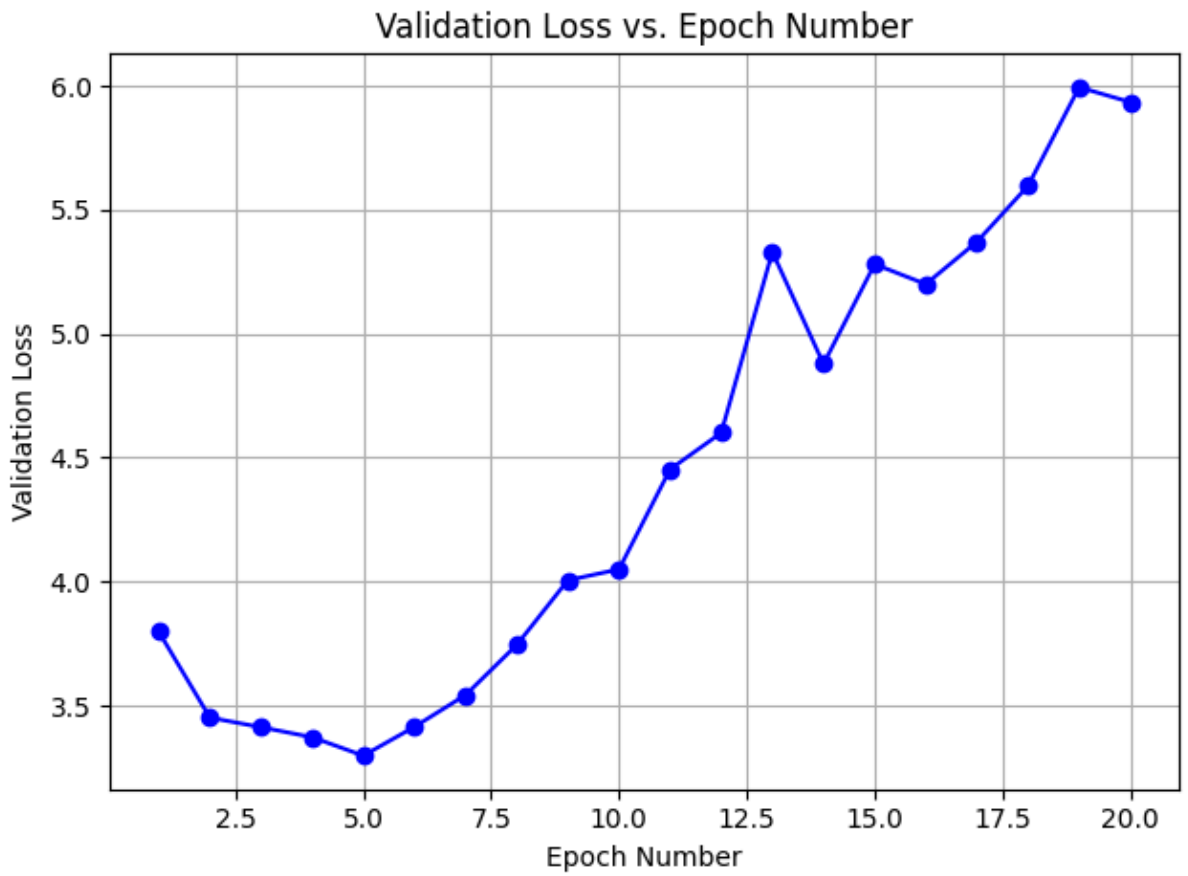
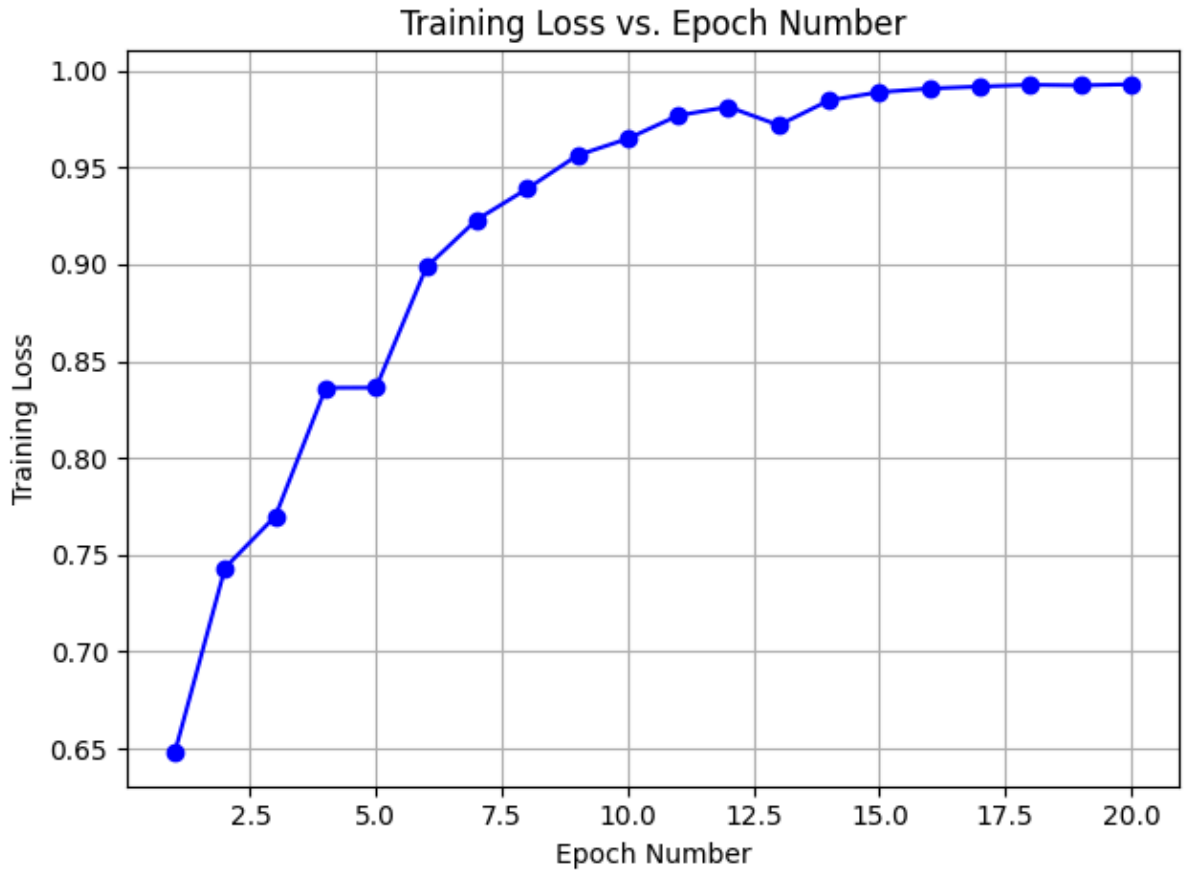
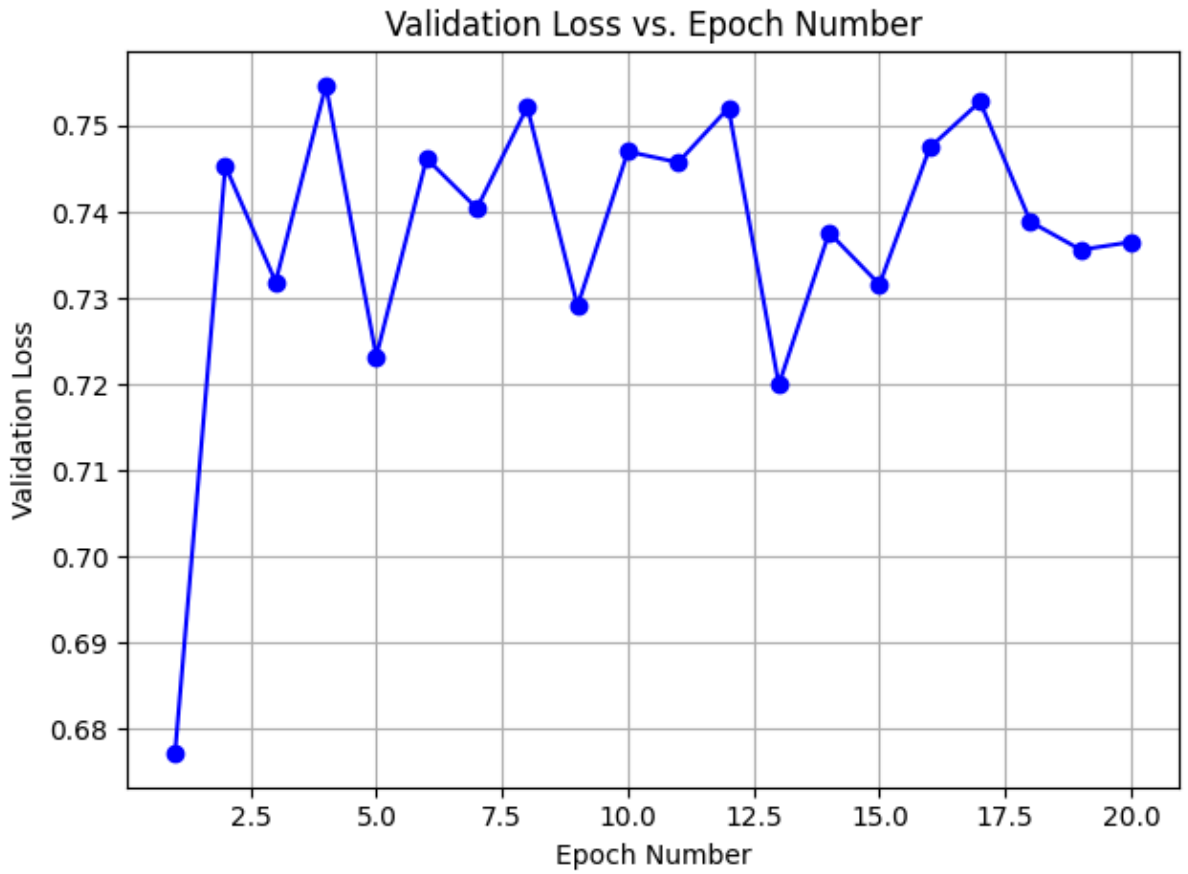
ASSIGNMENT-2 REPORT

GRU
GRU_t2_glove.ipynb



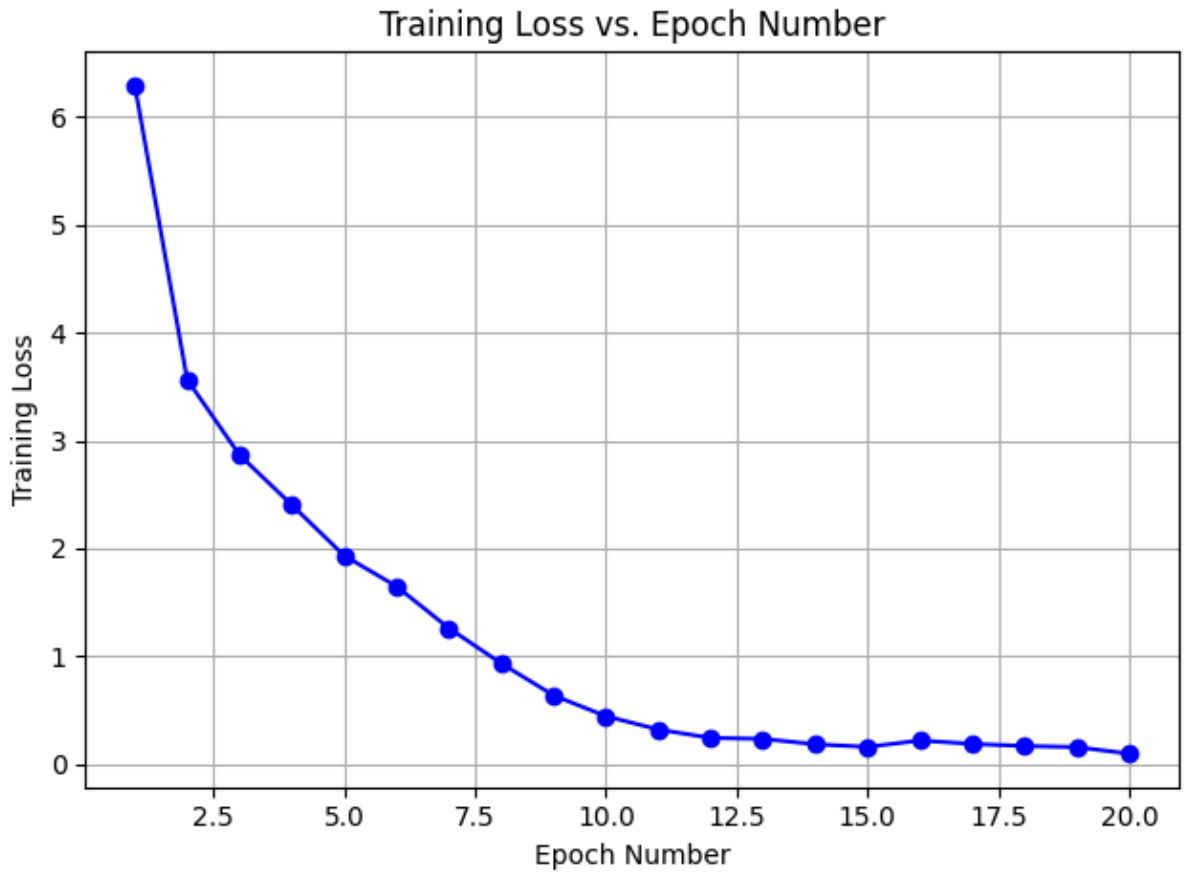

Validation Loss vs. Epoch Number
Validation Loss
Epoch Number
6.0
5.5
5.0
4.5
4.0
3.5
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

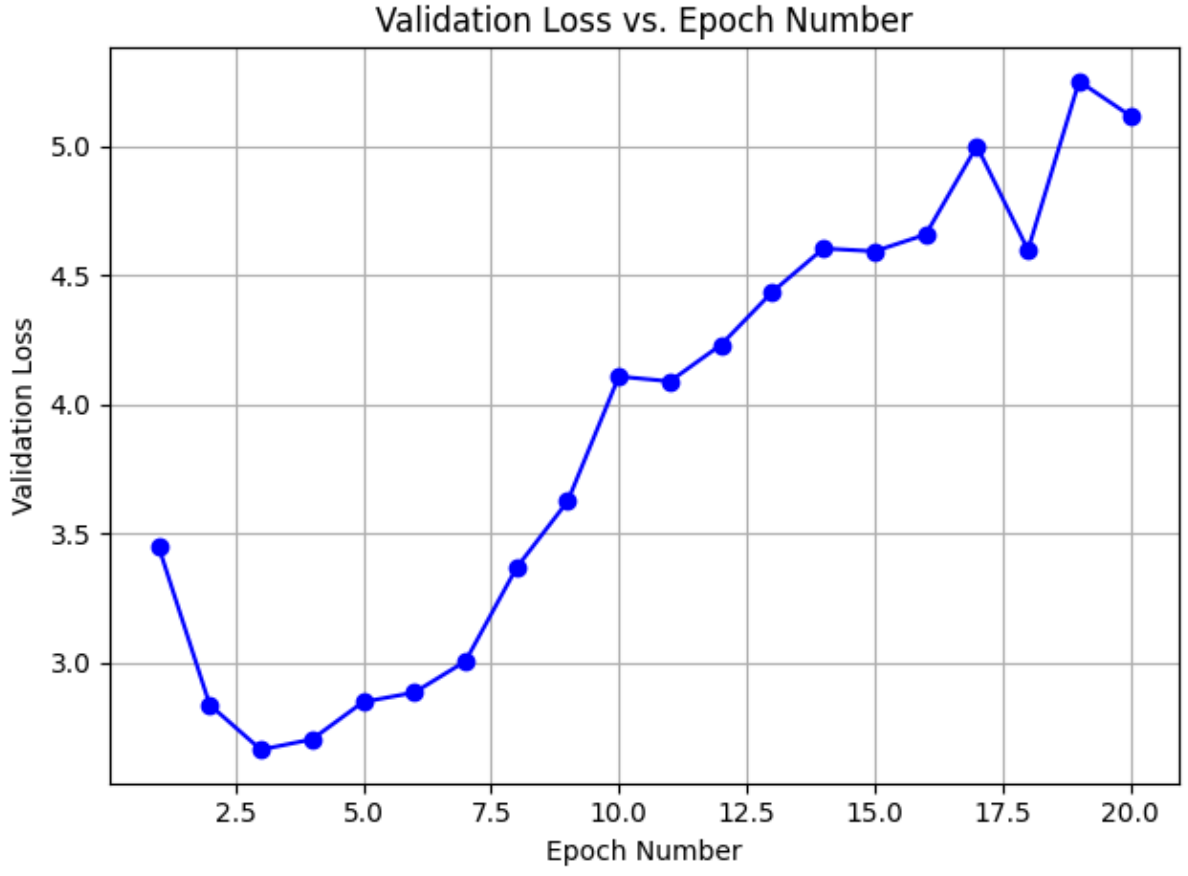

Training Loss vs. Epoch Number
Training Loss
Epoch Number
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

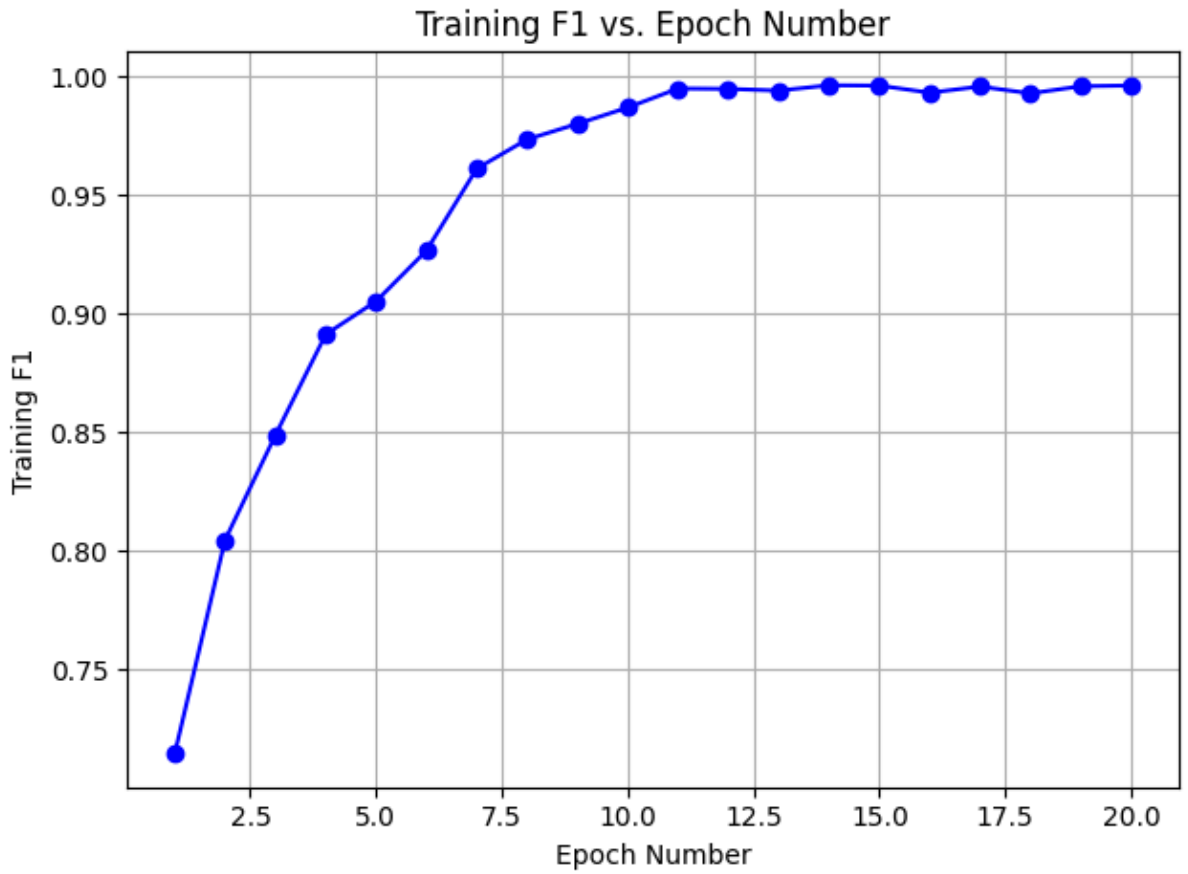


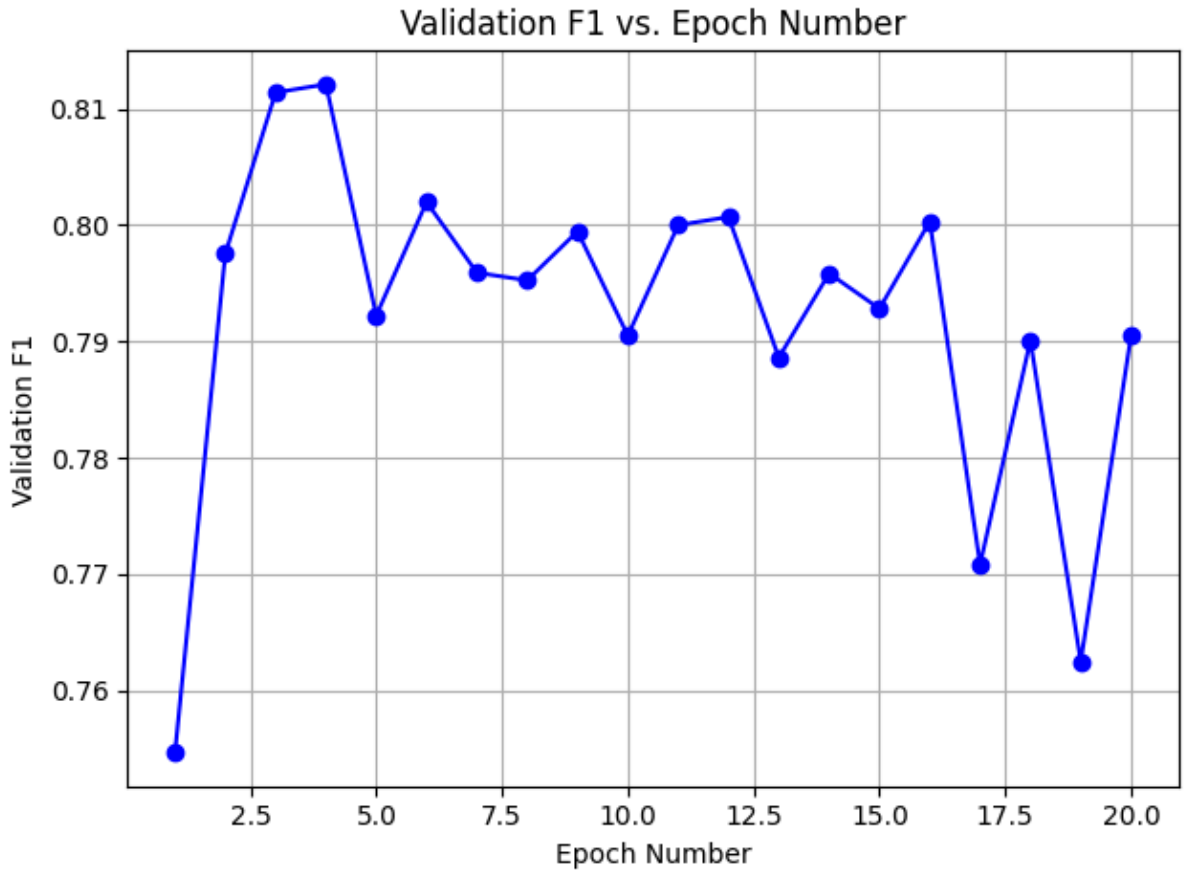
The accuracy of the model is: 0.9067878553471282
The Macro-F1 score obtained is: 0.6960879848463777

GRU_t2_word2vec.ipynb


Training Loss vs. Epoch Number
Training Loss
Epoch Number
0
1
2
3
4
5
6
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0


Validation Loss vs. Epoch Number
Validation Loss
Epoch Number
3.0
3.5
4.0
4.5
5.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

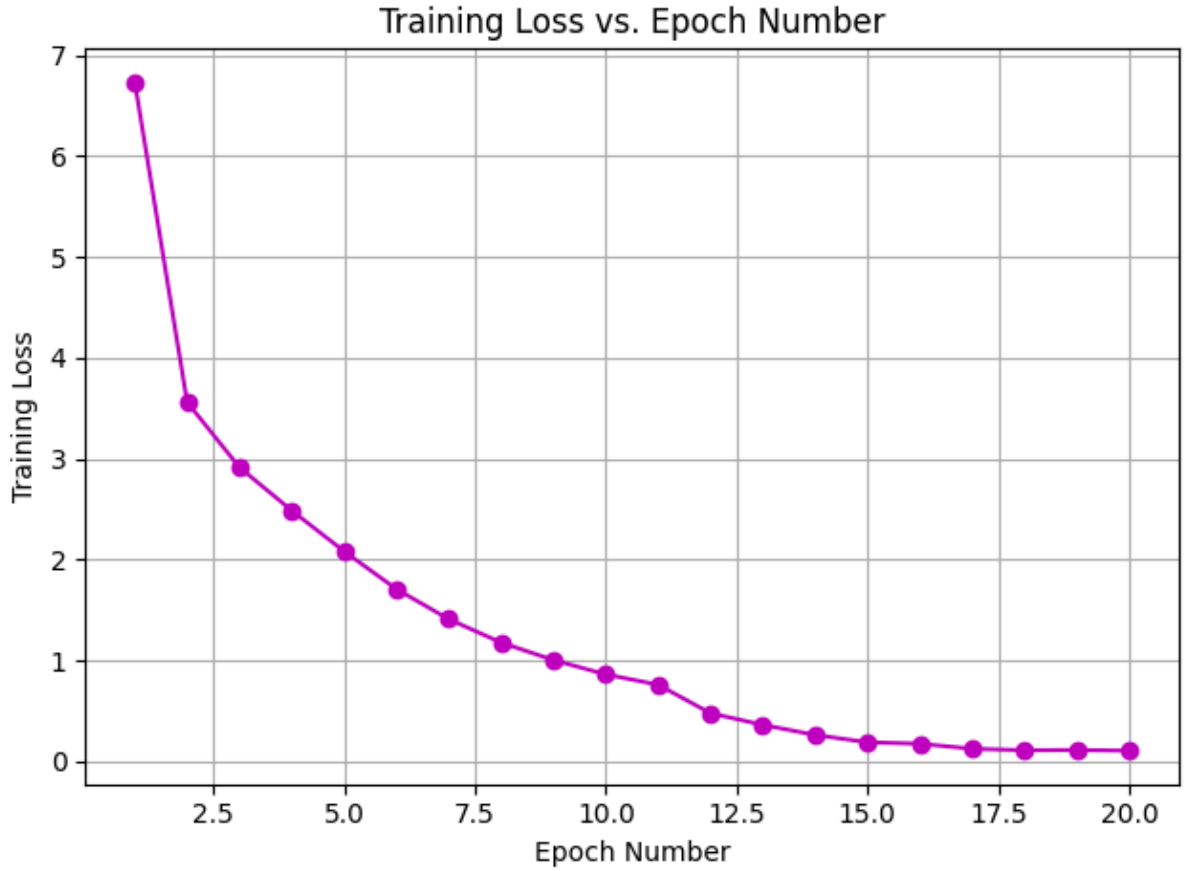


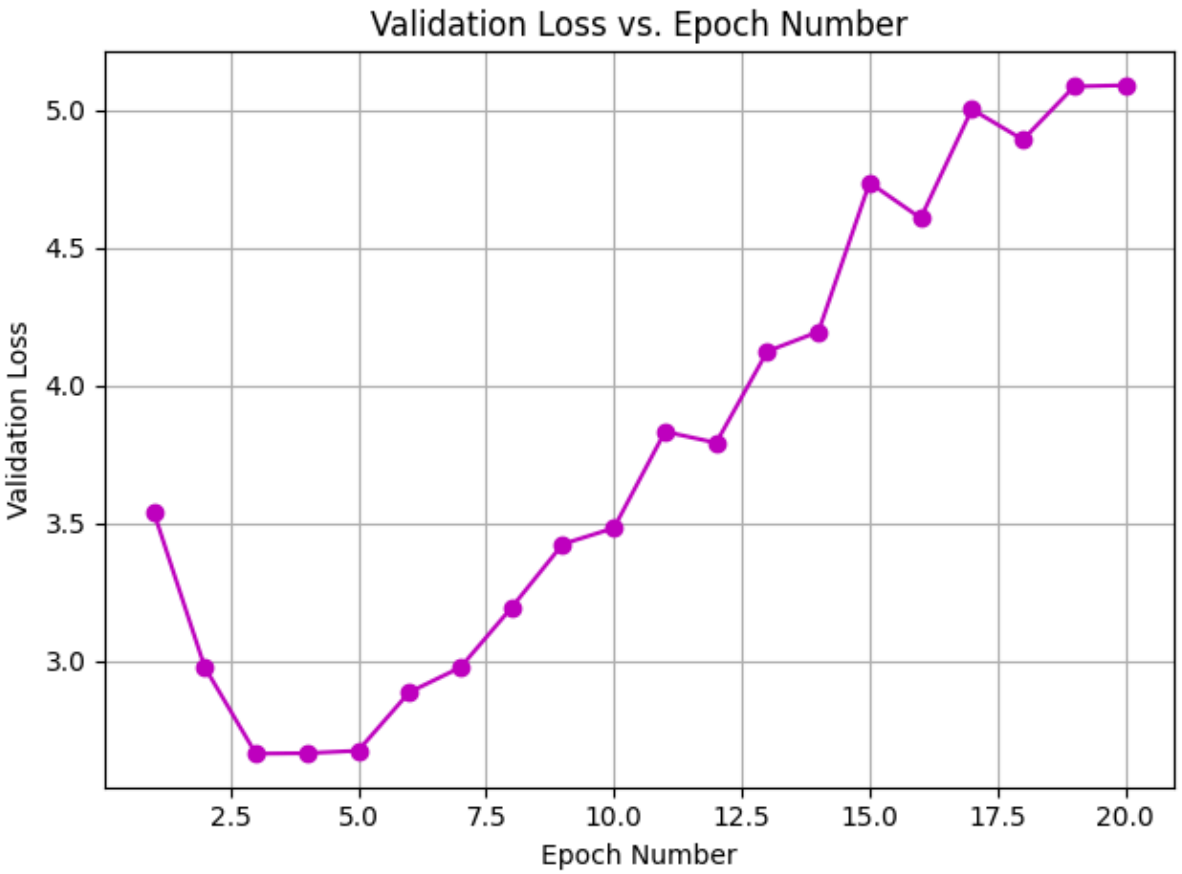


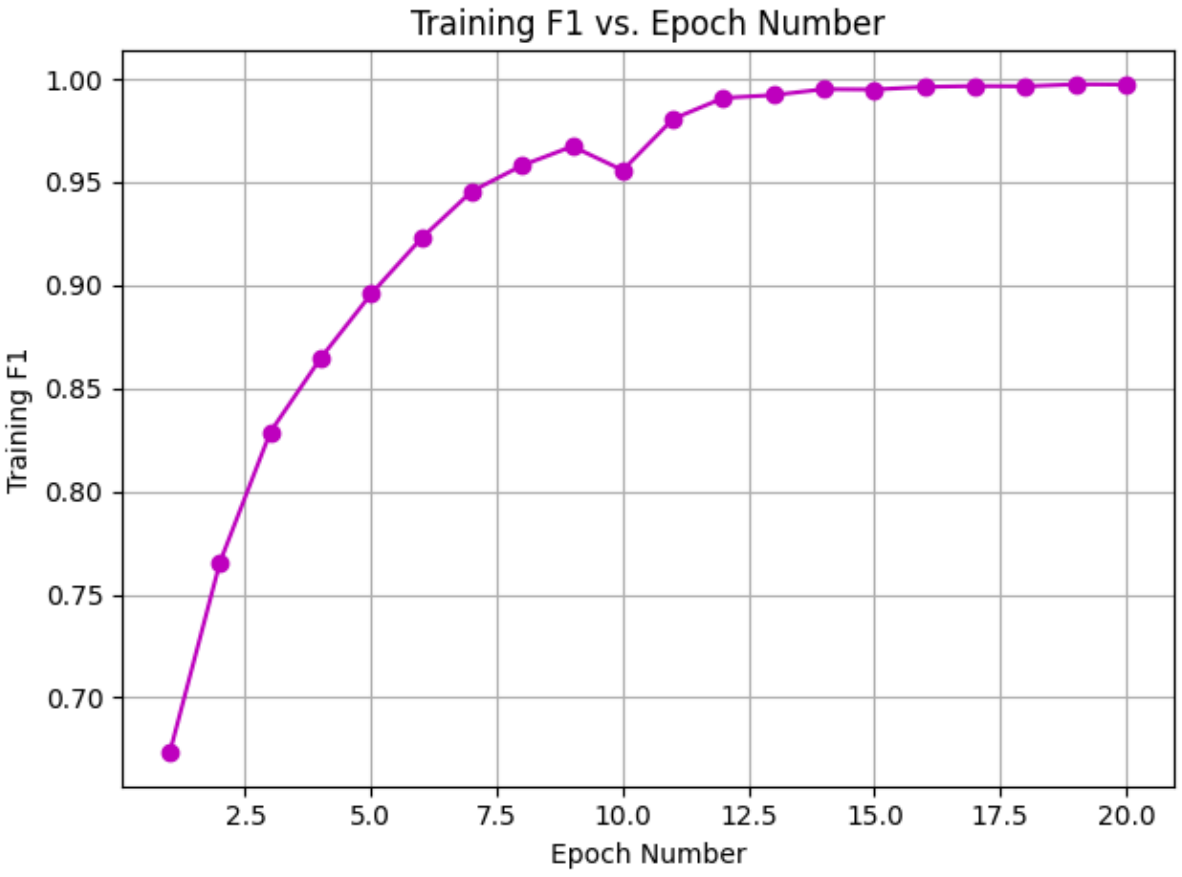
The accuracy of the model is: 0.9301875846064591

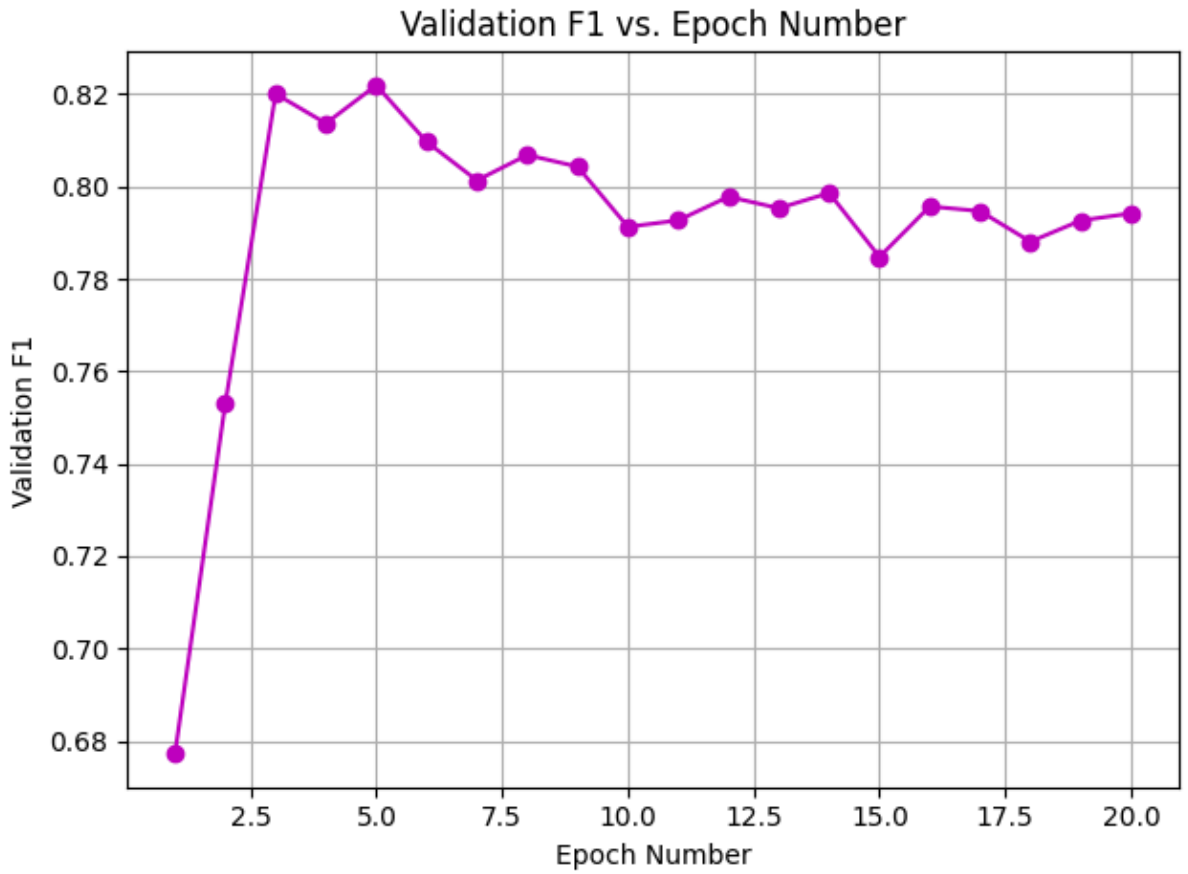
The Macro-F1 score obtained is: 0.7810392752488257

GRU_t2_fasttext.ipynb



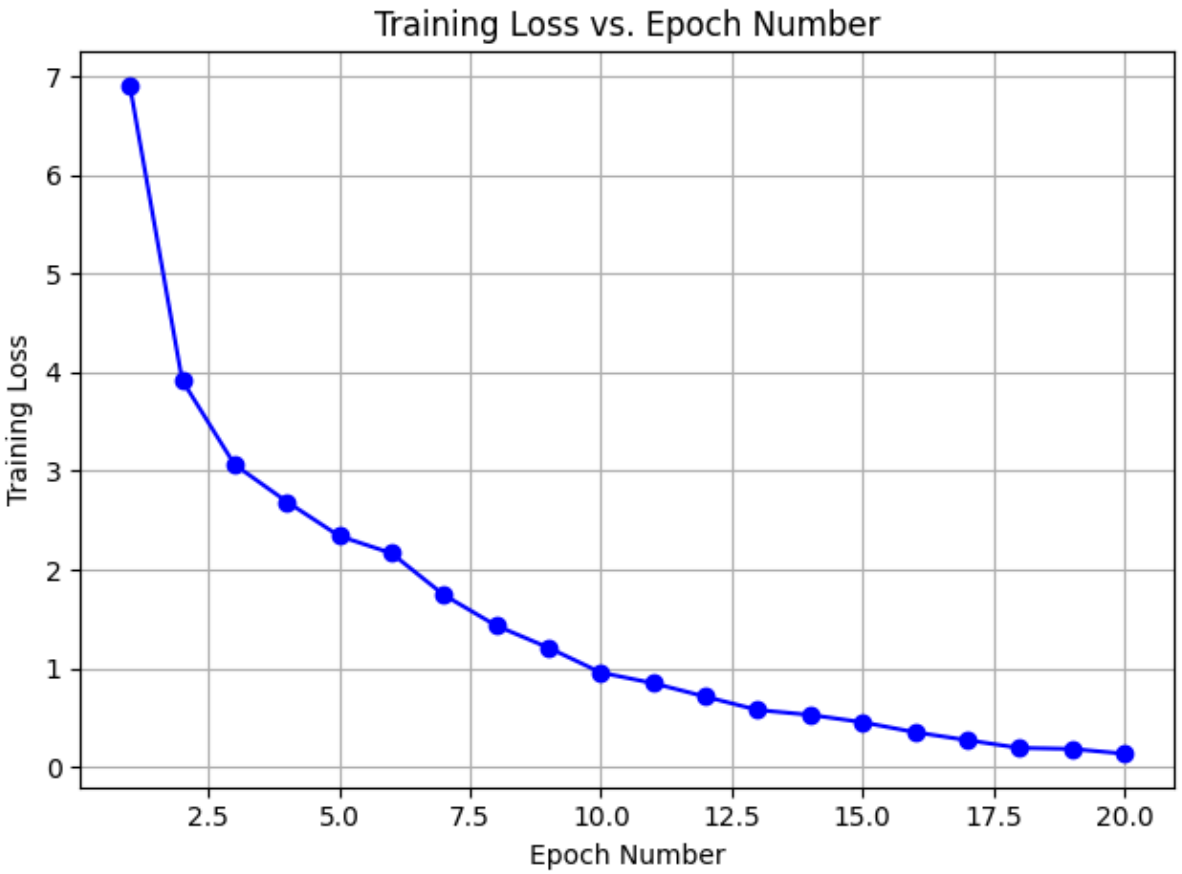

Validation Loss vs. Epoch Number
Validation Loss
5.0
4.5
4.0
3.5
3.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epoch Number

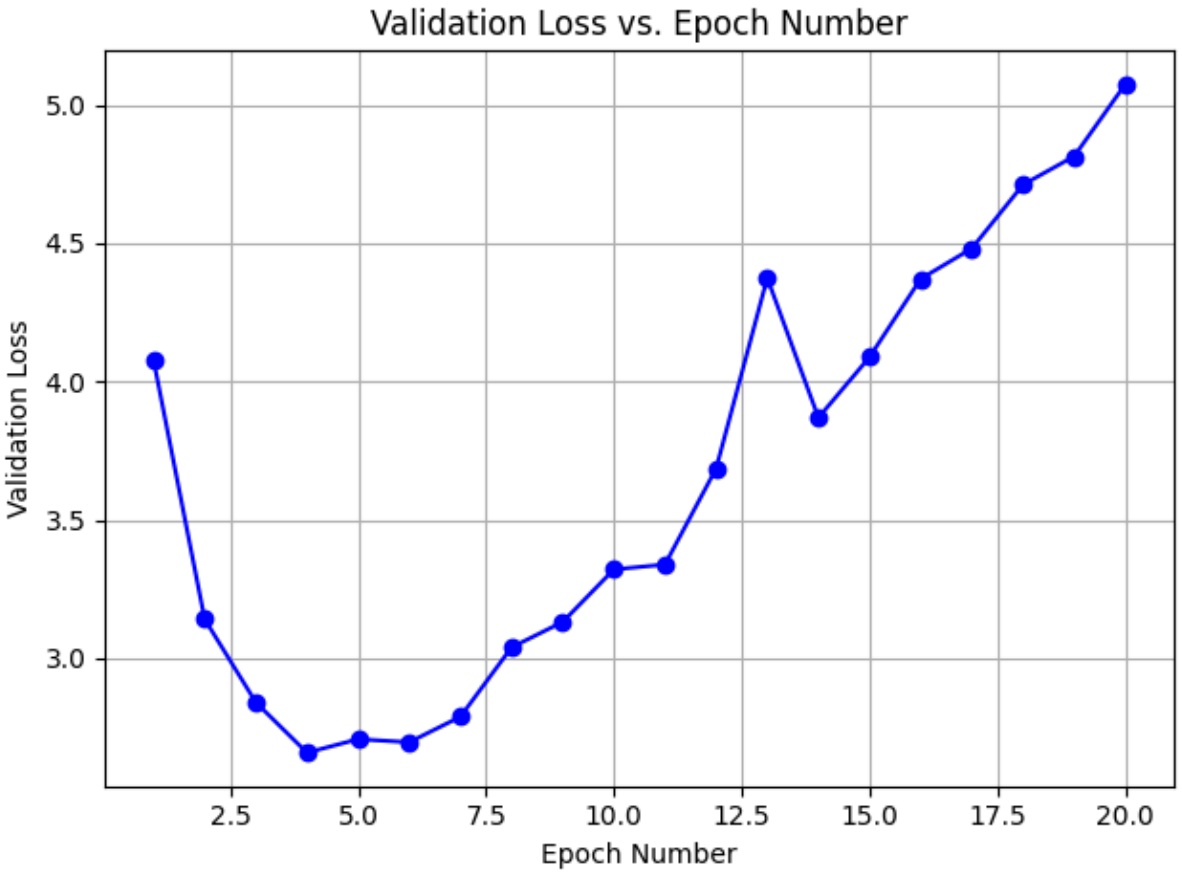

Training F1 vs. Epoch Number
Training F1
1.00
0.95
0.90
0.85
0.80
0.75
0.70
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epoch Number

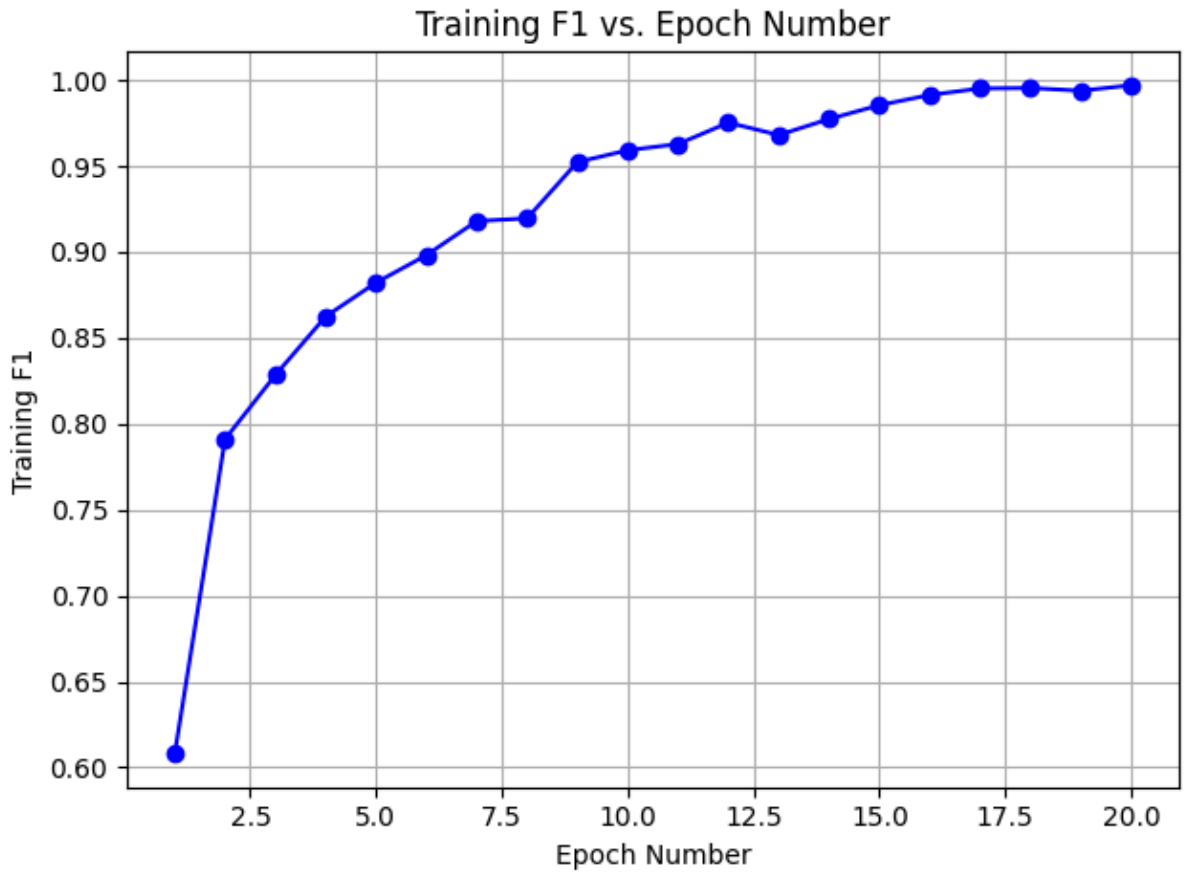


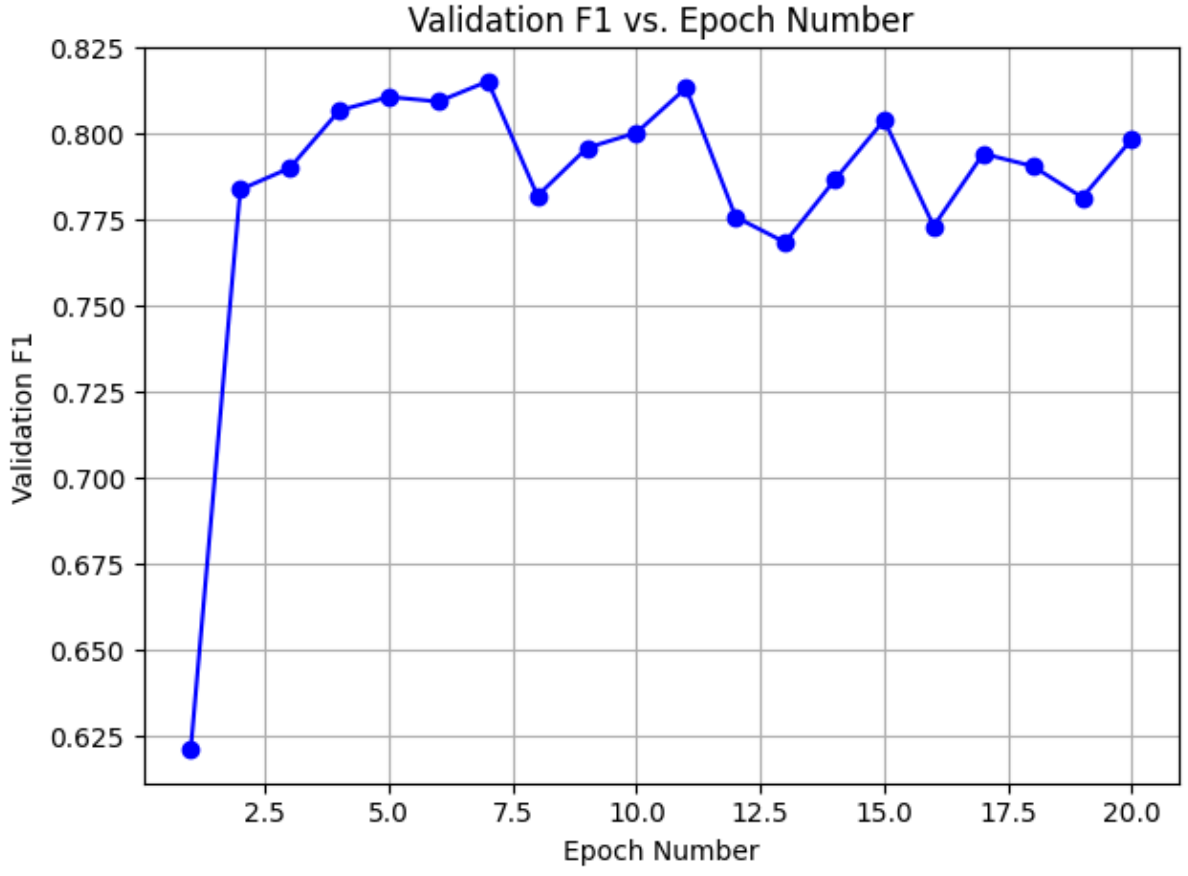
The accuracy of the model is: 0.9305743569909108
The Macro-F1 score obtained is: 0.7847591309117993

LSTM_t2_word2vec.ipynb


Training Loss vs. Epoch Number
Training Loss
Epoch Number
0
1
2
3
4
5
6
7
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

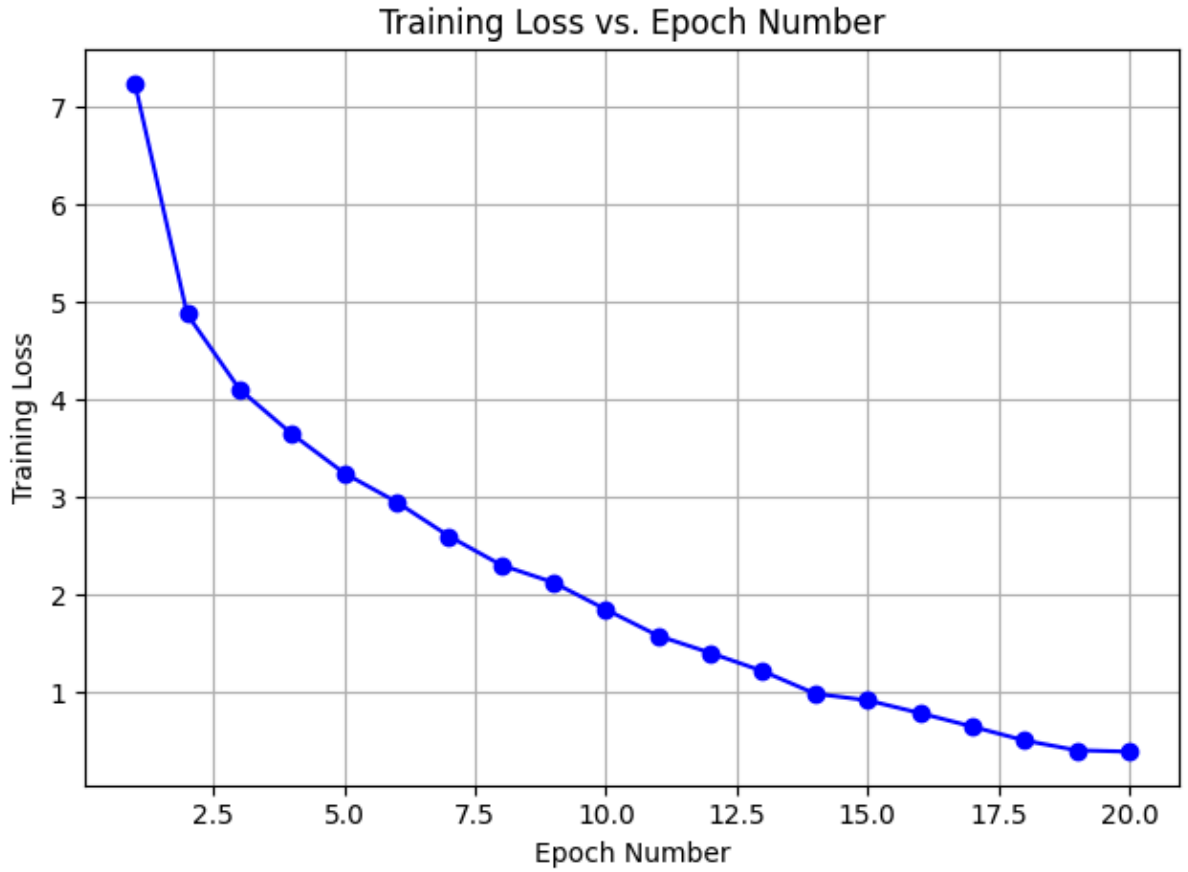

Validation Loss vs. Epoch Number
Validation Loss
Epoch Number
3.0
3.5
4.0
4.5
5.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

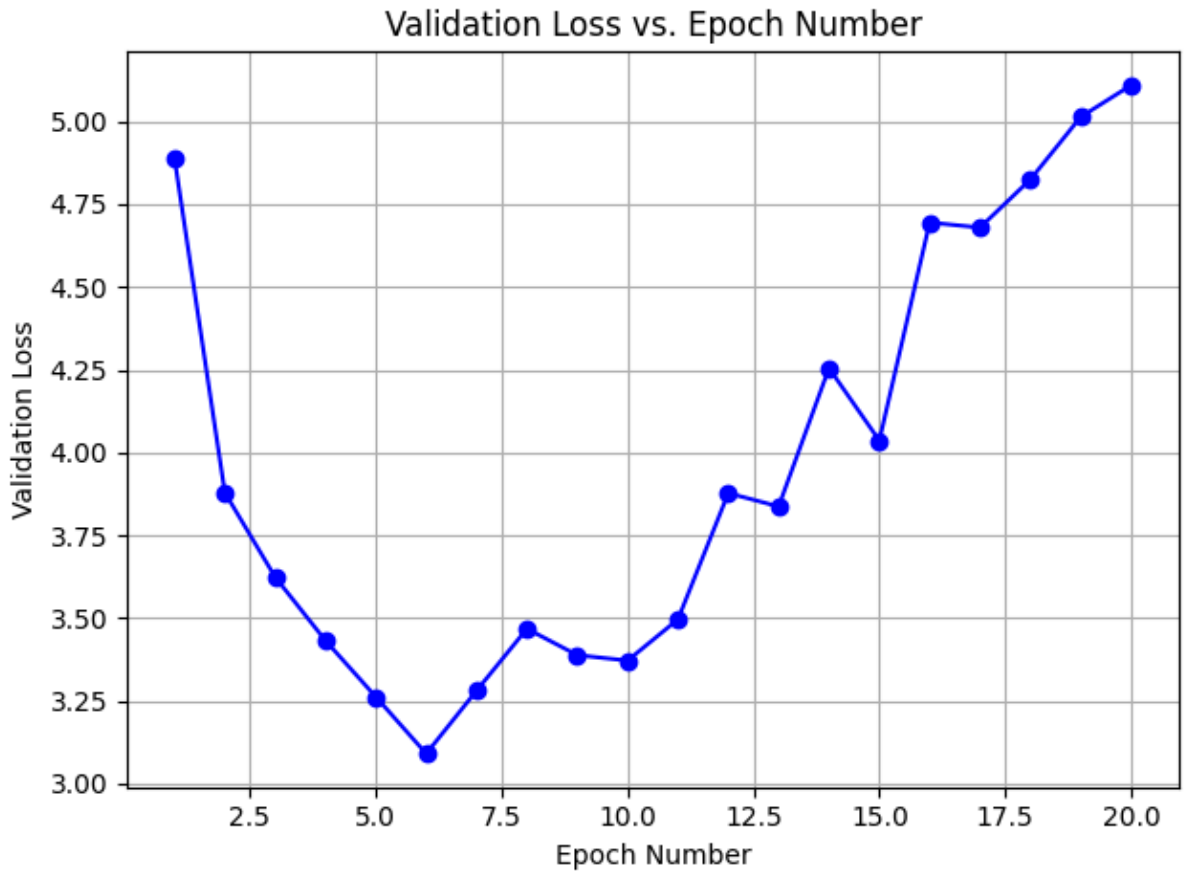

Training F1 vs. Epoch Number
Training F1
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epoch Number

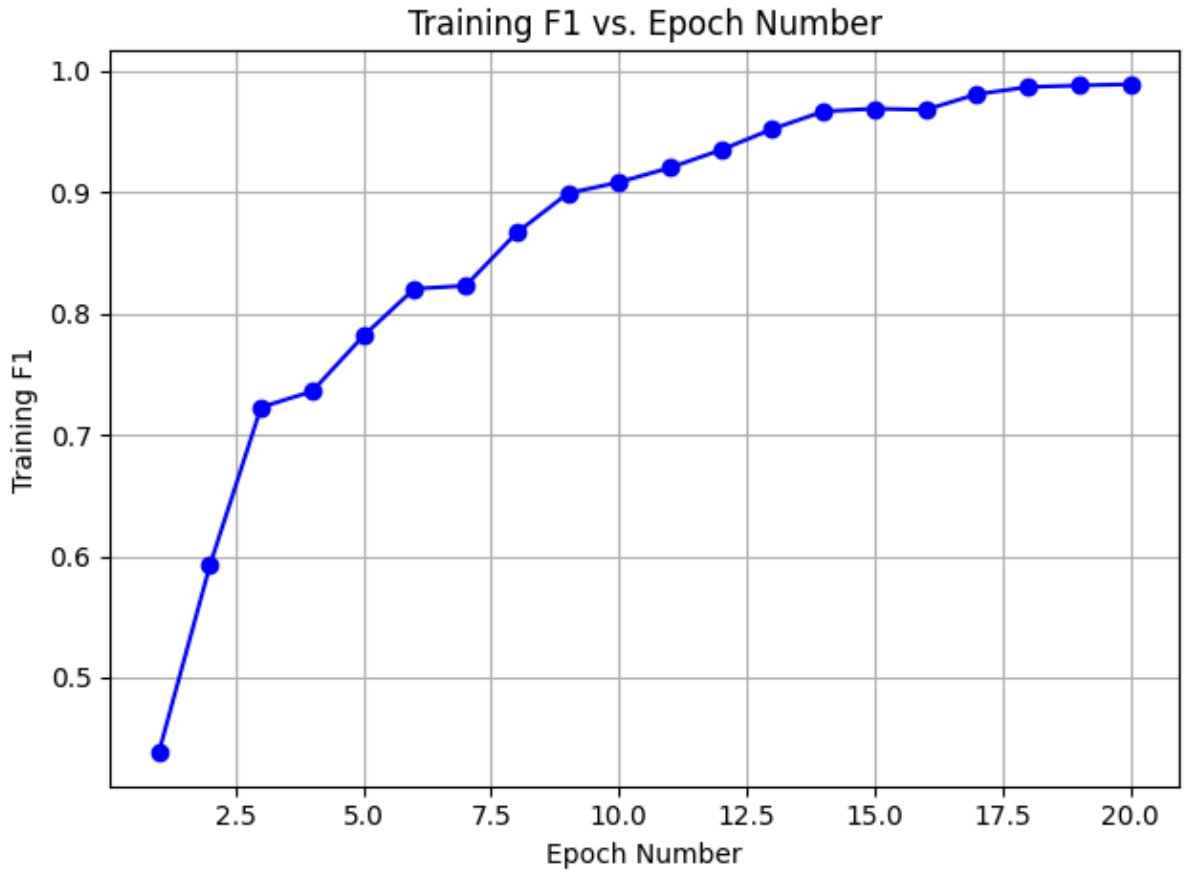

Validation F1 vs. Epoch Number
Validation F1
0.825
0.800
0.775
0.750
0.725
0.700
0.675
0.650
0.625
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epoch Number

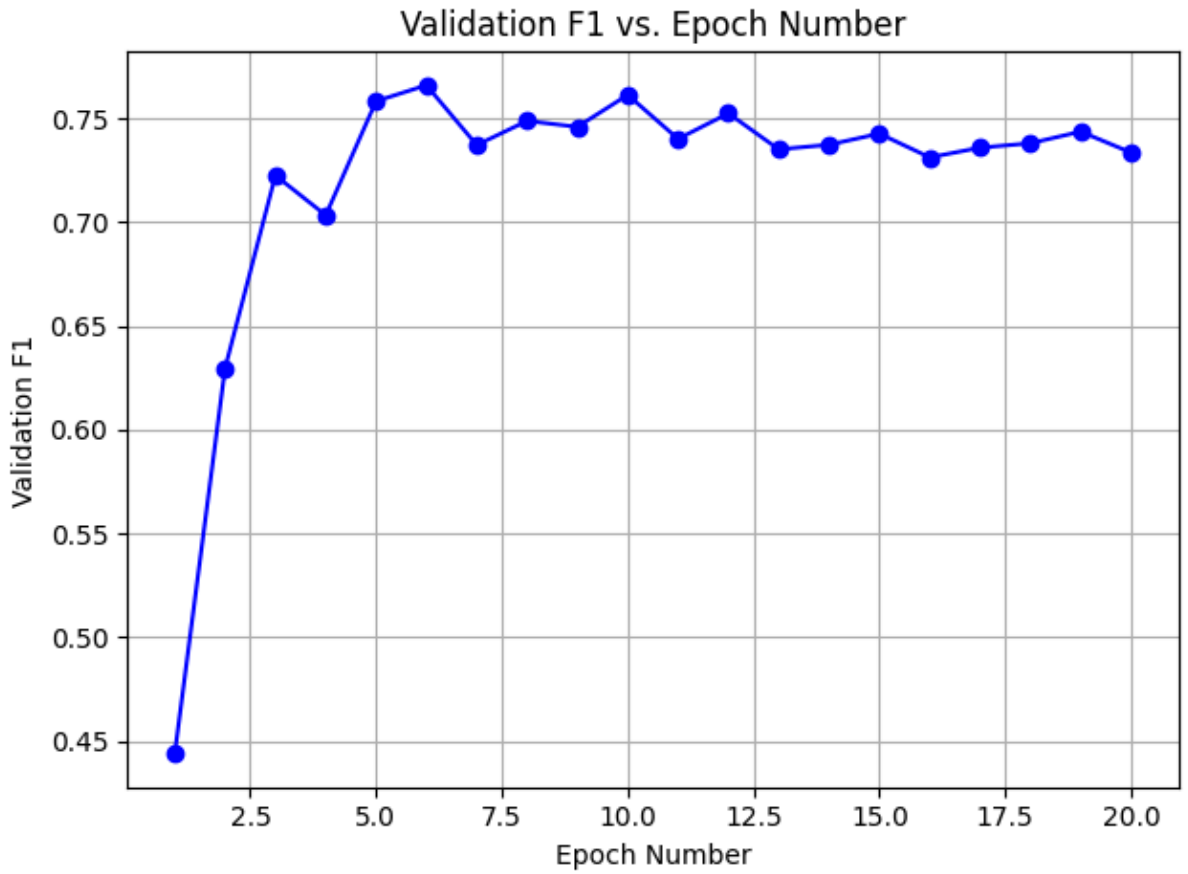
The accuracy of the model is: 0.9284471088764262
The Macro-F1 score obtained is: 0.7696524633352634

LSTM_t2_glove.ipynb



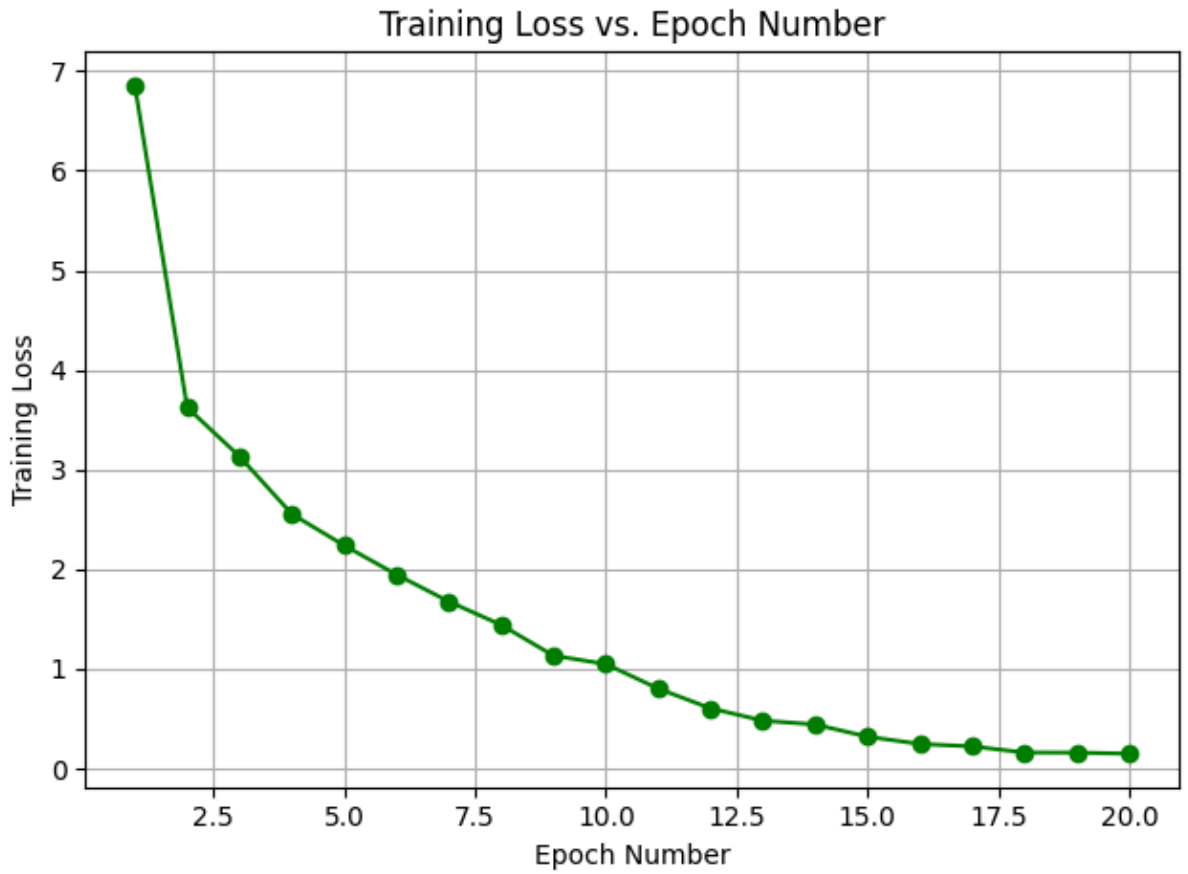

Validation Loss vs. Epoch Number
Validation Loss
Epoch Number
5.00
4.75
4.50
4.25
4.00
3.75
3.50
3.25
3.00
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

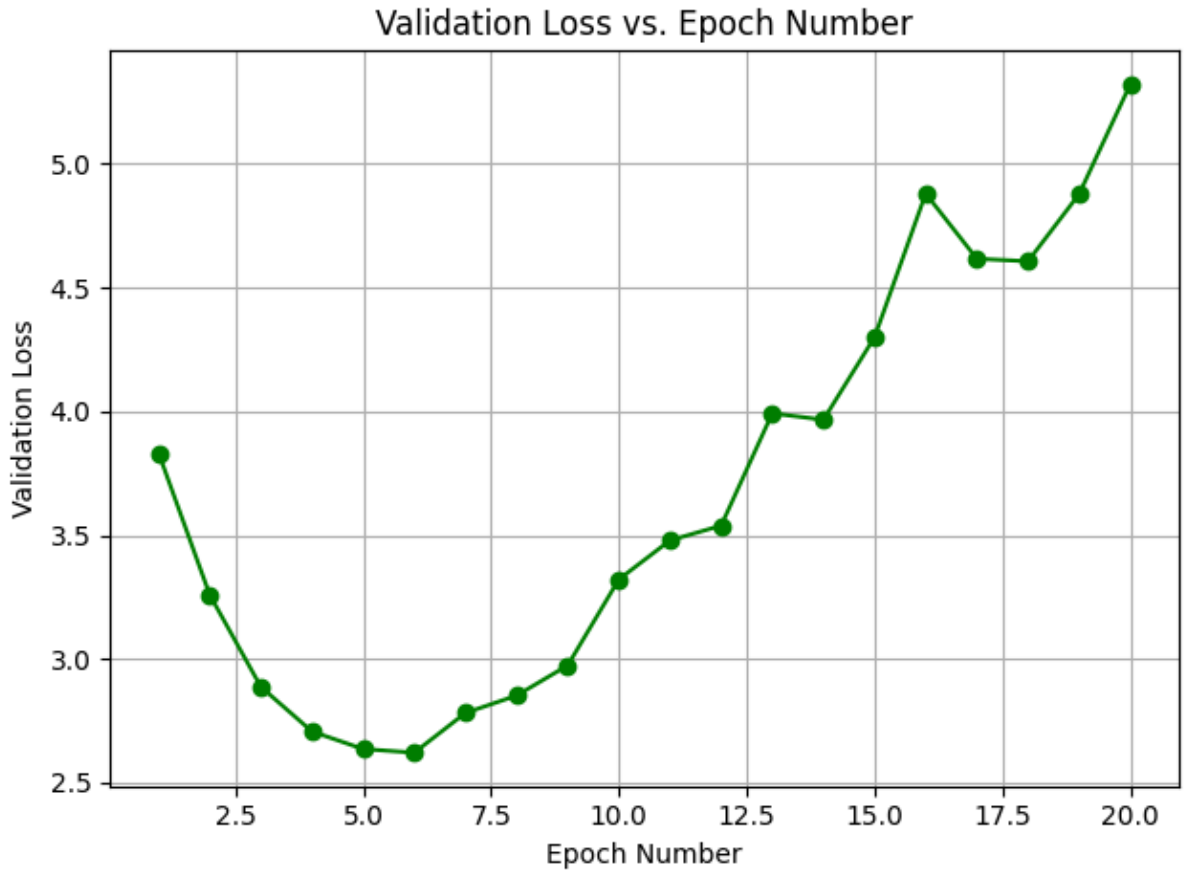

Training F1 vs. Epoch Number
Training F1
Epoch Number
1.0
0.9
0.8
0.7
0.6
0.5
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

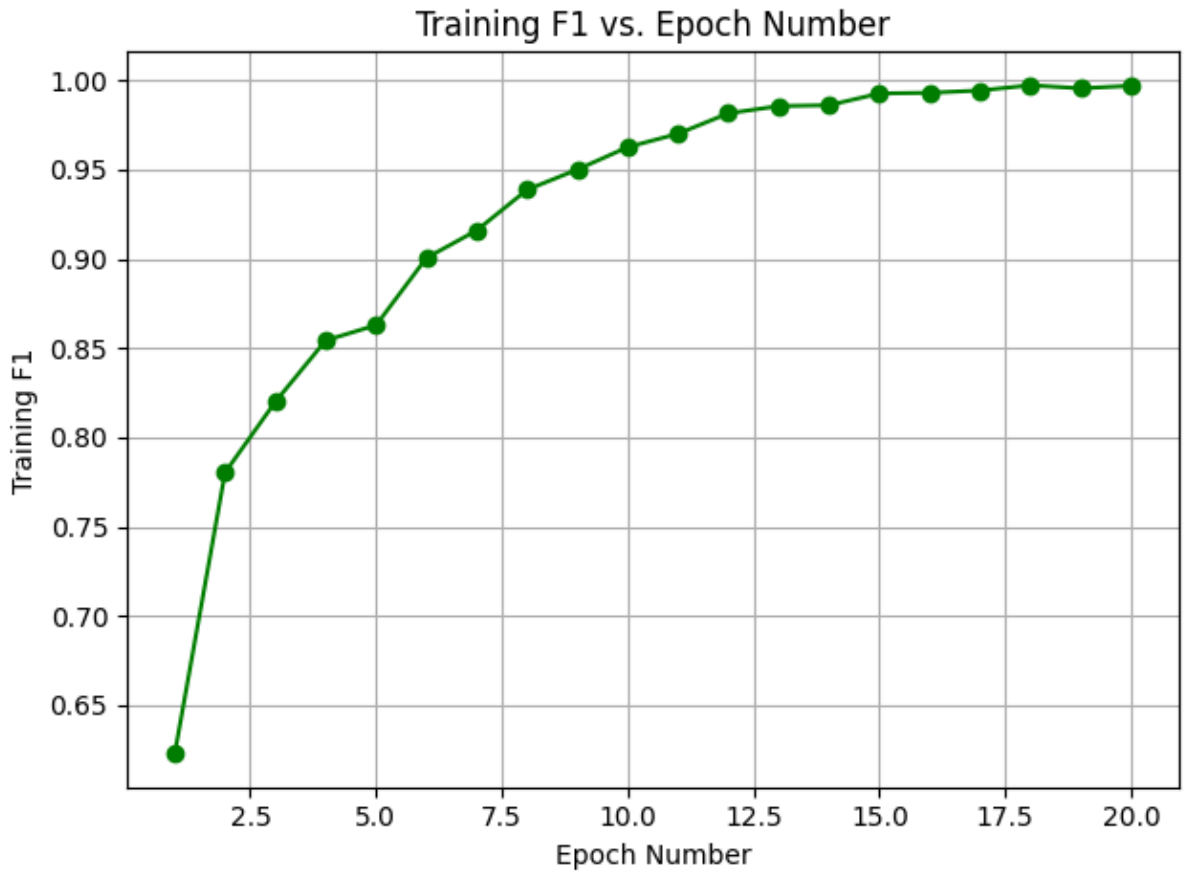


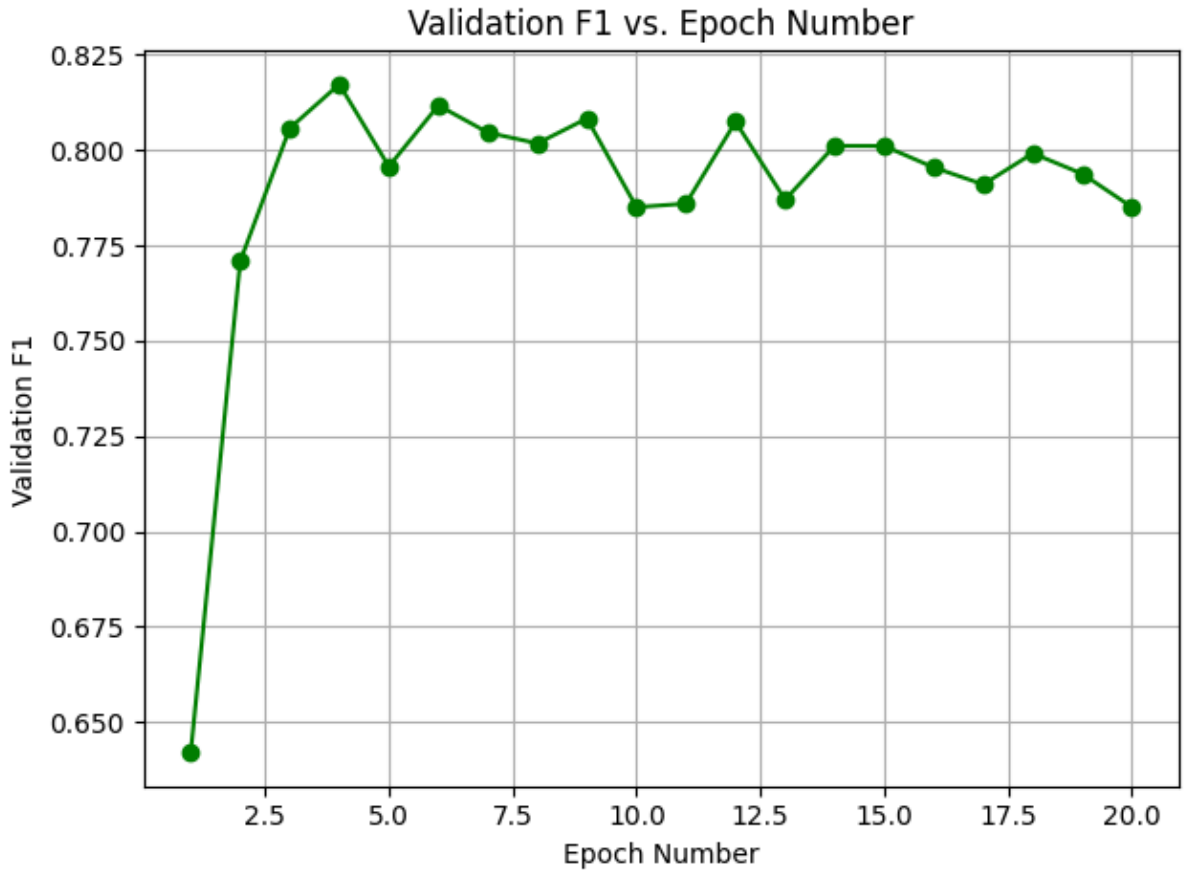
The accuracy of the model is: 0.9073680139238058
The Macro-F1 score obtained is: 0.6979182752229057

LSTM_t2_fasttext.ipynb


Training Loss vs. Epoch Number
Training Loss
Epoch Number
0
1
2
3
4
5
6
7
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0


Validation Loss vs. Epoch Number
Validation Loss
Epoch Number
2.5
3.0
3.5
4.0
4.5
5.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

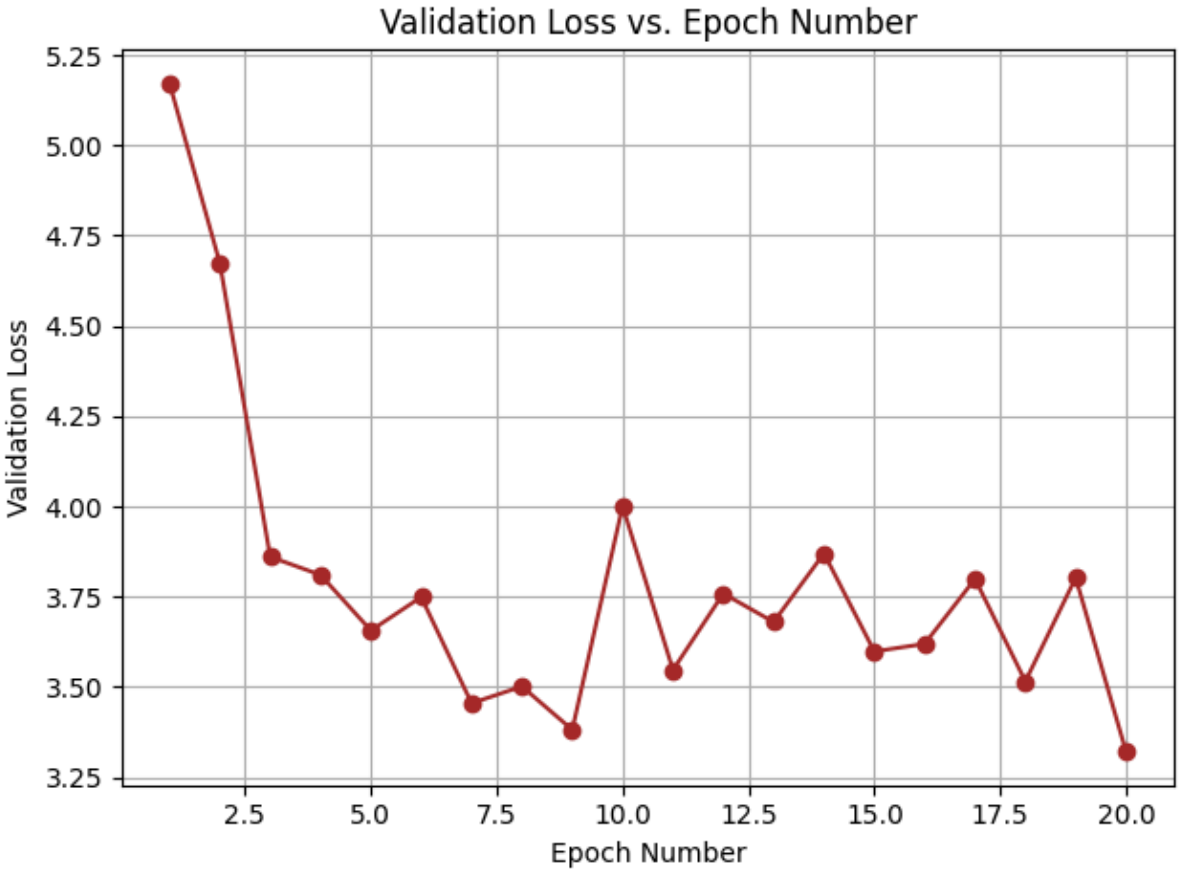




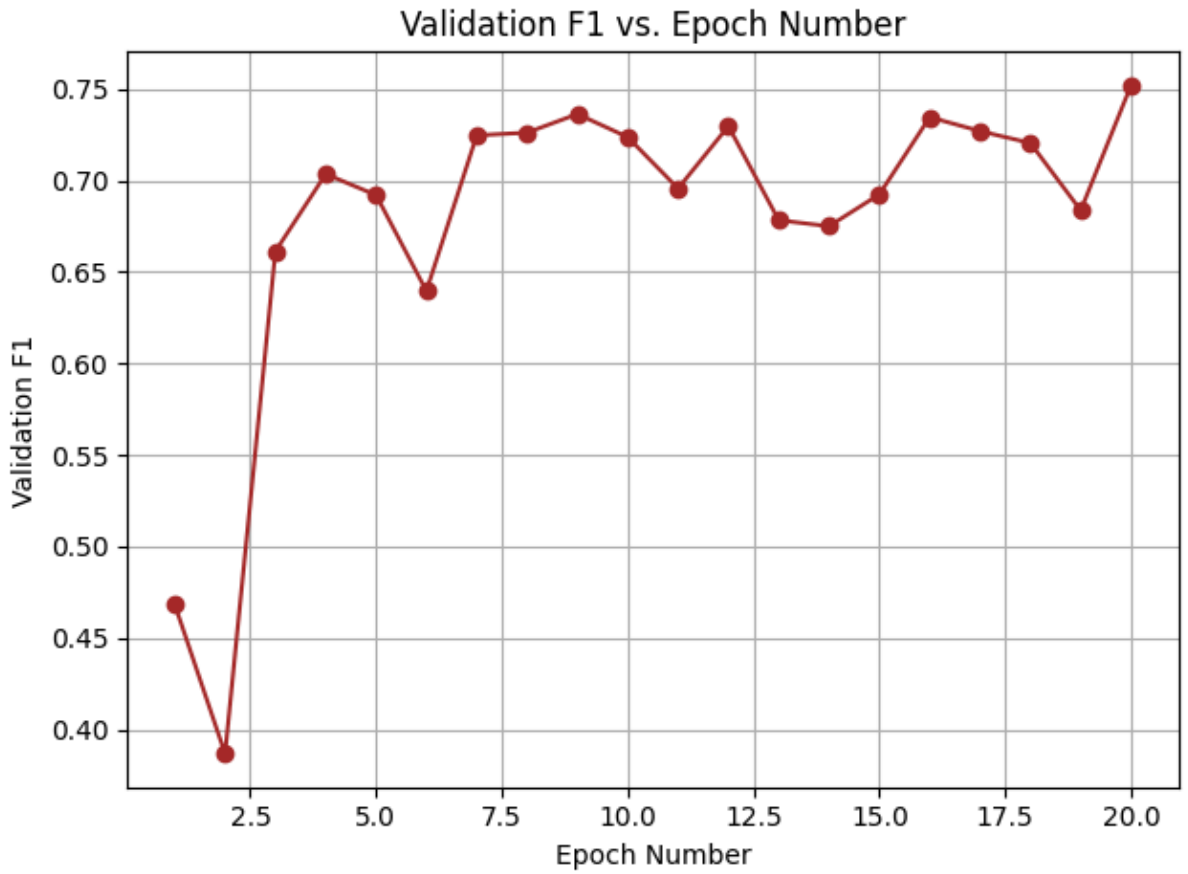
The accuracy of the model is: 0.9272867917230709

The Macro-F1 score obtained is: 0.7643203585344831

RNN_glove_task2.ipynb

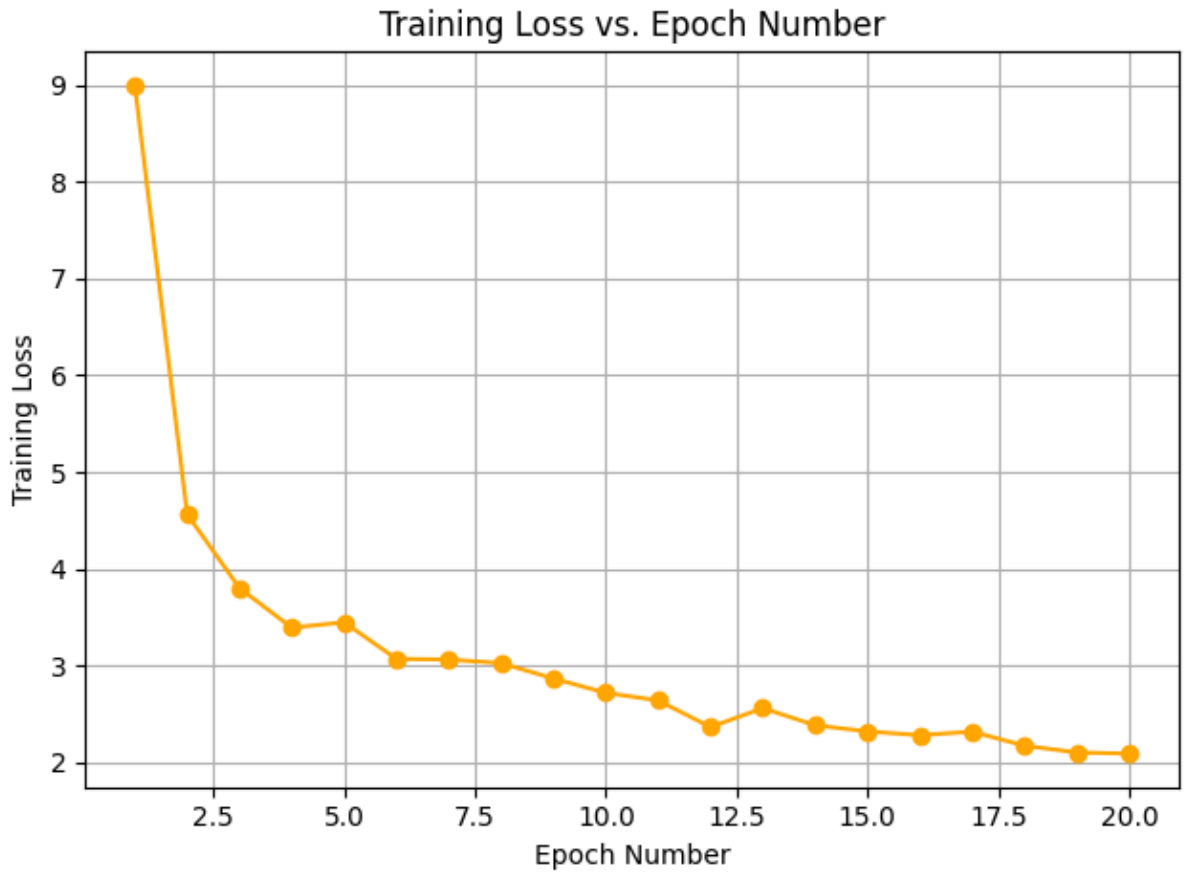

Validation Loss vs. Epoch Number
Validation Loss
Epoch Number
5.25
5.00
4.75
4.50
4.25
4.00
3.75
3.50
3.25
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

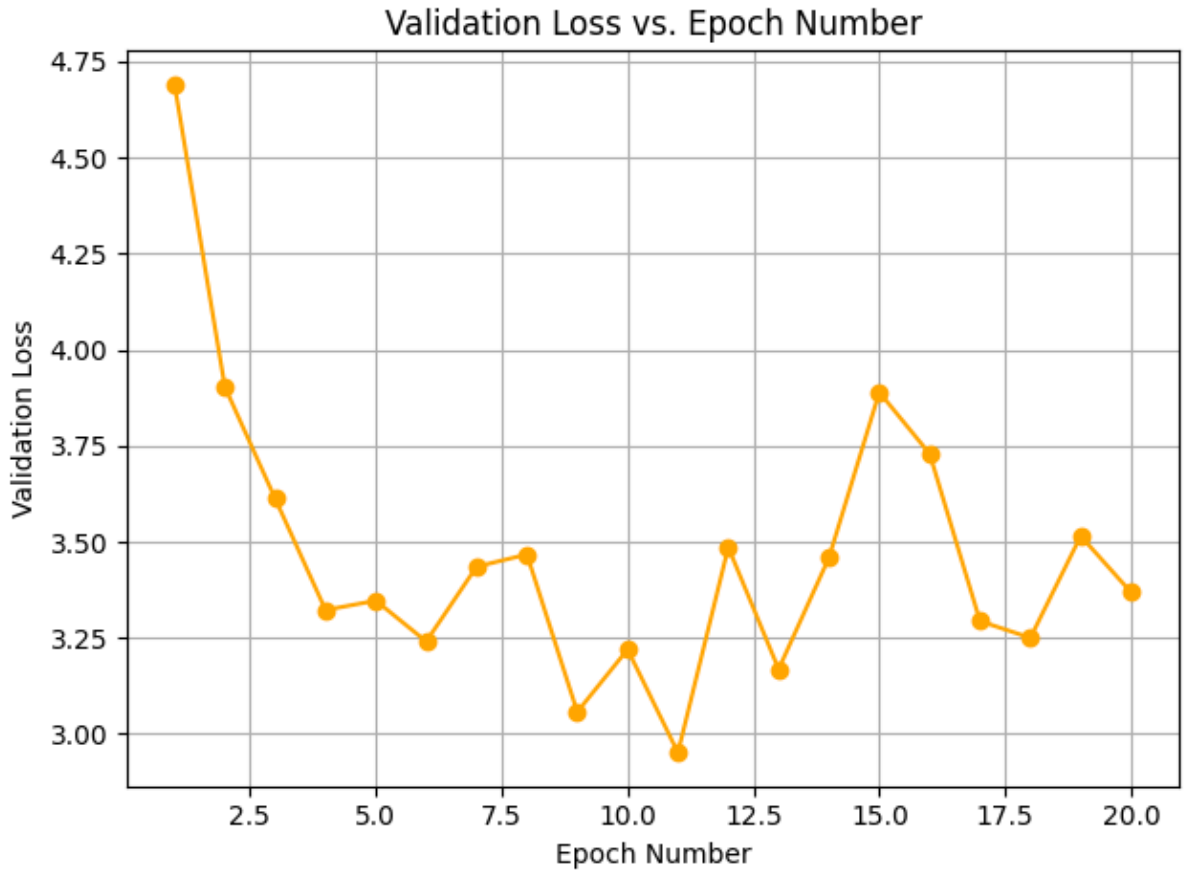

Training F1 vs. Epoch Number
Training F1
Epoch Number
0.8
0.7
0.6
0.5
0.4
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

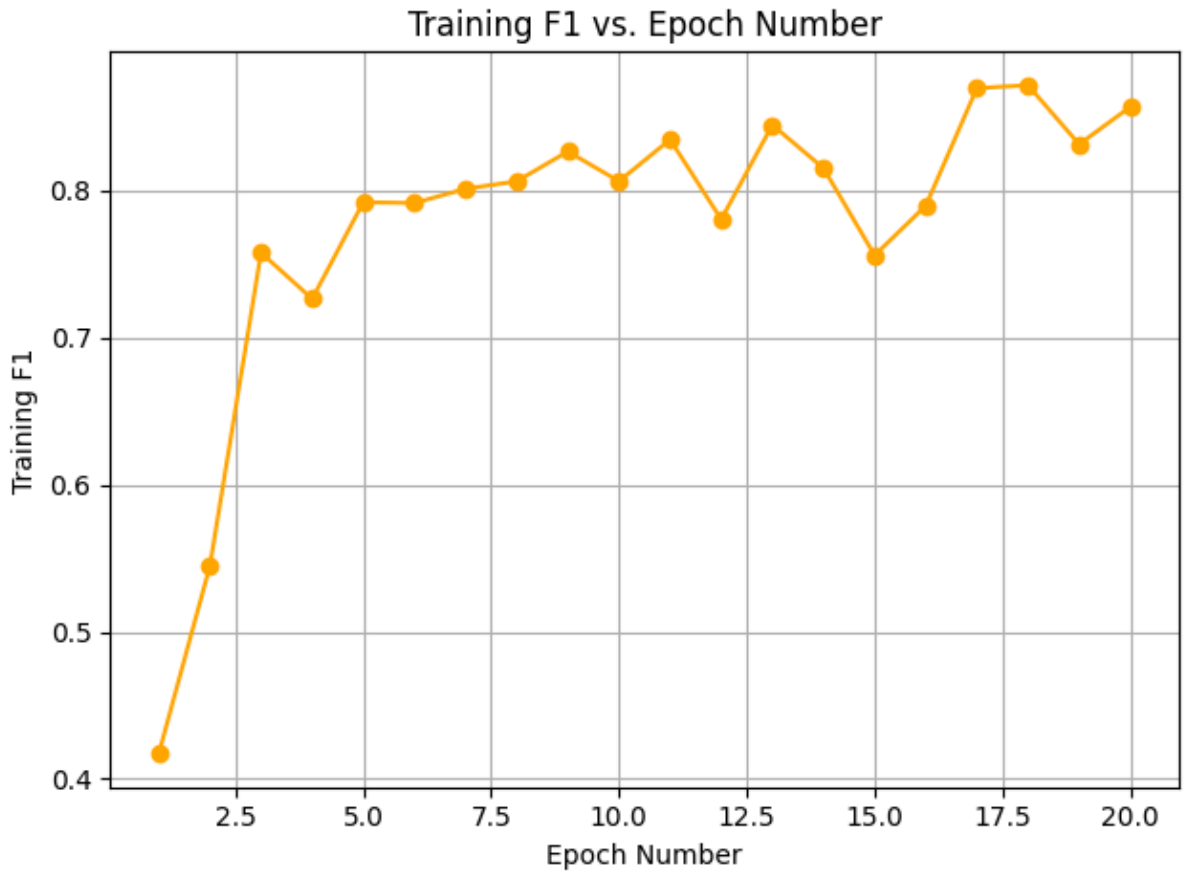


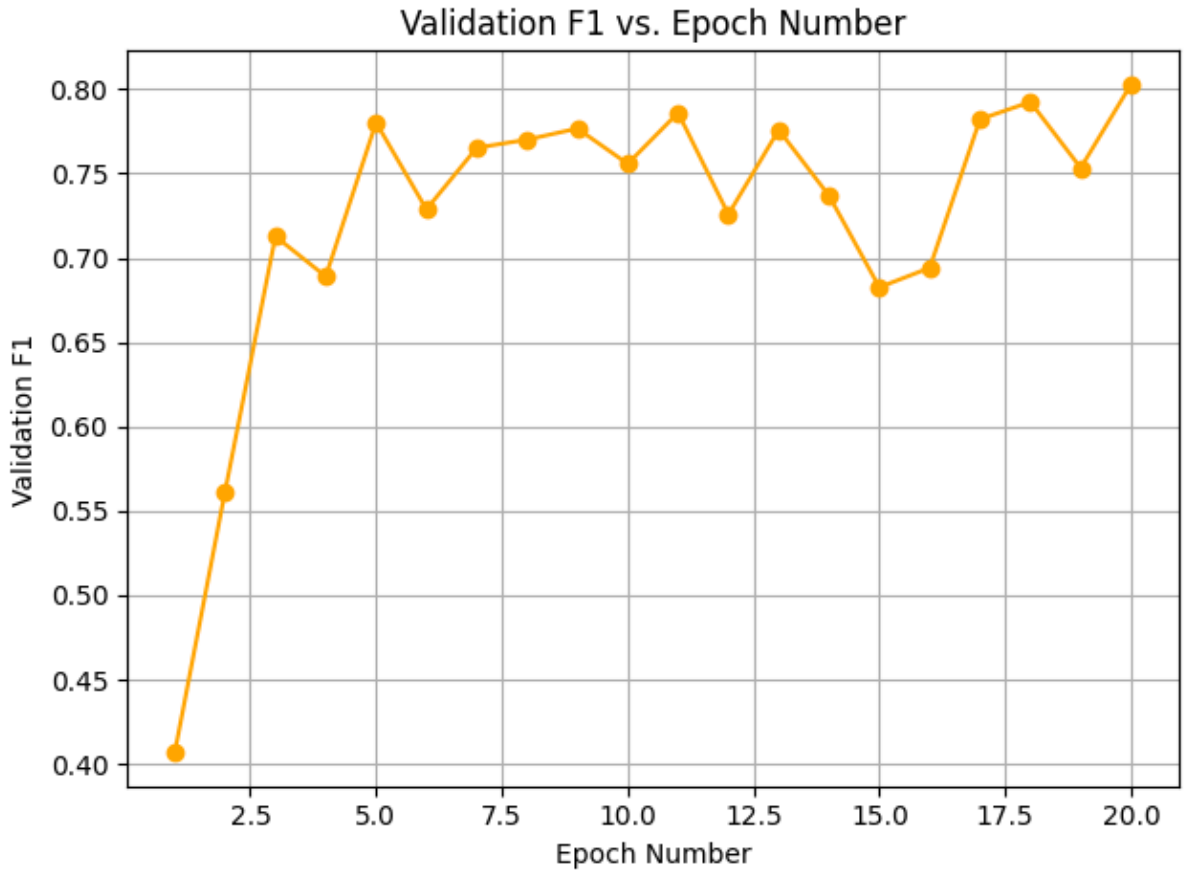
The accuracy of the model is: 0.9031135176948366
The Macro-F1 score obtained is: 0.659785788424339

RNN_fasttext_task2.ipynb


Training Loss vs. Epoch Number
Training Loss
Epoch Number
9
8
7
6
5
4
3
2
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0


Validation Loss vs. Epoch Number
Validation Loss
Epoch Number
4.75
4.50
4.25
4.00
3.75
3.50
3.25
3.00
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

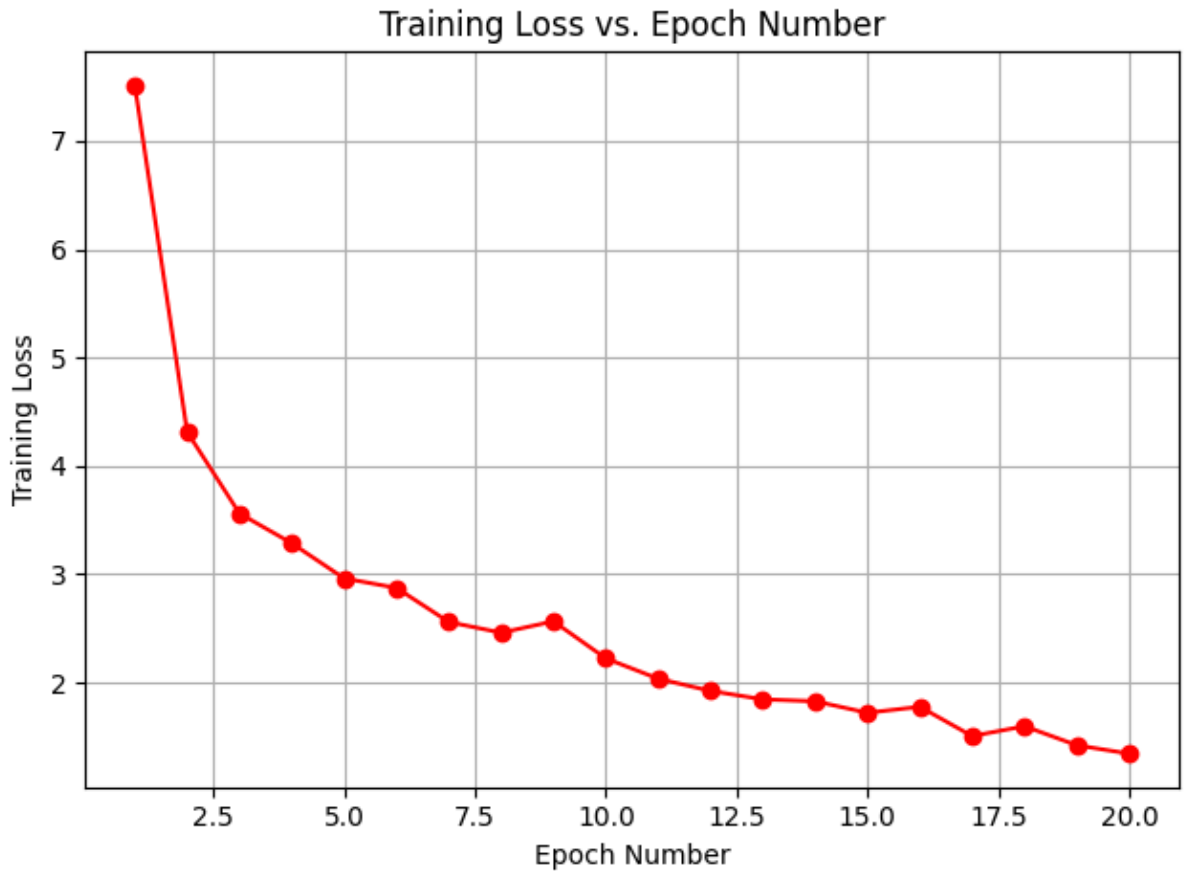


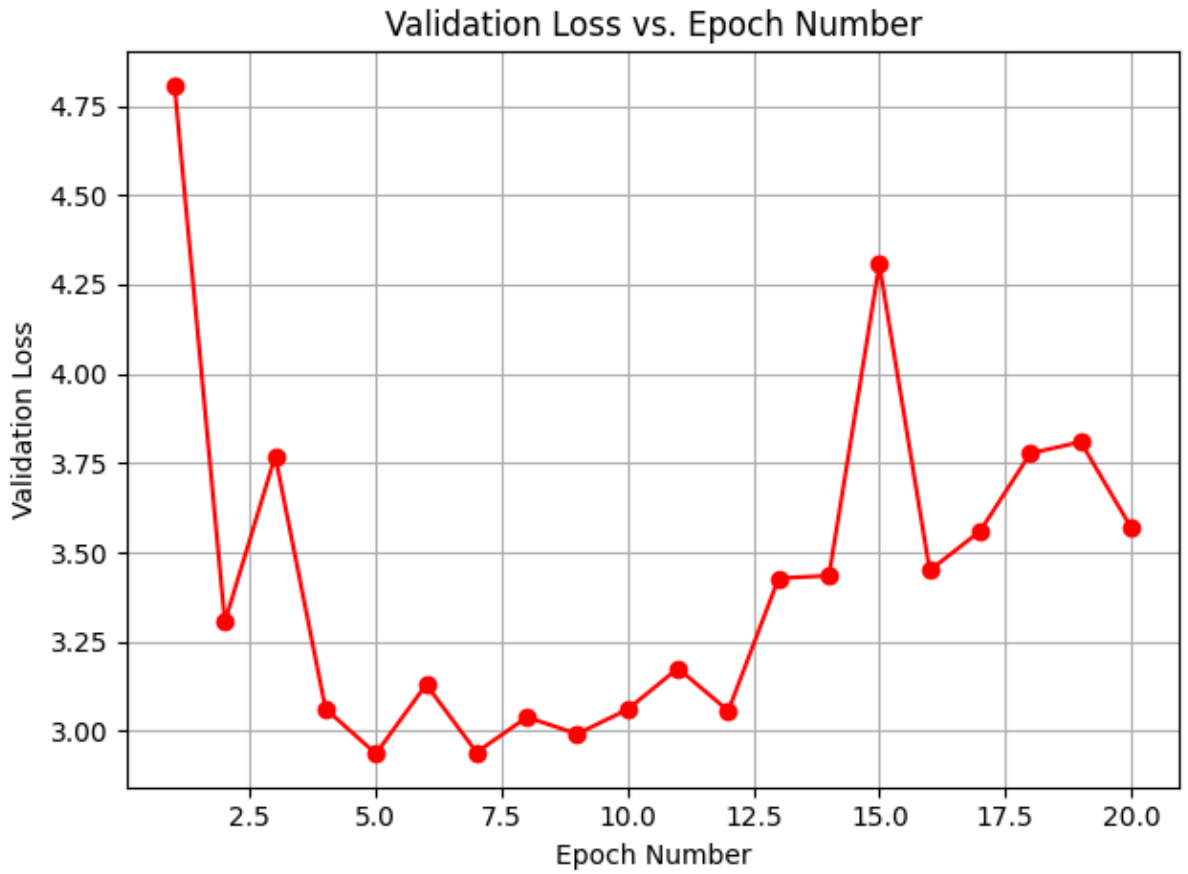


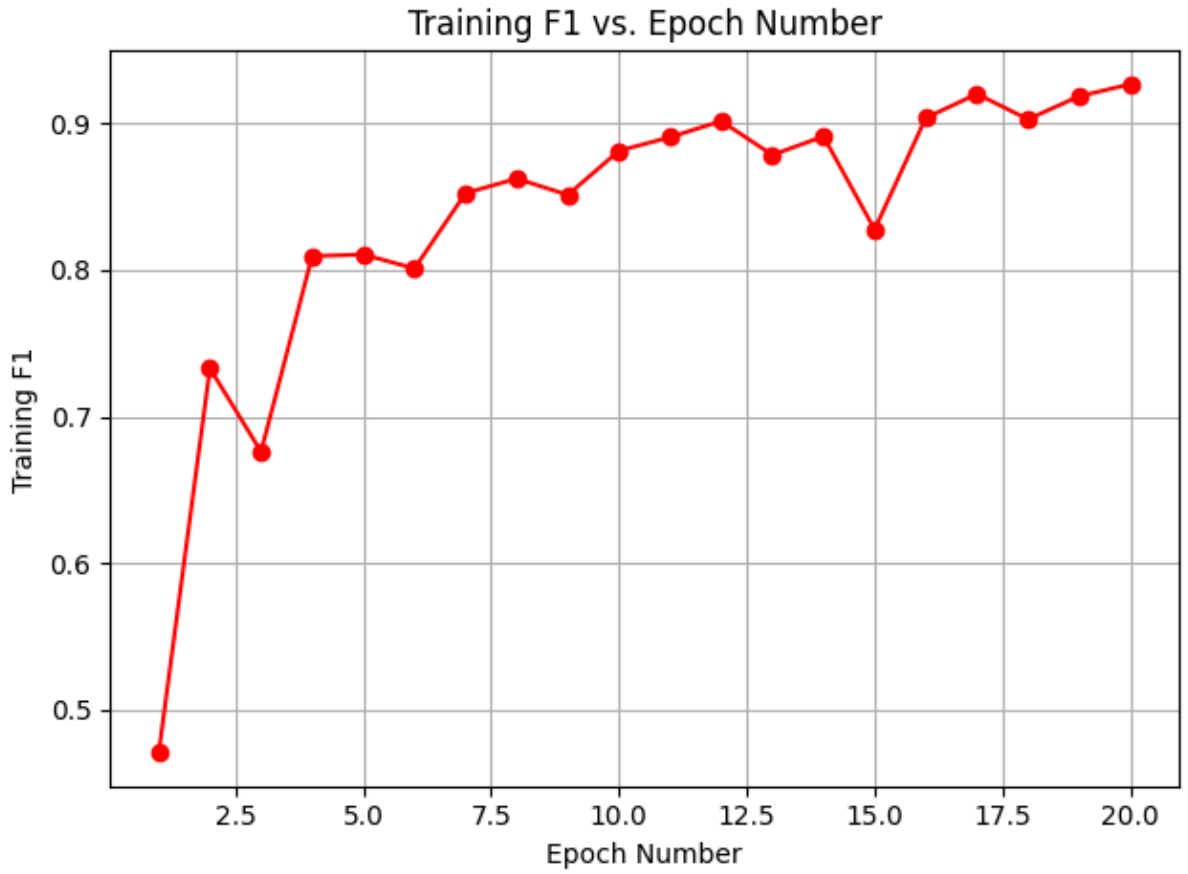
The accuracy of the model is: 0.9269000193386192

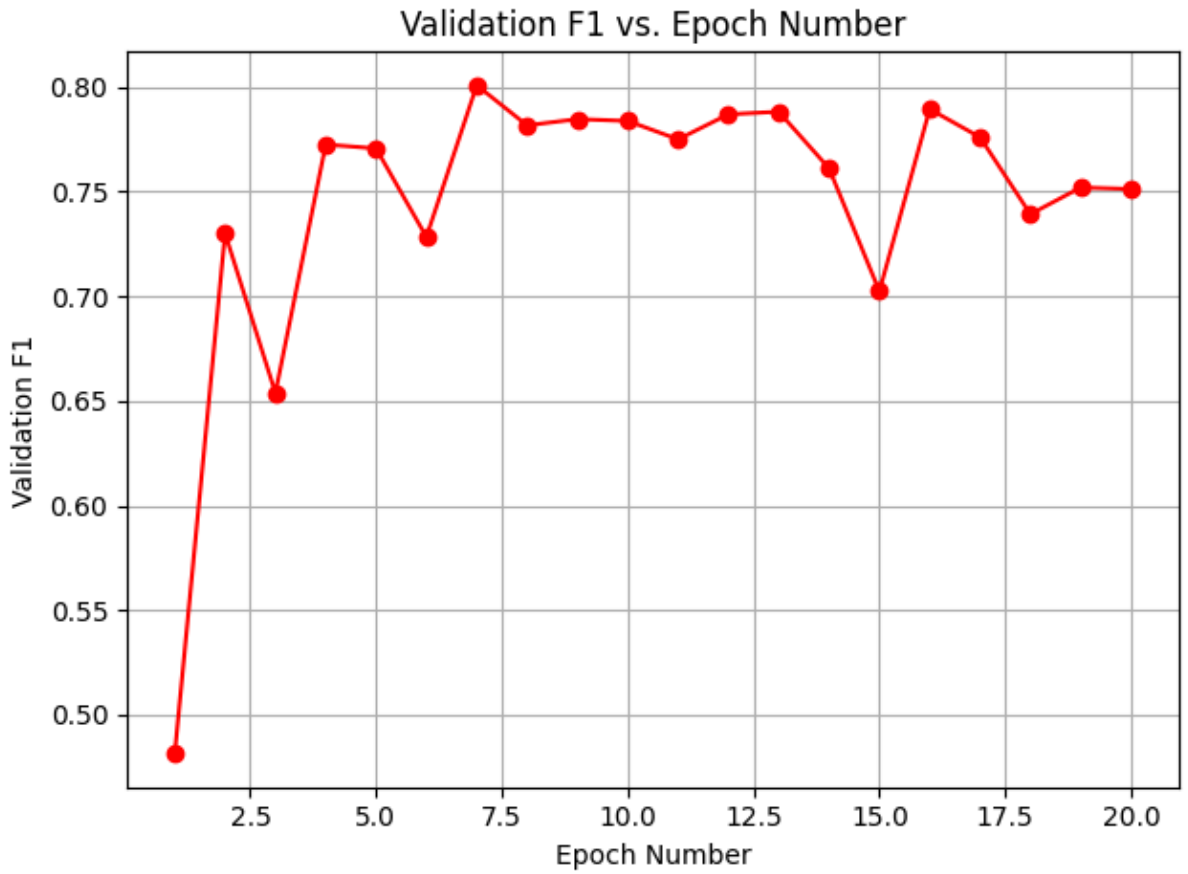
The Macro-F1 score obtained is: 0.7874690756939237

RNN_word2vec_task2.ipynb



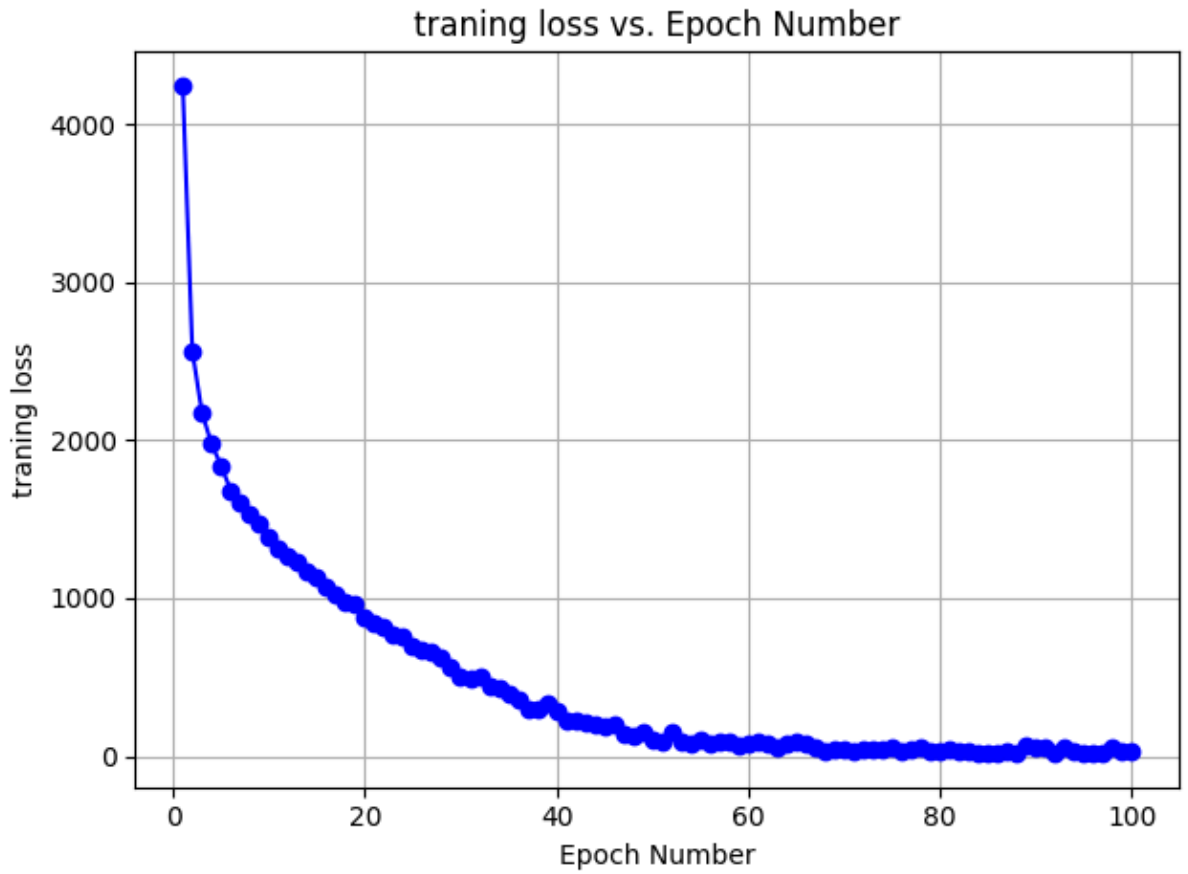

Validation Loss vs. Epoch Number
Validation Loss
4.75
4.50
4.25
4.00
3.75
3.50
3.25
3.00
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epoch Number

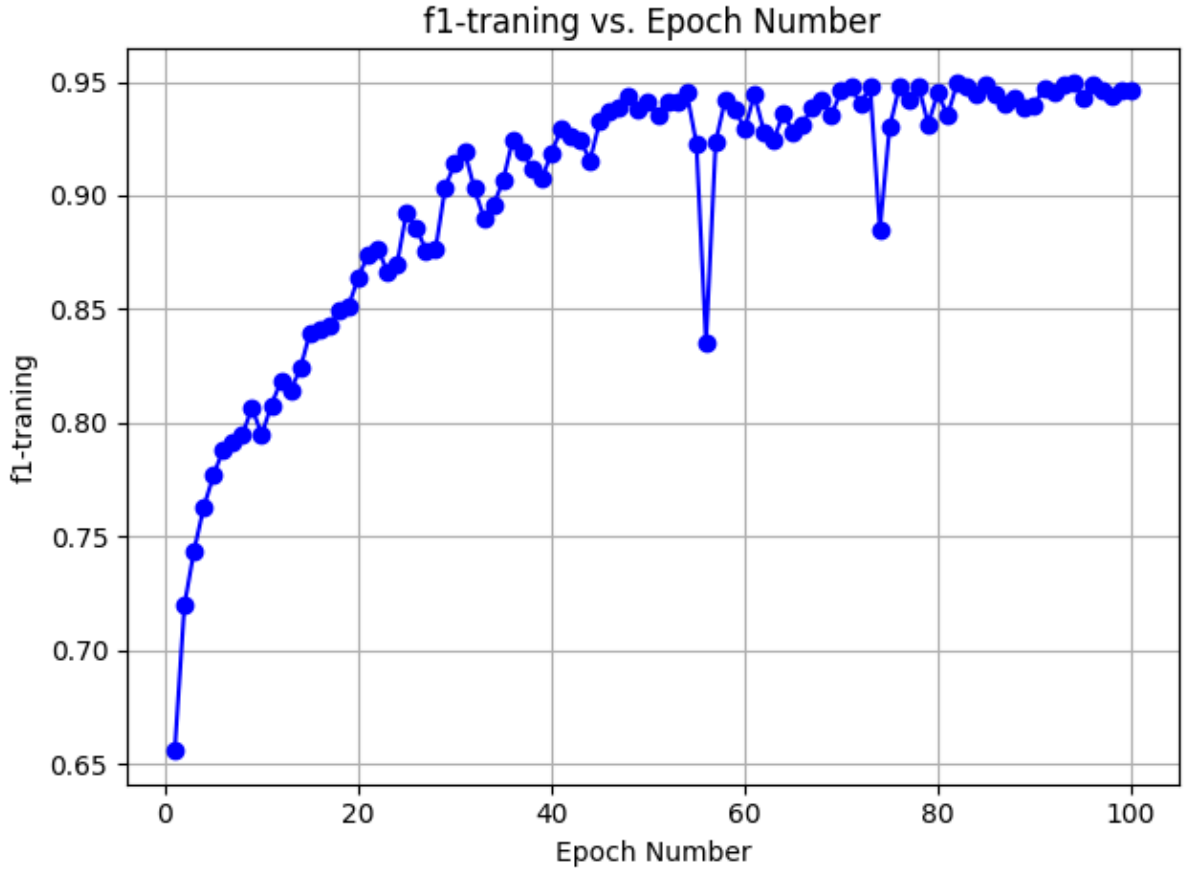

Training F1 vs. Epoch Number
Training F1
0.9
0.8
0.7
0.6
0.5
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Epoch Number

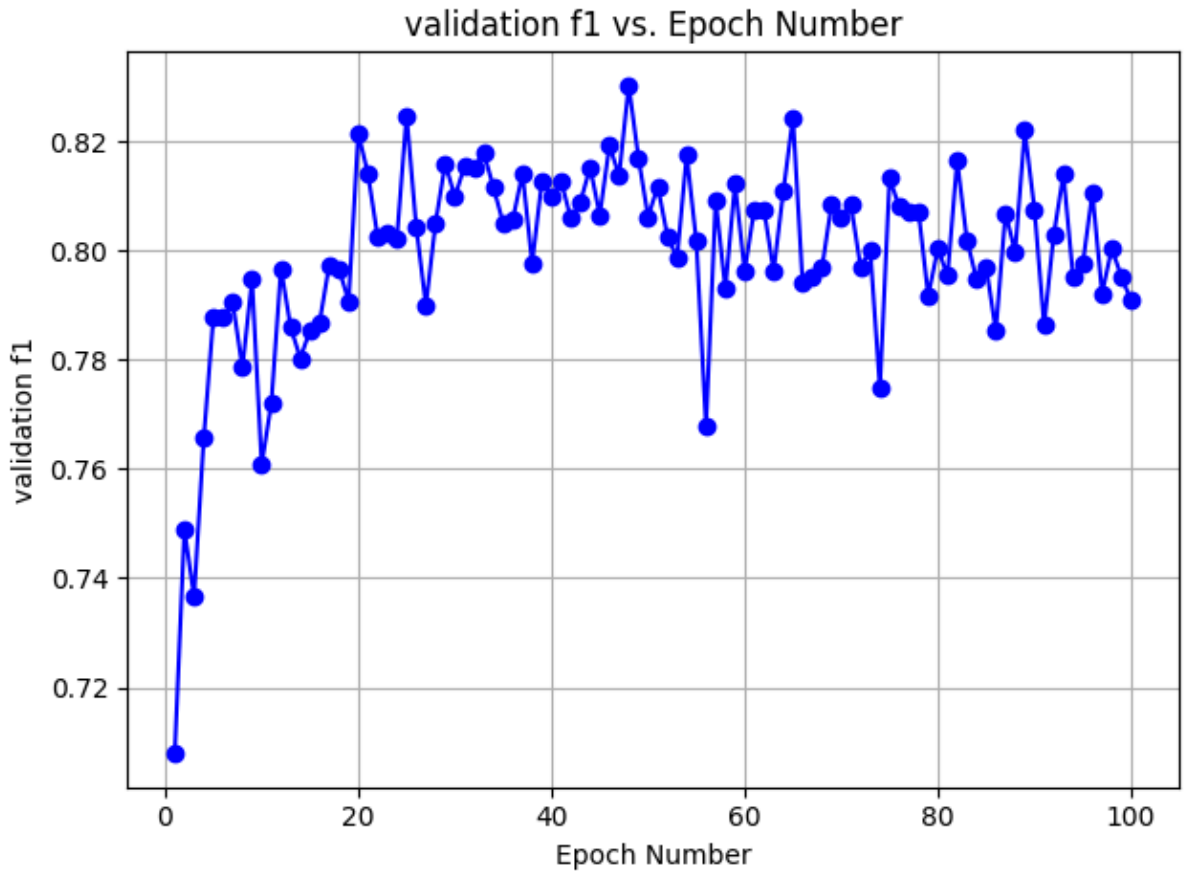


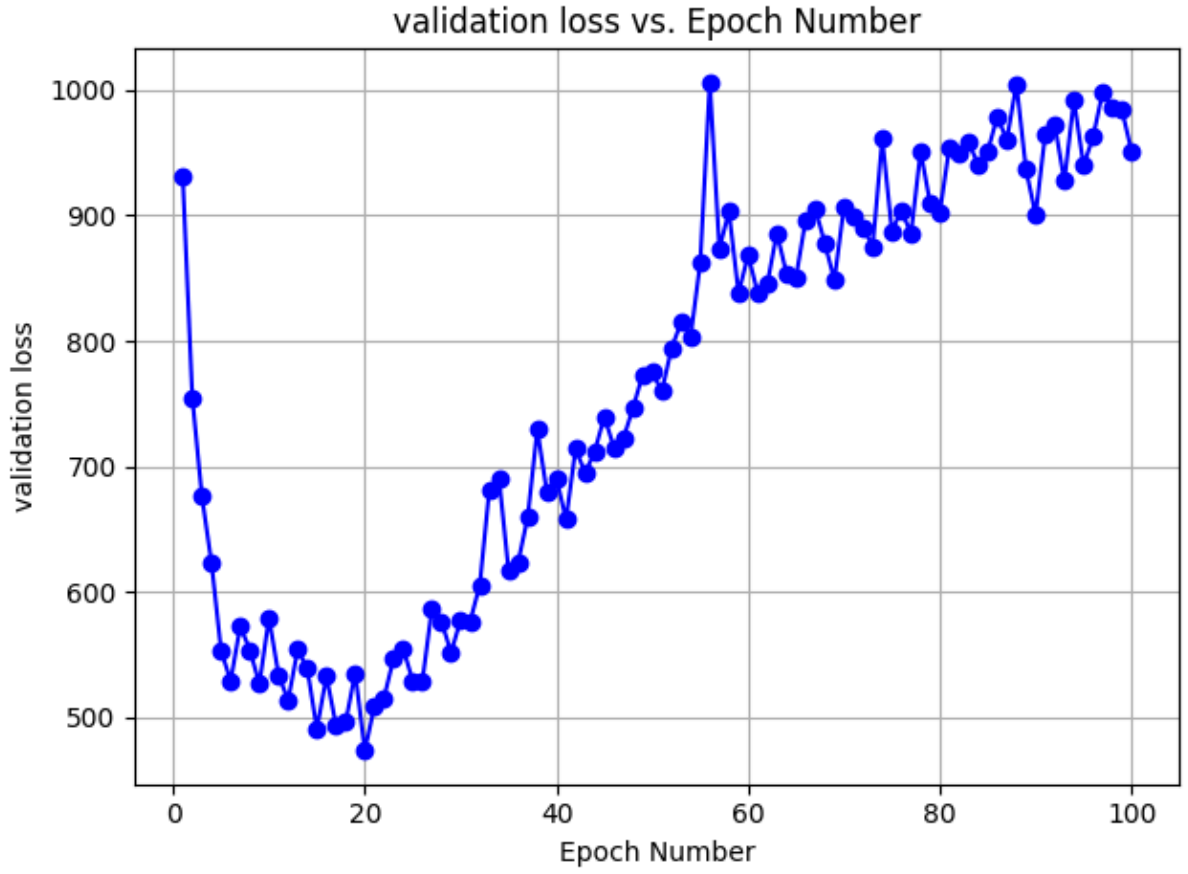
Testing f1=0.7204053804398965

Bilstm-crf(word2vec)


traning loss vs. Epoch Number
traning loss
4000
3000
2000
1000
0
0
20
40
60
80
100
Epoch Number

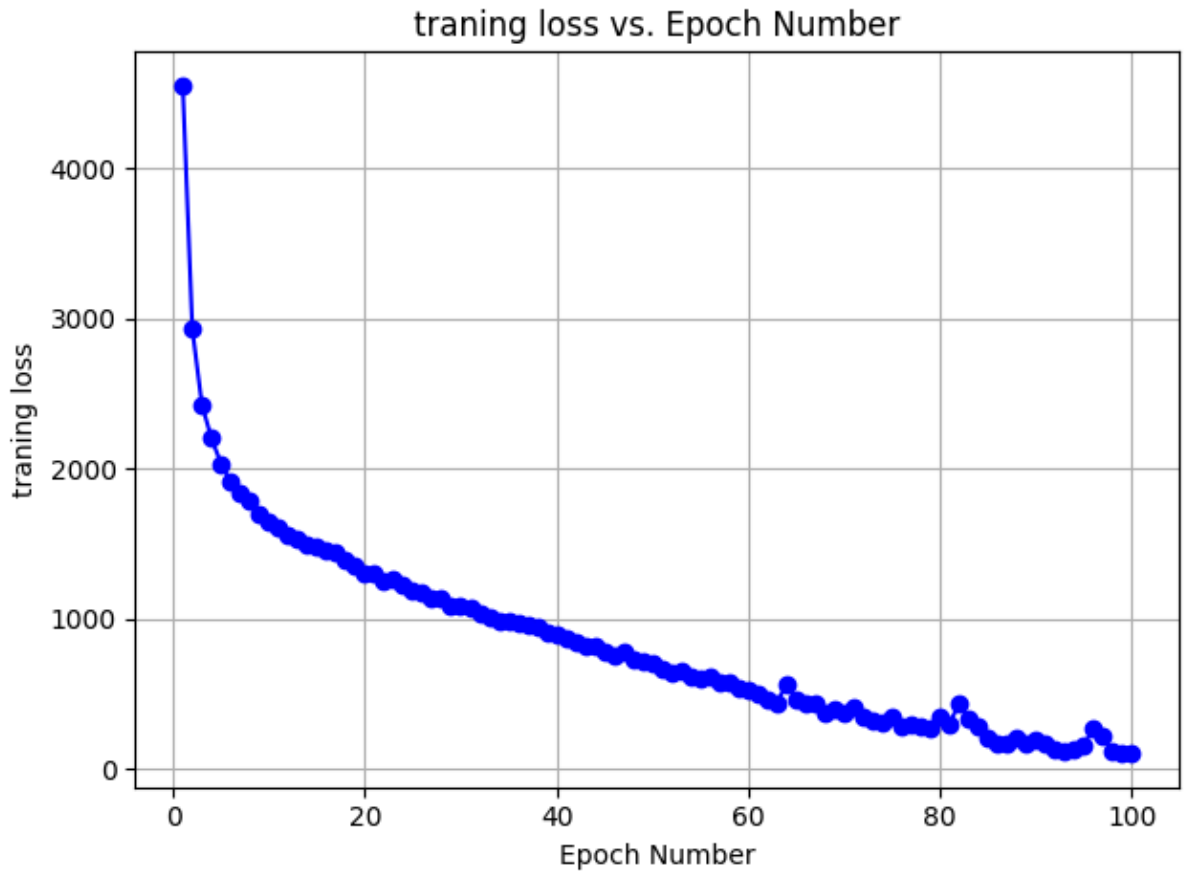

f1-traning vs. Epoch Number
f1-traning
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0
20
40
60
80
100
Epoch Number

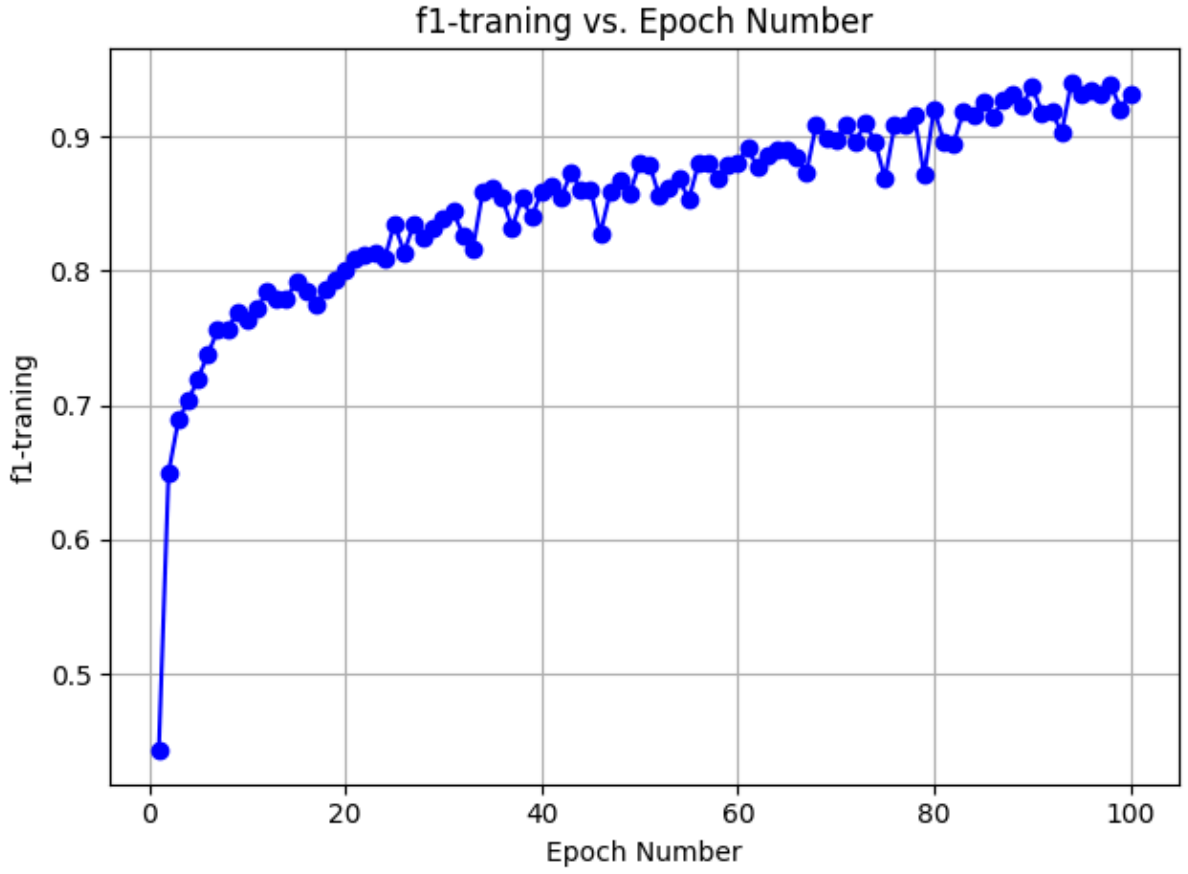

validation f1 vs. Epoch Number
validation f1
0.82
0.80
0.78
0.76
0.74
0.72
0
20
40
60
80
100
Epoch Number

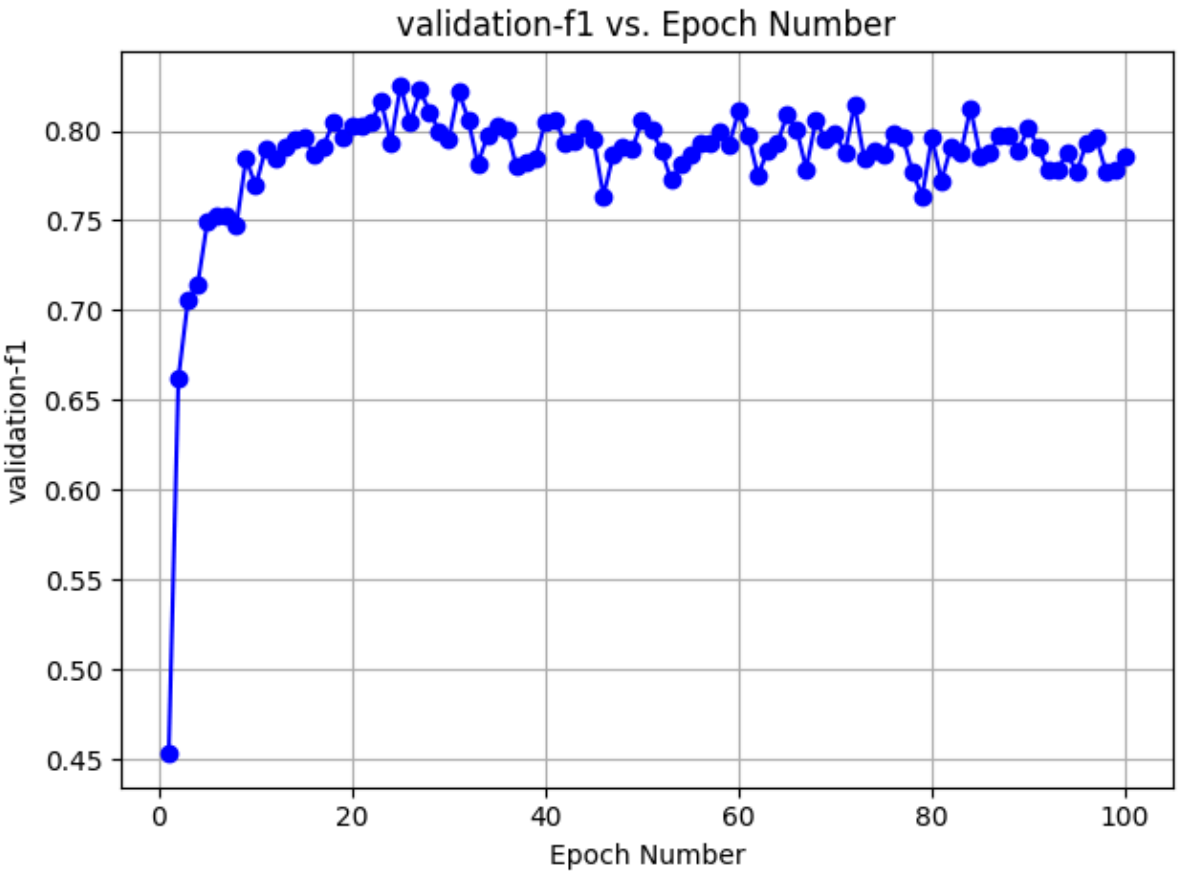

validation loss vs. Epoch Number
validation loss
1000
900
800
700
600
500
0
20
40
60
80
100
Epoch Number

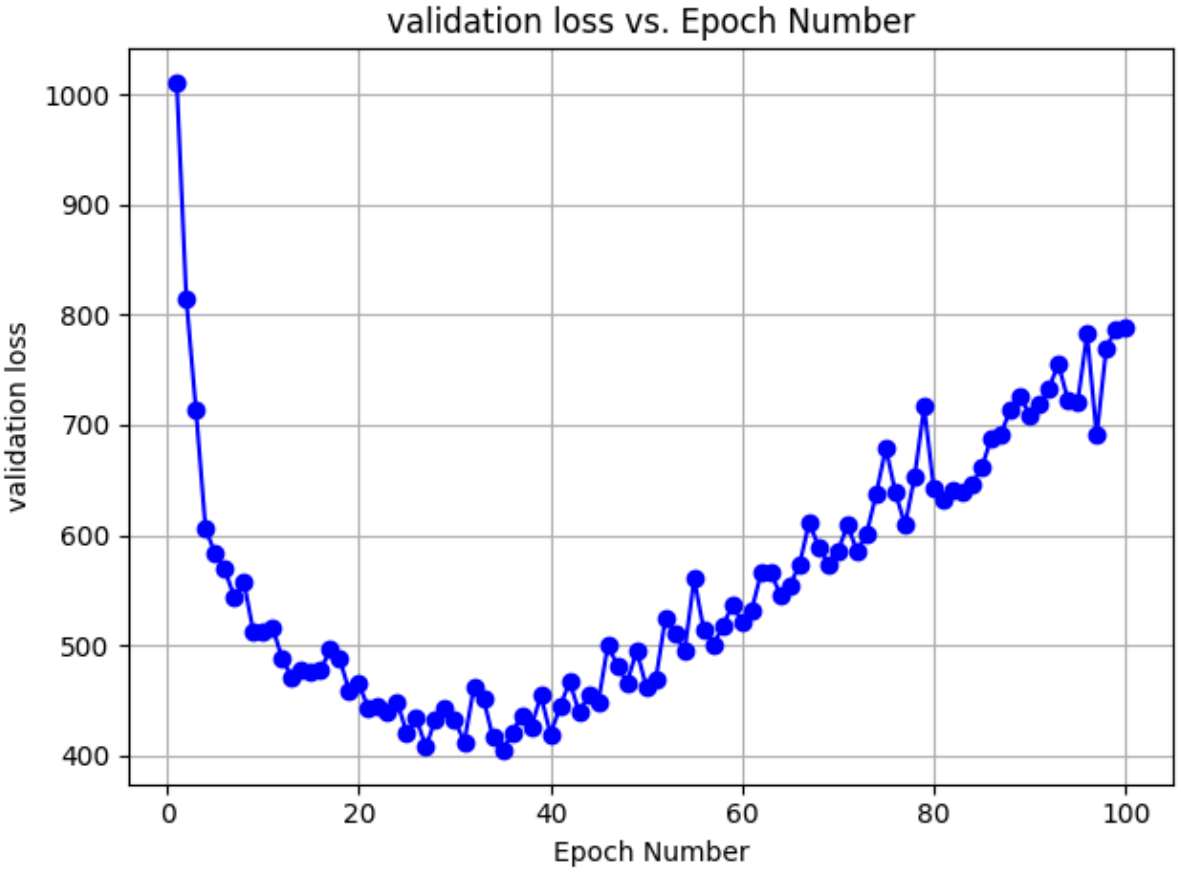
Testing f1=0.789073814564027

bilstm-crf(glove)


traning loss vs. Epoch Number
traning loss
4000
3000
2000
1000
0
0
20
40
60
80
100
Epoch Number

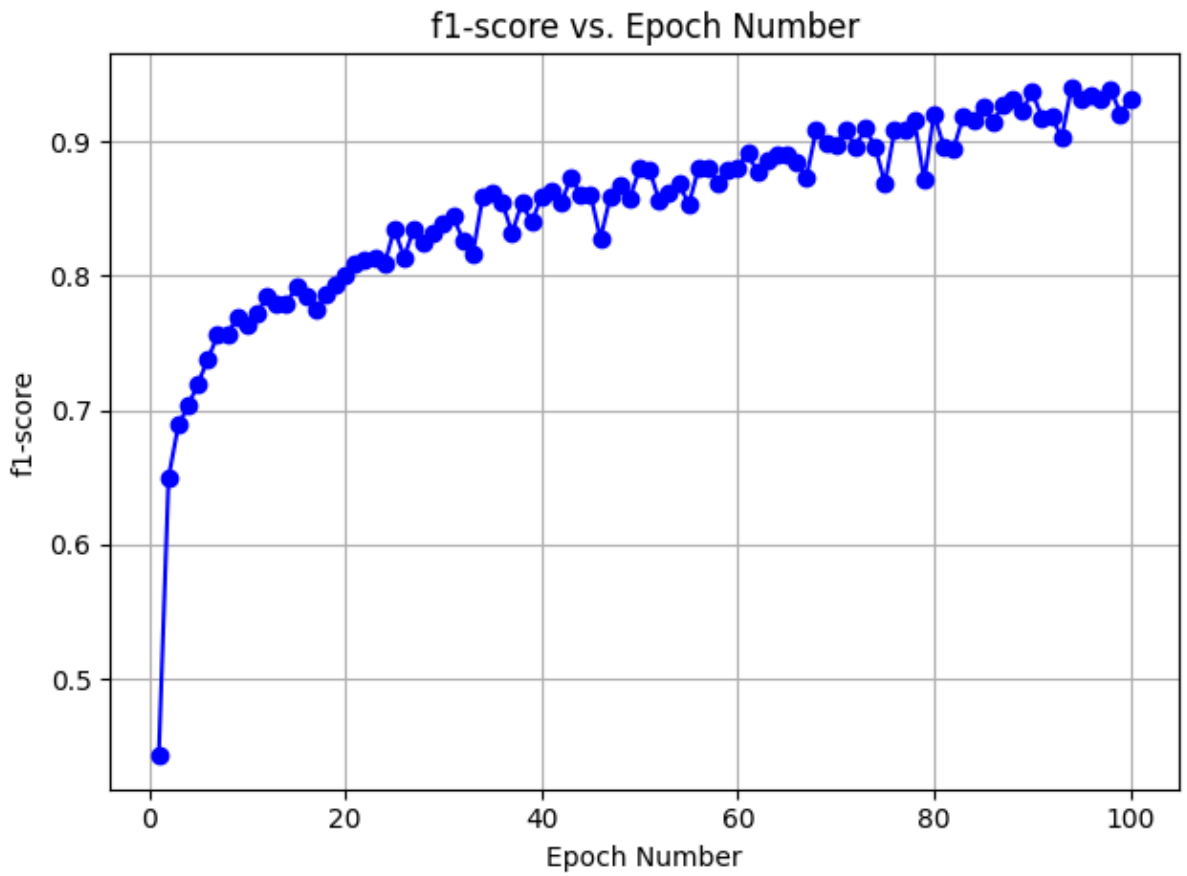

f1-traning vs. Epoch Number
f1-traning
0.9
0.8
0.7
0.6
0.5
0
20
40
60
80
100
Epoch Number

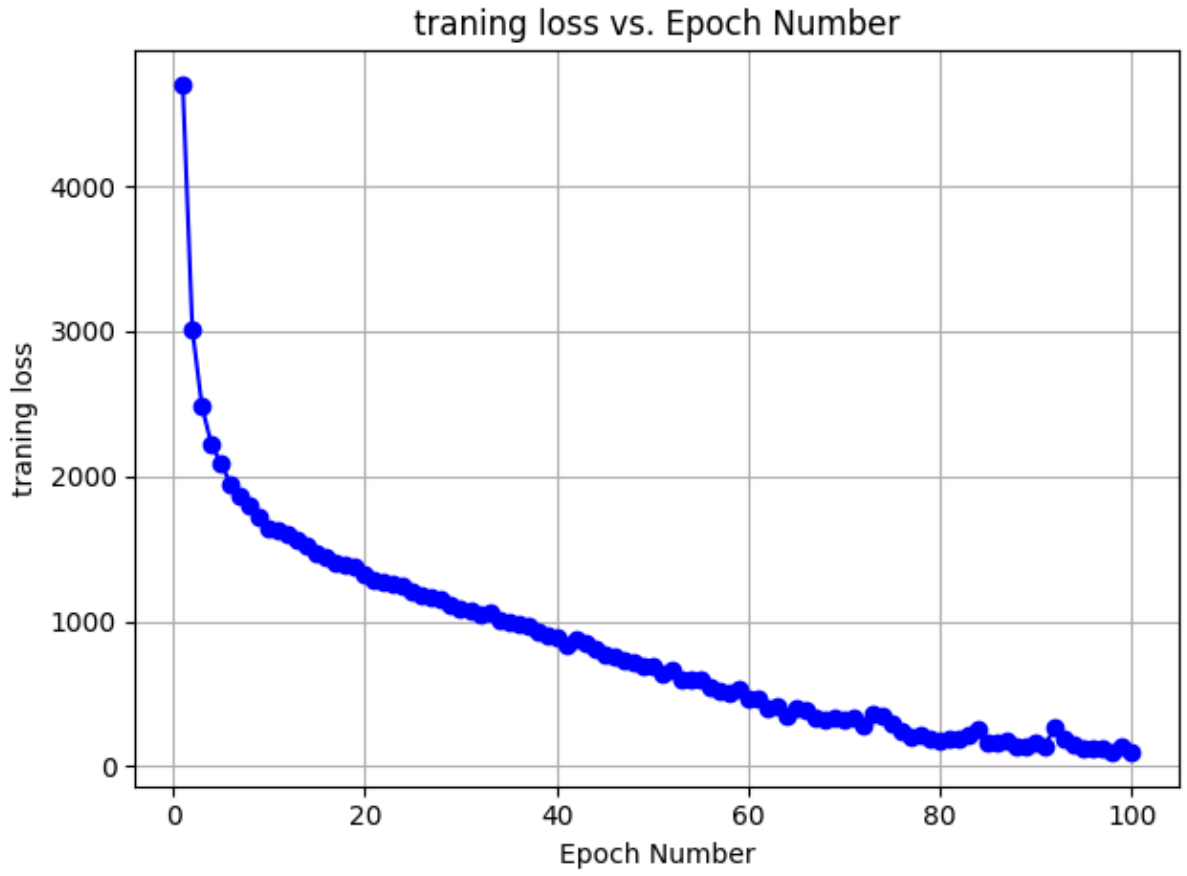

validation-f1 vs. Epoch Number
validation-f1
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0
20
40
60
80
100
Epoch Number

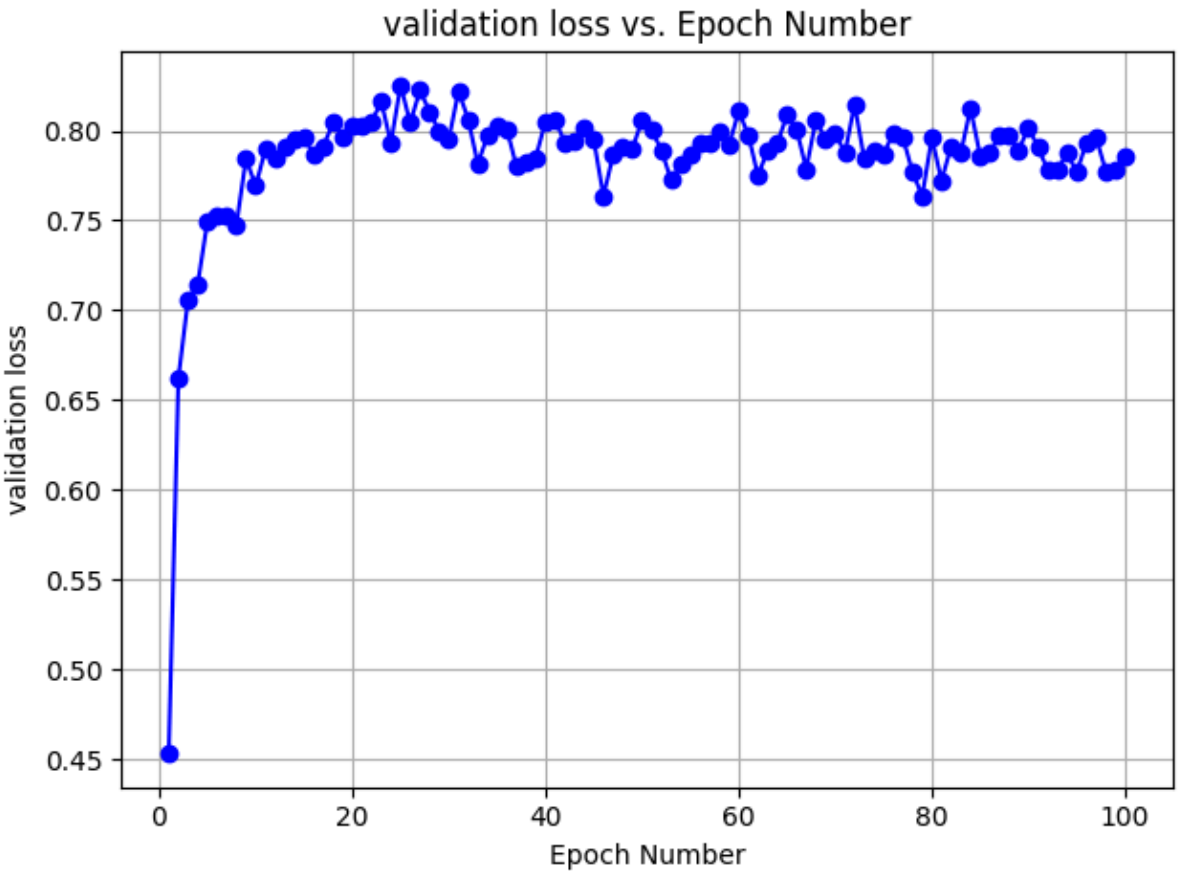

validation loss vs. Epoch Number
validation loss
1000
900
800
700
600
500
400
0
20
40
60
80
100
Epoch Number

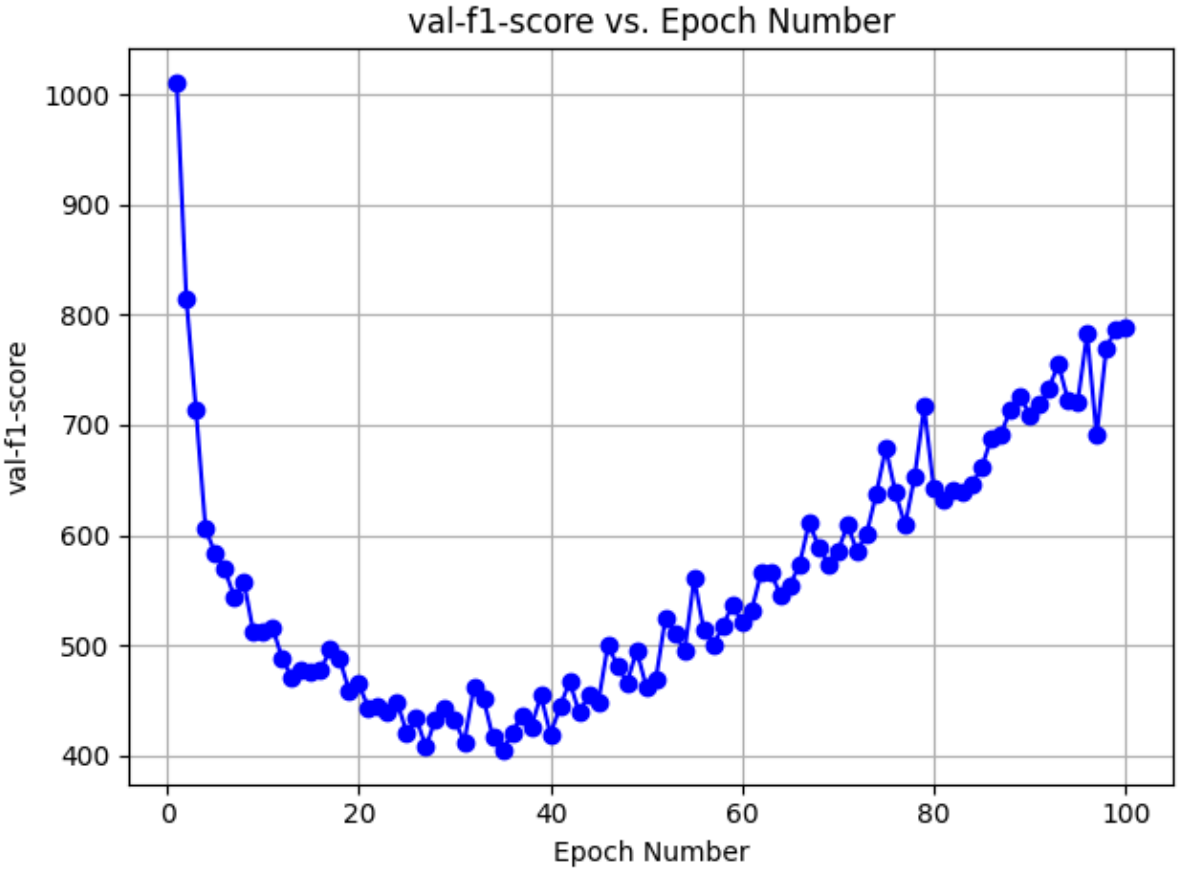
Testing f1=0.8079759051222668

bilstm-crf(fasttext)


f1-score vs. Epoch Number
f1-score
Epoch Number
0.5
0.6
0.7
0.8
0.9
0
20
40
60
80
100


traning loss vs. Epoch Number
traning loss
Epoch Number
0
1000
2000
3000
4000
0
20
40
60
80
100


validation loss vs. Epoch Number
validation loss
Epoch Number
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0
20
40
60
80
100


val-f1-score vs. Epoch Number
val-f1-score
Epoch Number
1000
900
800
700
600
500
400
0
20
40
60
80
100

Testing score=0.8032702843359497

F1 Score Table for Task-2:

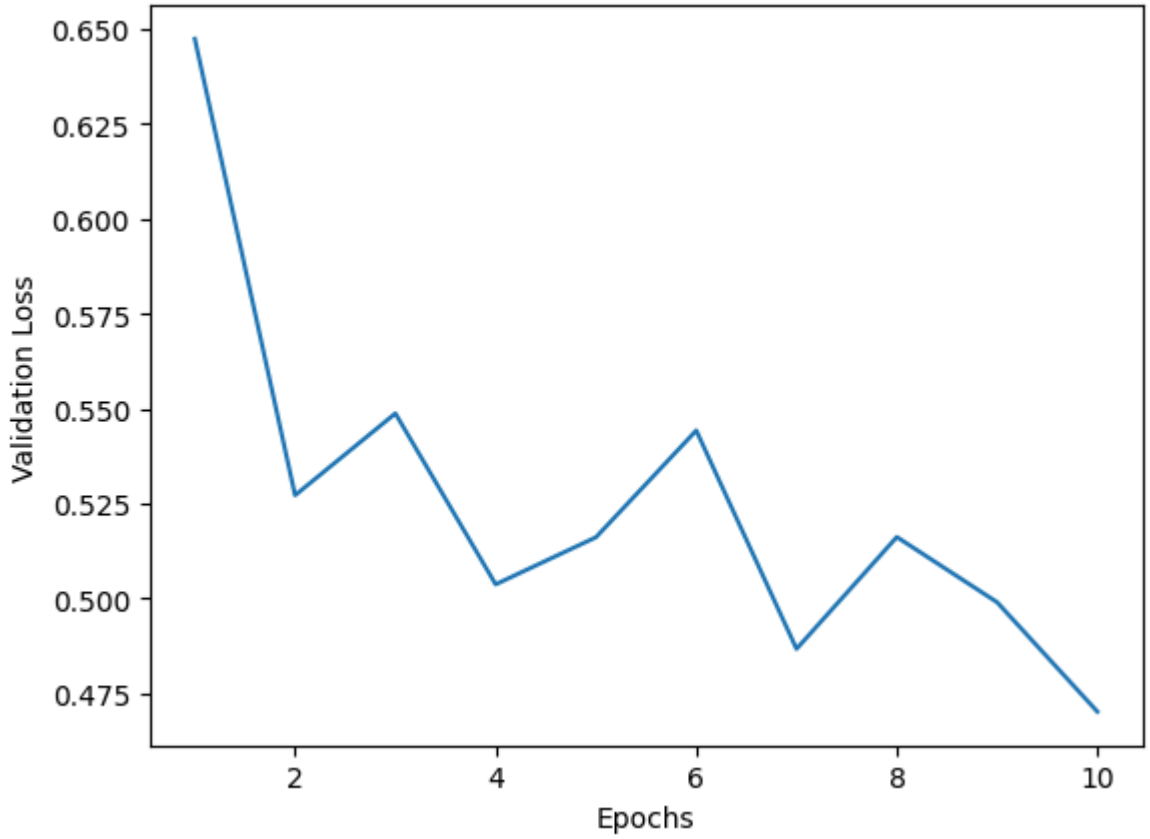
Model Used	Word Embedding Used		
	Word2Vec	GloVe	FastText
Vanilla RNN	0.738680210590	0.615371975757	0.775791260404
LSTM	0.776924574441	0.709430534213	0.755002579094
GRU	0.761749808372	0.706967476331	0.787246601667
BiLSTM-CRF	0.789073814564	0.807975905122	0.803270284335

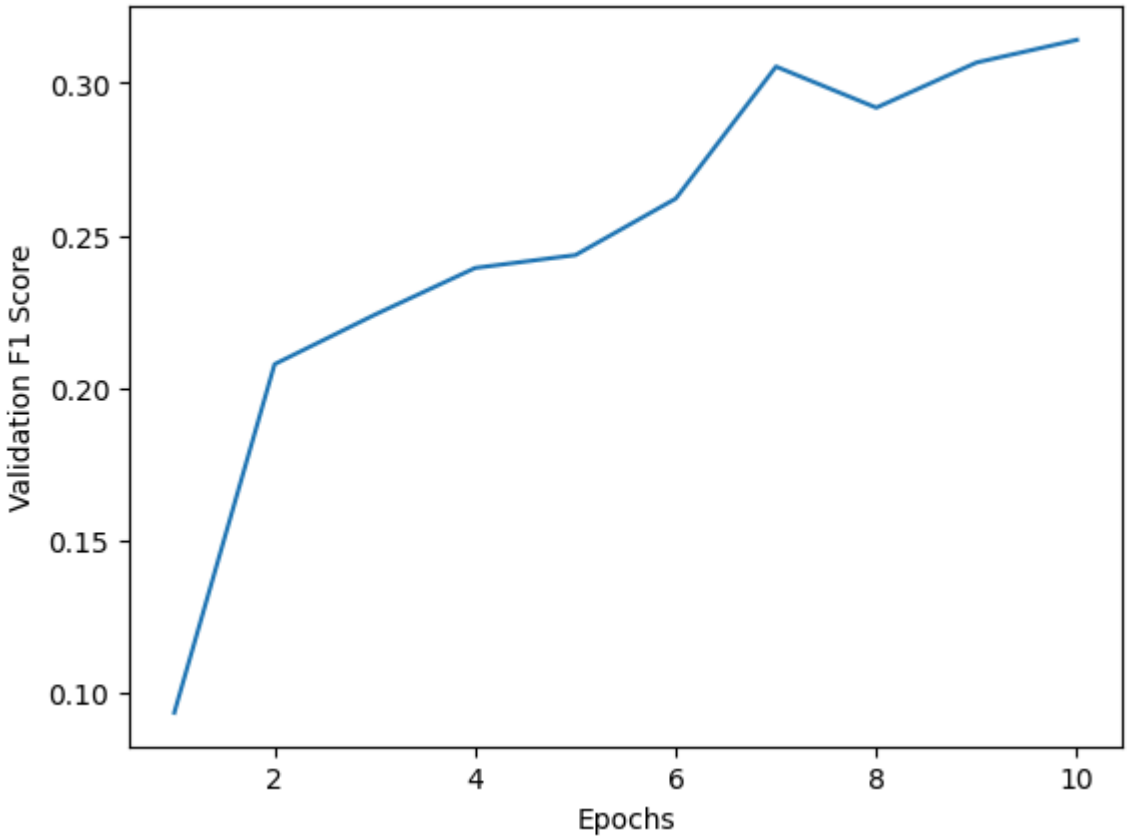
F1 Score Table for Task-1:

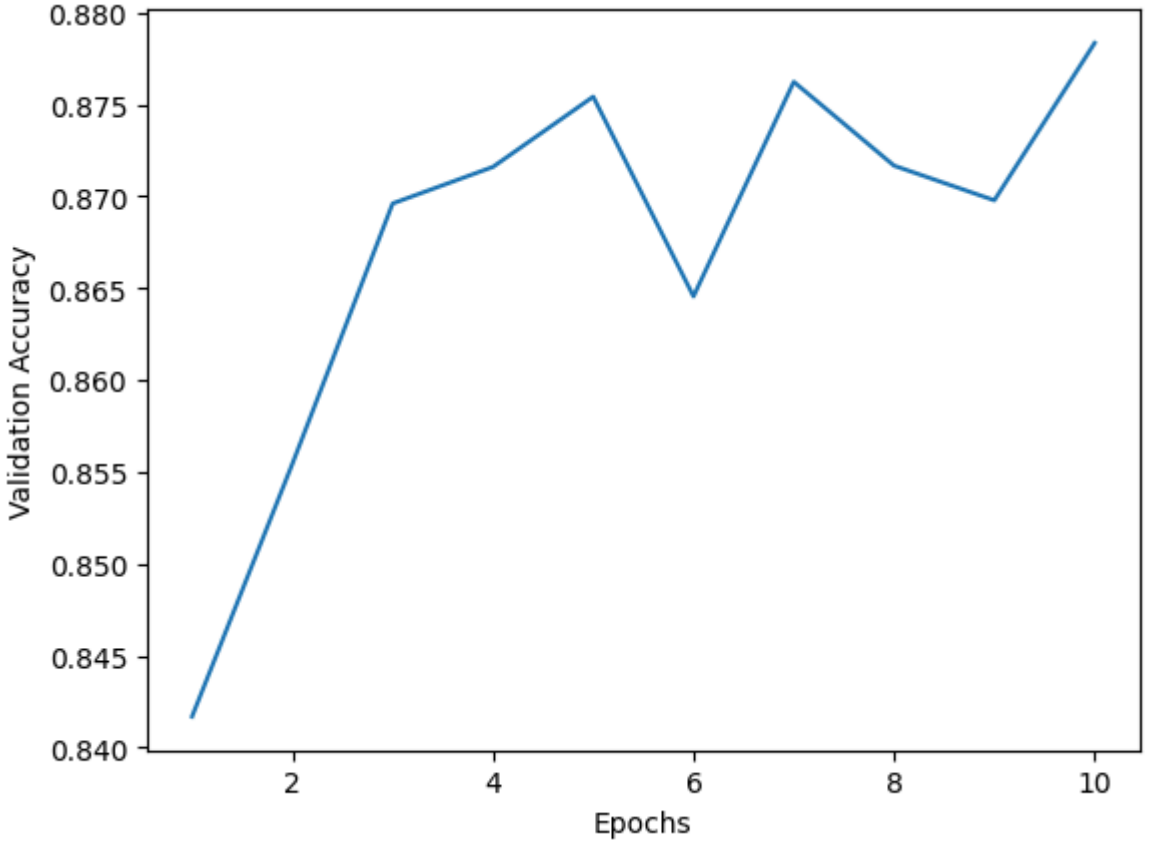
Model Used	Word Embedding Used		
	Word2Vec	GloVe	FastText
Vanilla RNN	0.335854560592052	0.381039036151791	0.232372399642
LSTM	0.589553484270634	0.629000508136211	0.545726889720
GRU	0.587956081680481	0.620198650768853	0.561156907260
BiLSTM-CRF	-	-	-

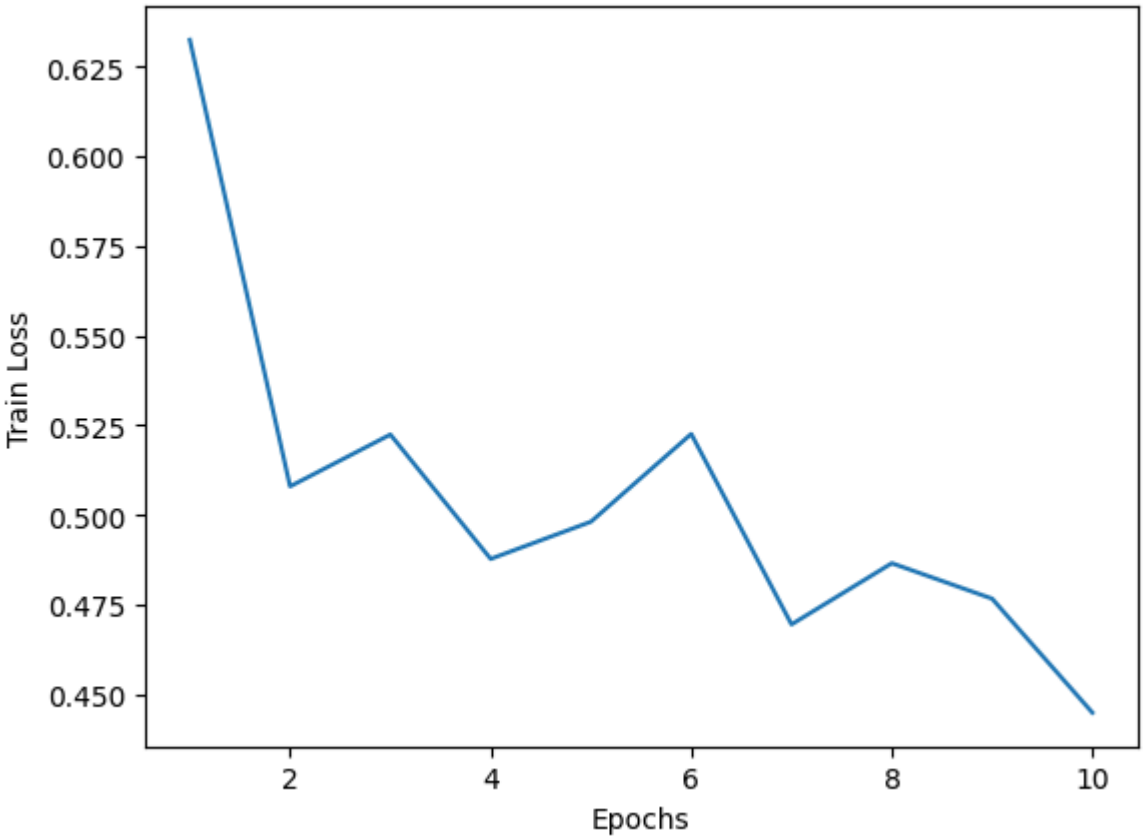
TASK-1 GRAPHS:

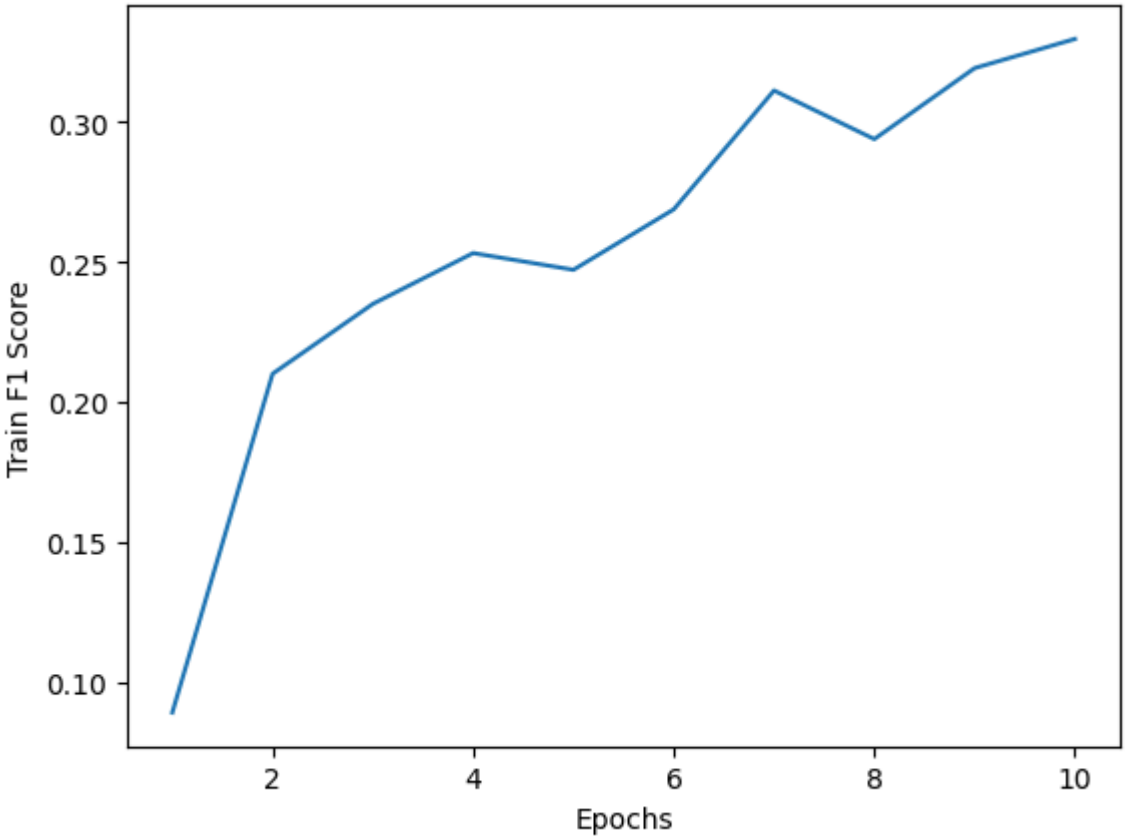
1) RNN With Word2Vec:

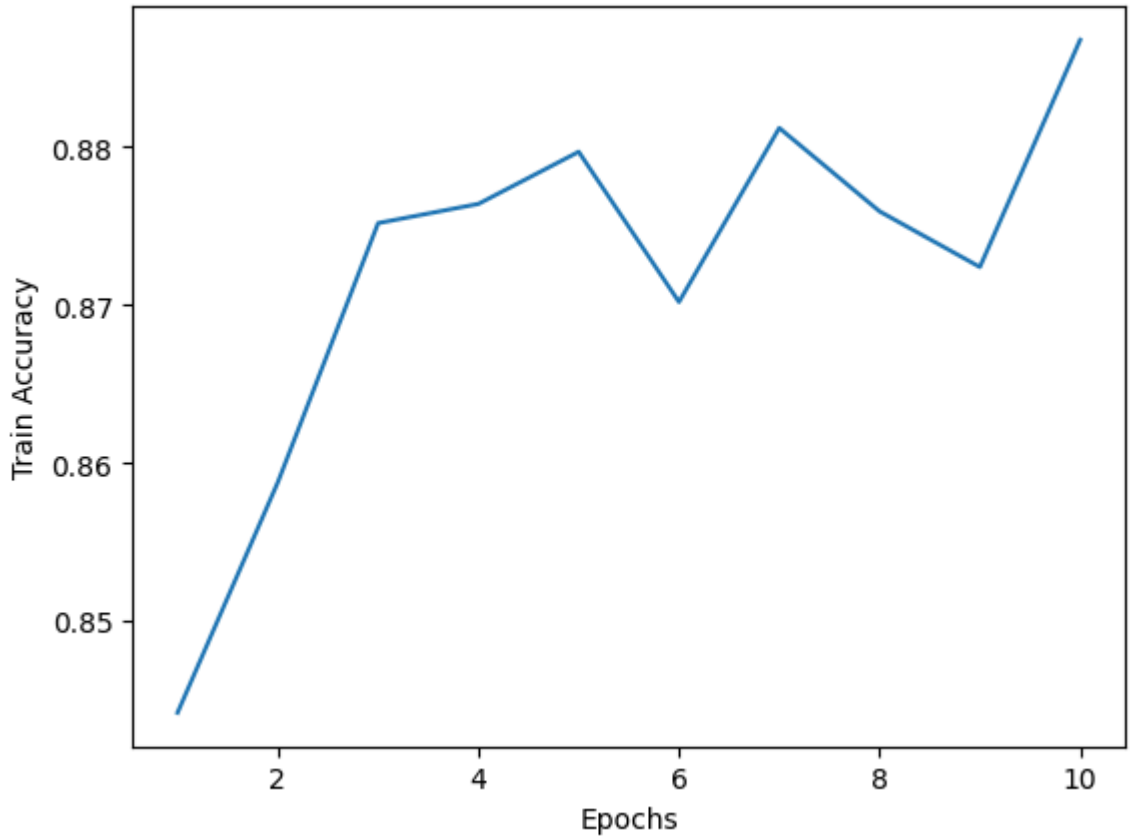



Validation F1 Score
0.30
0.25
0.20
0.15
0.10
2
4
6
8
10
Epochs

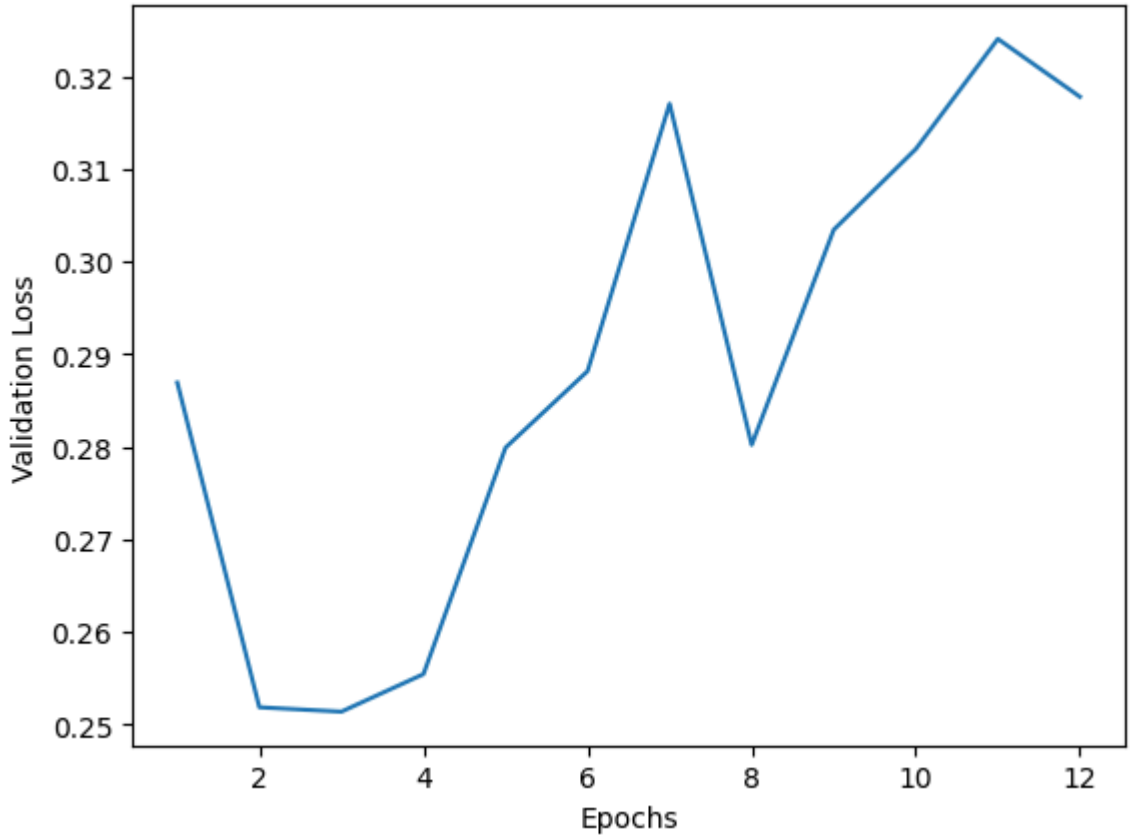

Validation Accuracy
0.880
0.875
0.870
0.865
0.860
0.855
0.850
0.845
0.840
2
4
6
8
10
Epochs

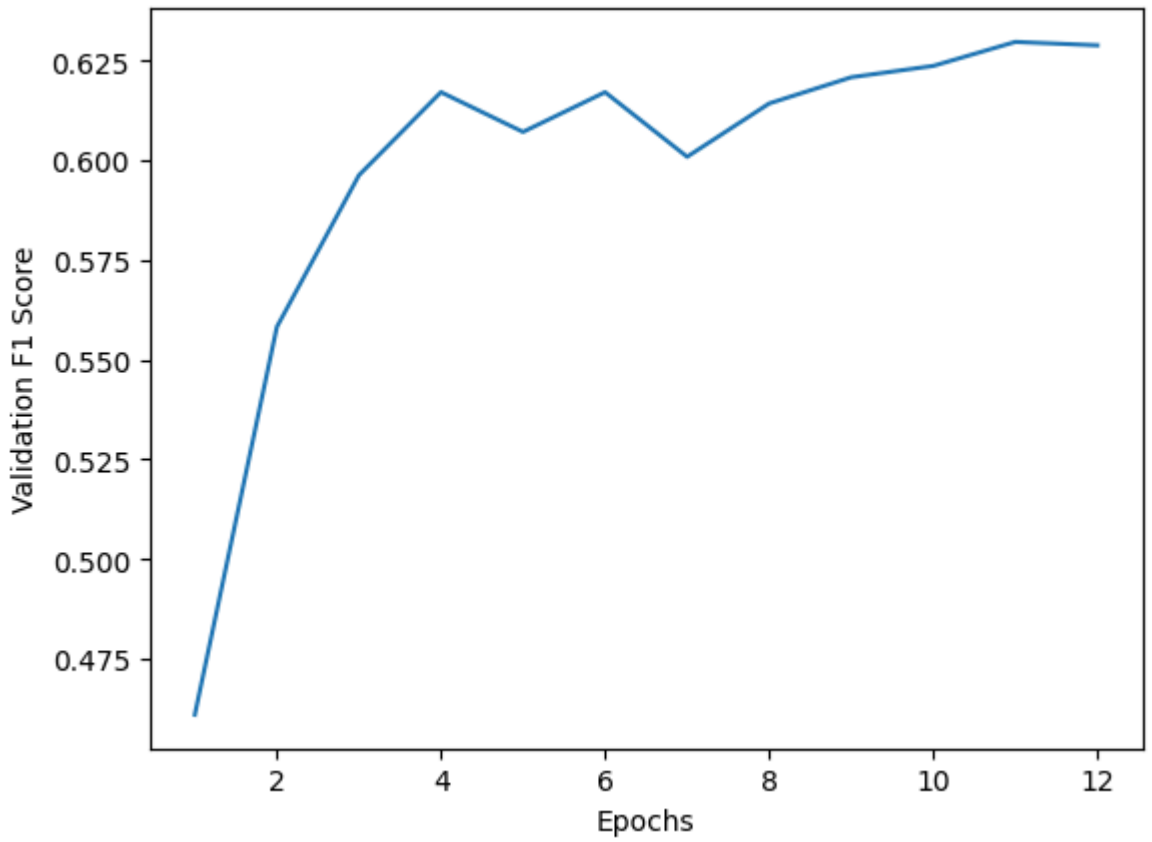

0.625
0.600
0.575
0.550
0.525
0.500
0.475
0.450
Train Loss
2
4
6
8
10
Epochs

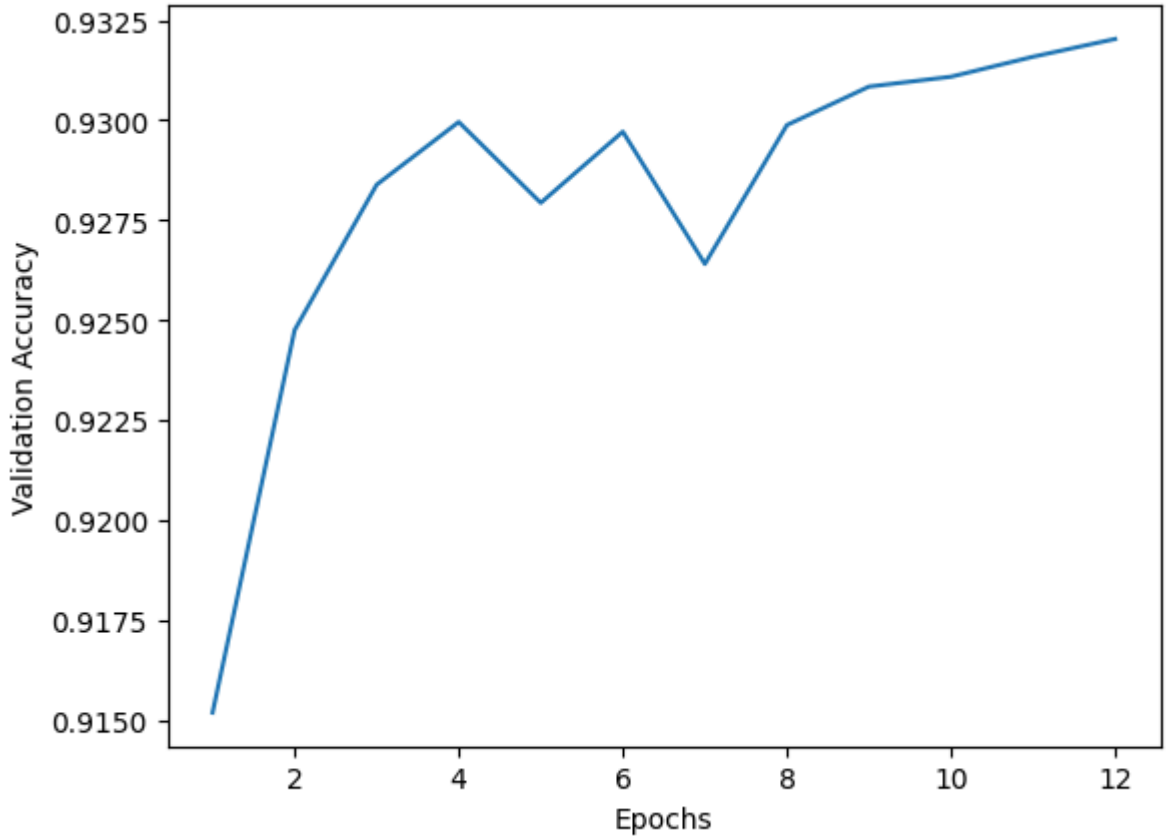

0.30
0.25
0.20
0.15
0.10
Train F1 Score
2
4
6
8
10
Epochs

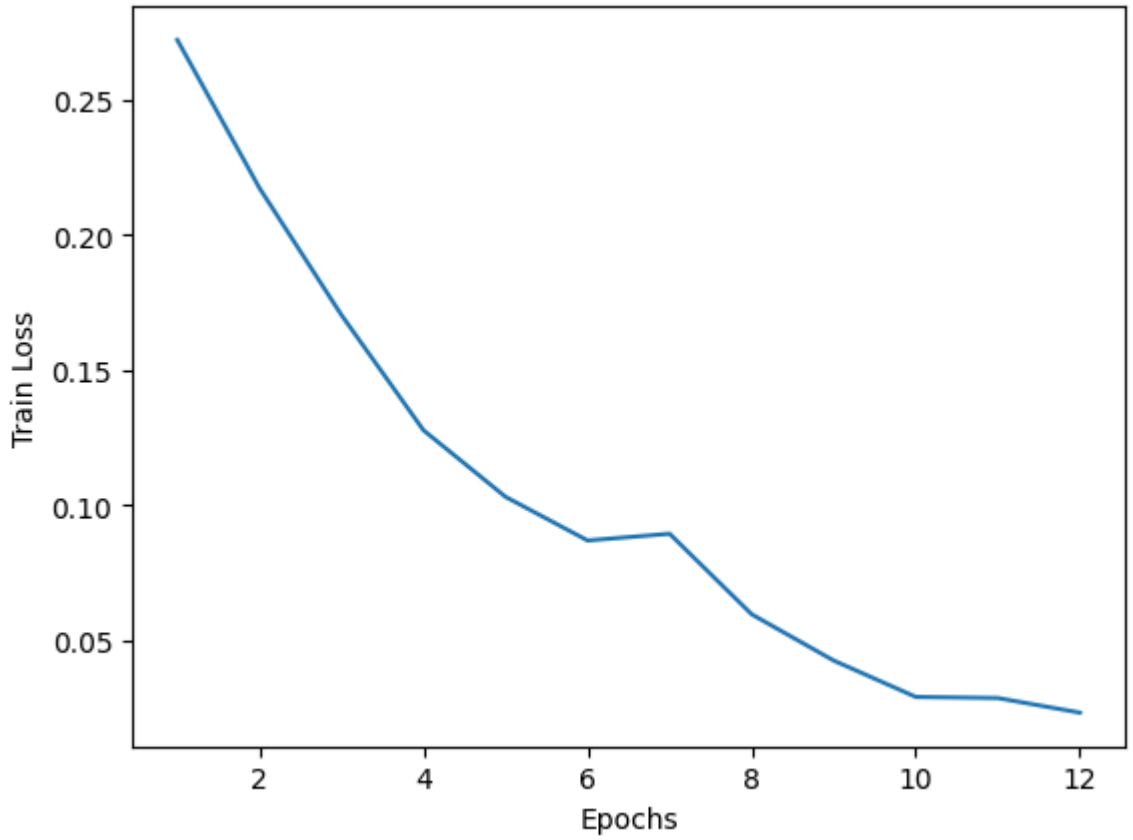


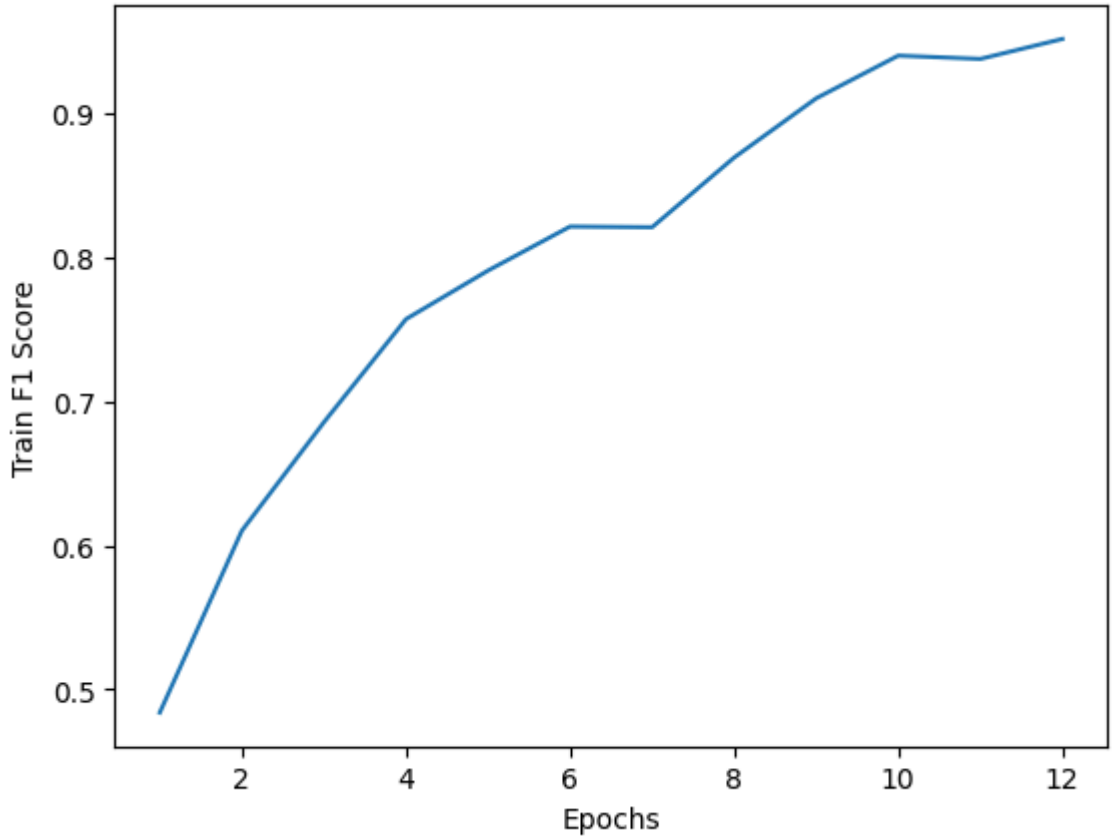
2) LSTM with Word2vec:

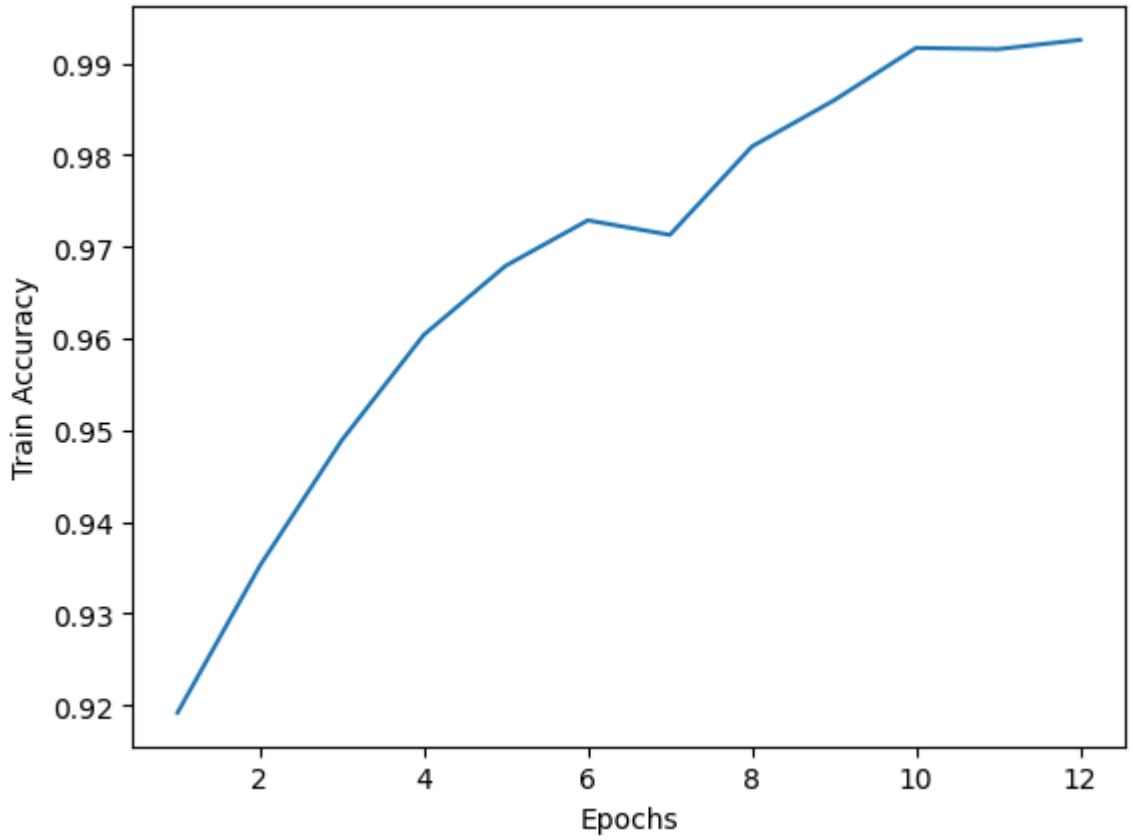



Validation F1 Score
0.625
0.600
0.575
0.550
0.525
0.500
0.475
2
4
6
8
10
12
Epochs

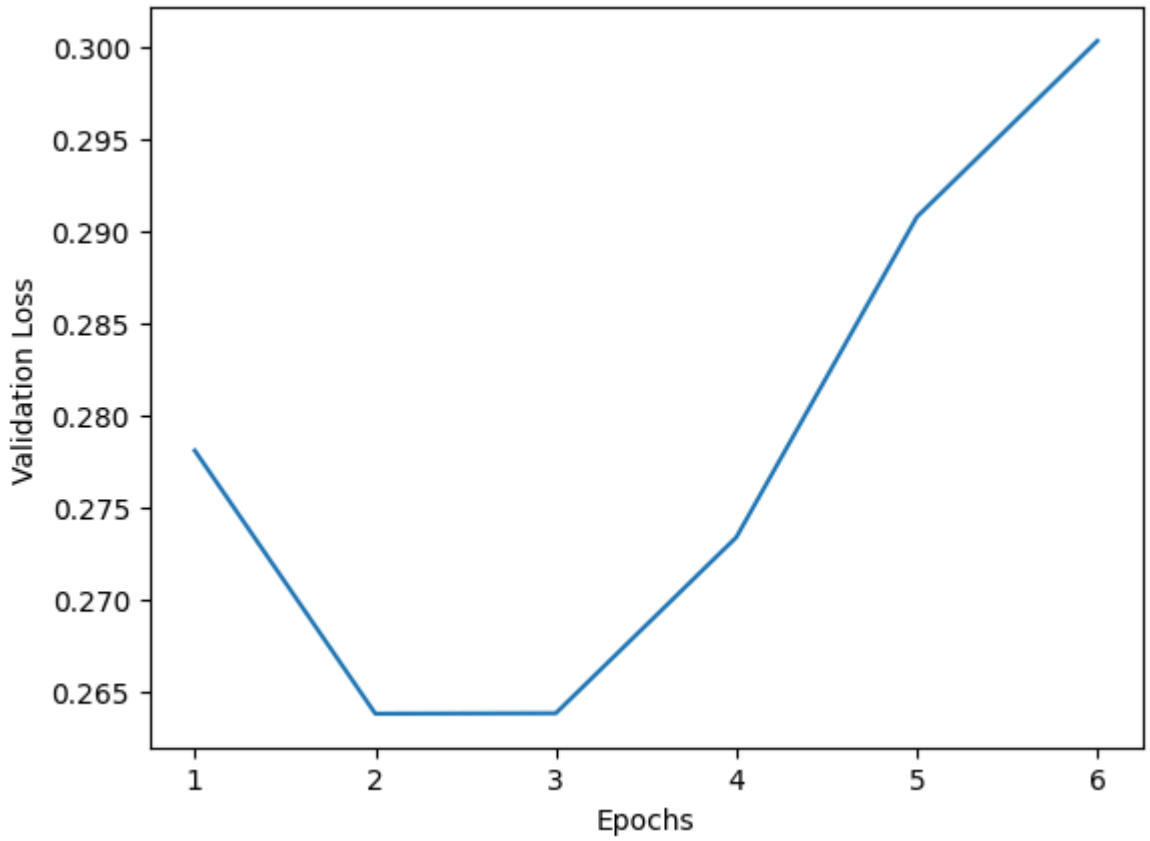

Validation Accuracy
0.9325
0.9300
0.9275
0.9250
0.9225
0.9200
0.9175
0.9150
2
4
6
8
10
12
Epochs

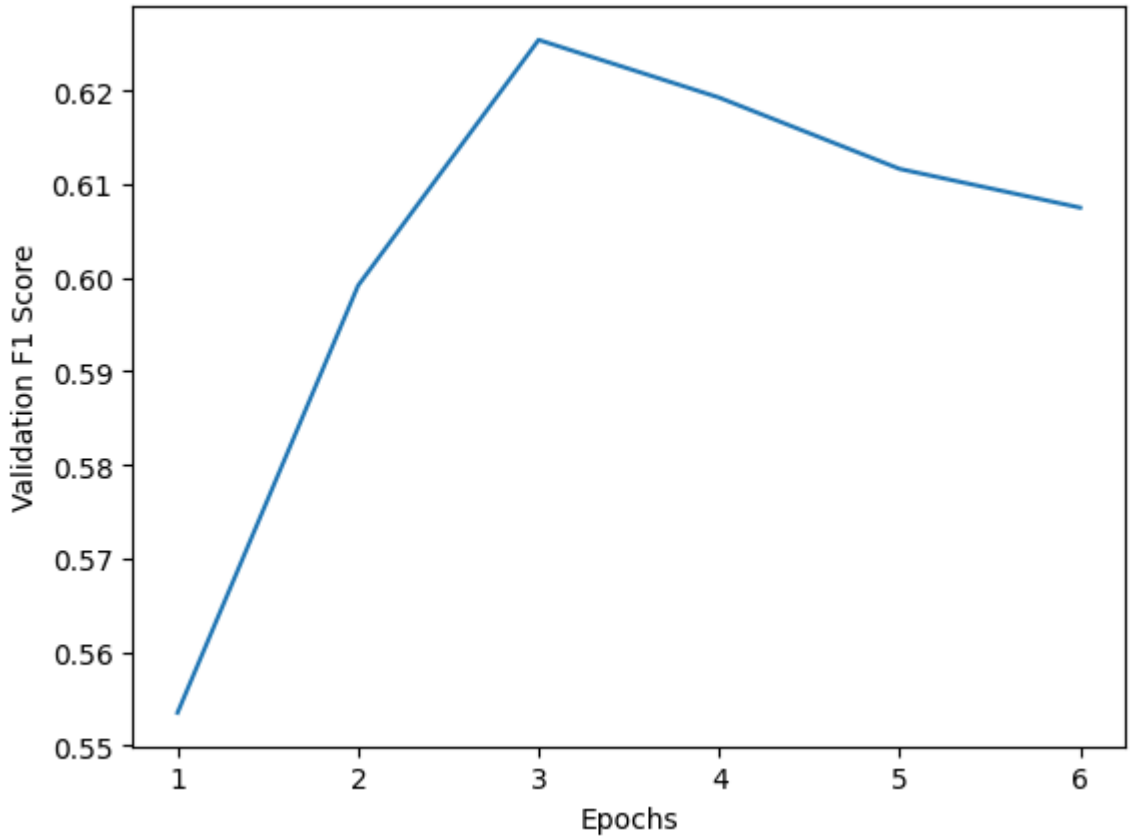

Train Loss
0.25
0.20
0.15
0.10
0.05
2
4
6
8
10
12
Epochs

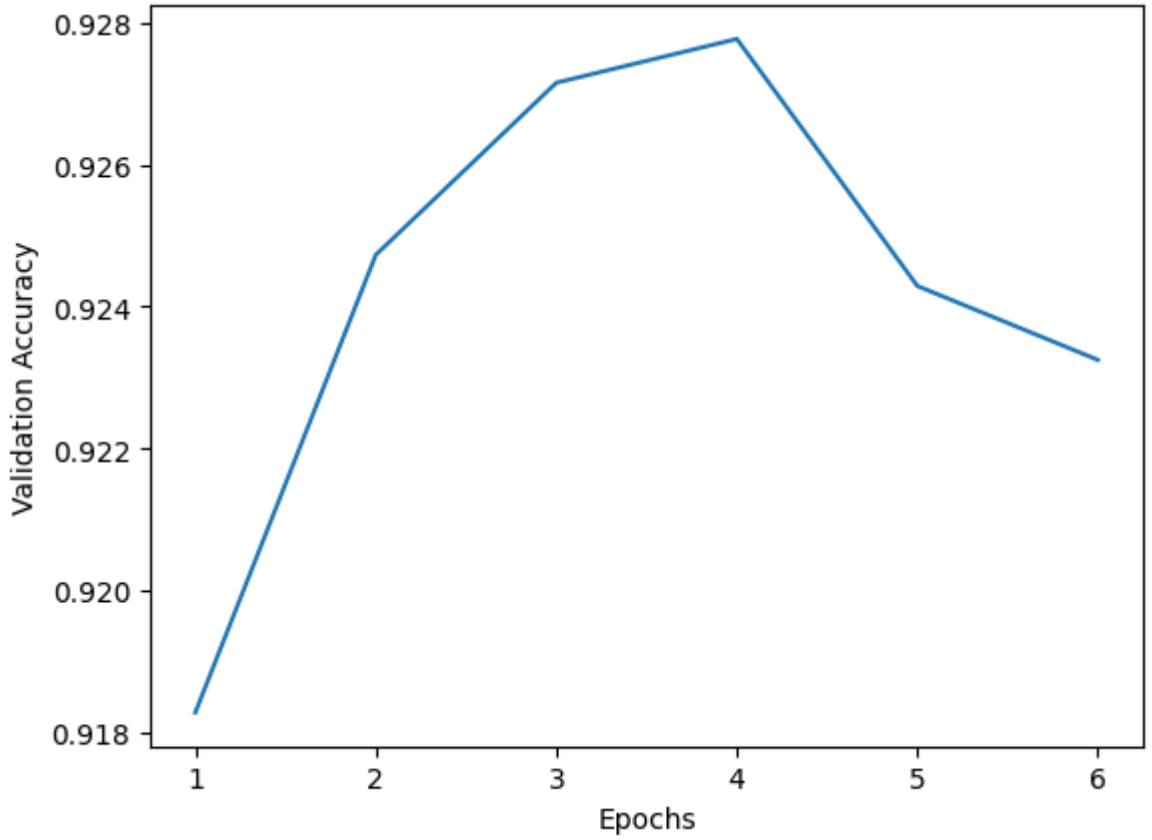

Train F1 Score
0.9
0.8
0.7
0.6
0.5
2
4
6
8
10
12
Epochs

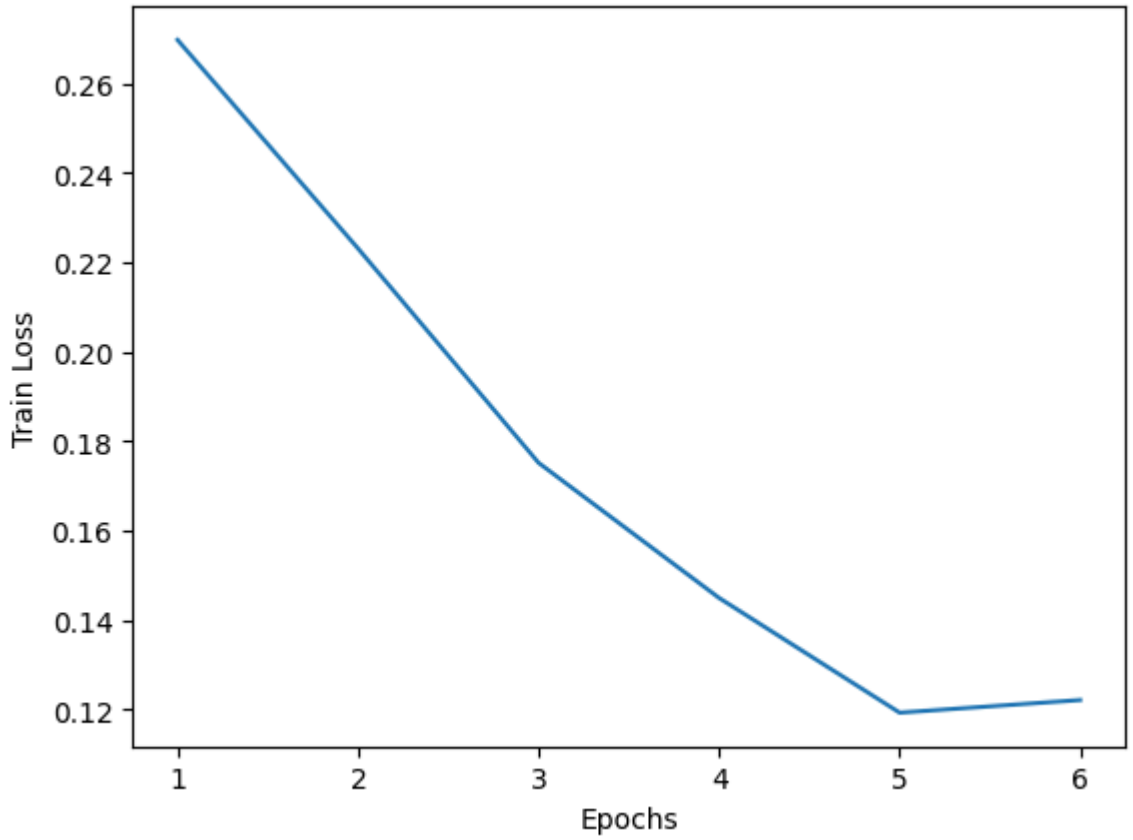


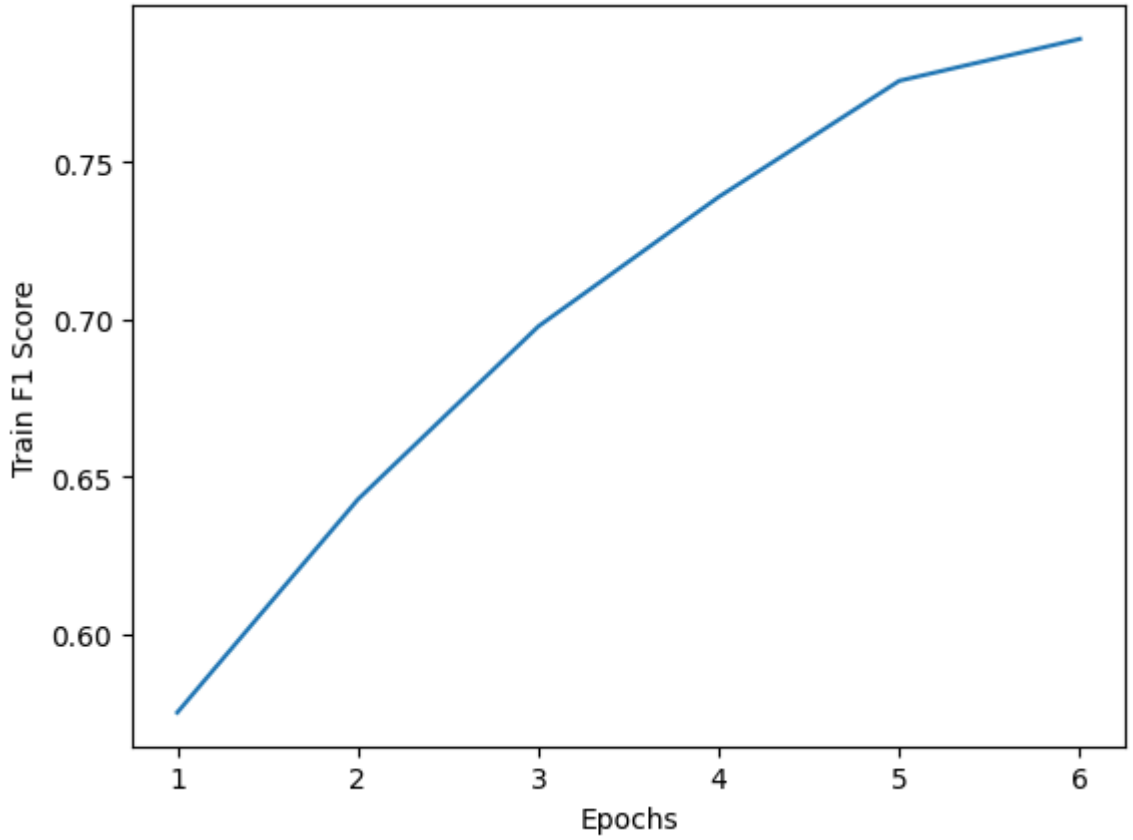
3) GRU With Word2Vec:

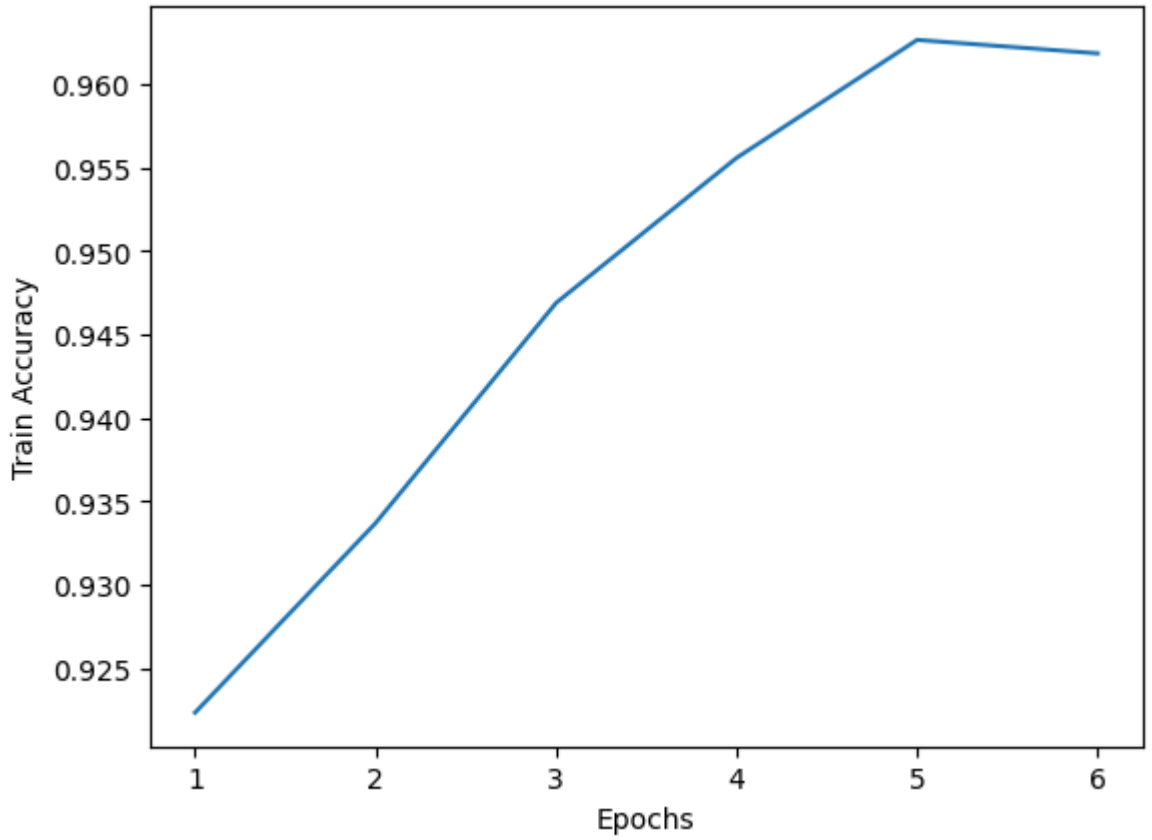



Validation F1 Score
0.62
0.61
0.60
0.59
0.58
0.57
0.56
0.55
1
2
3
4
5
6
Epochs

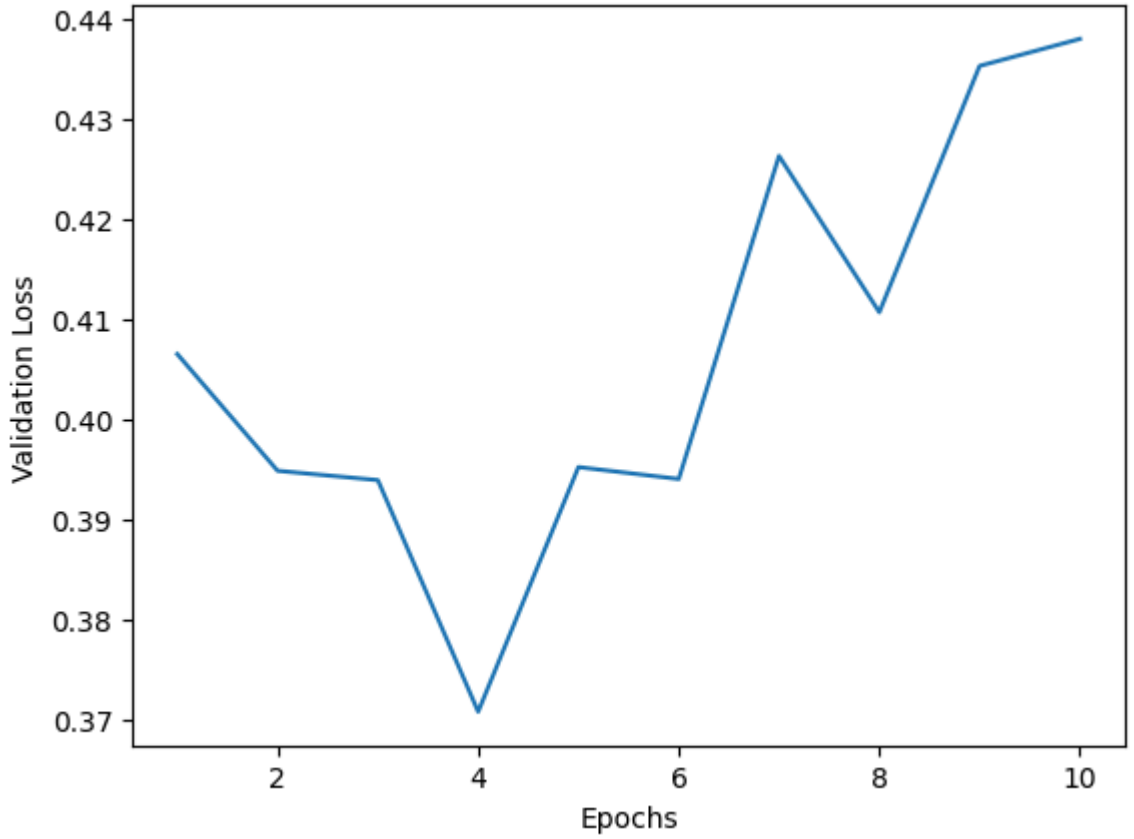

Validation Accuracy
0.928
0.926
0.924
0.922
0.920
0.918
1
2
3
4
5
6
Epochs

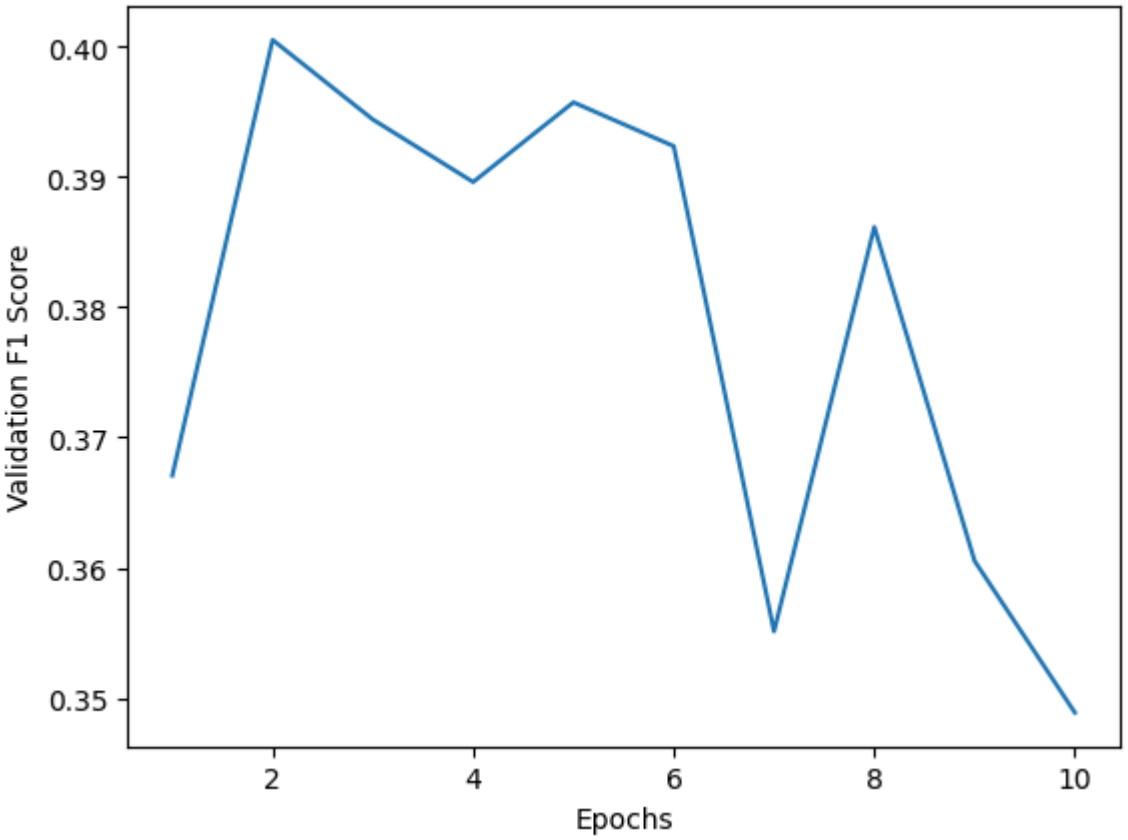

Train Loss
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
1
2
3
4
5
6
Epochs

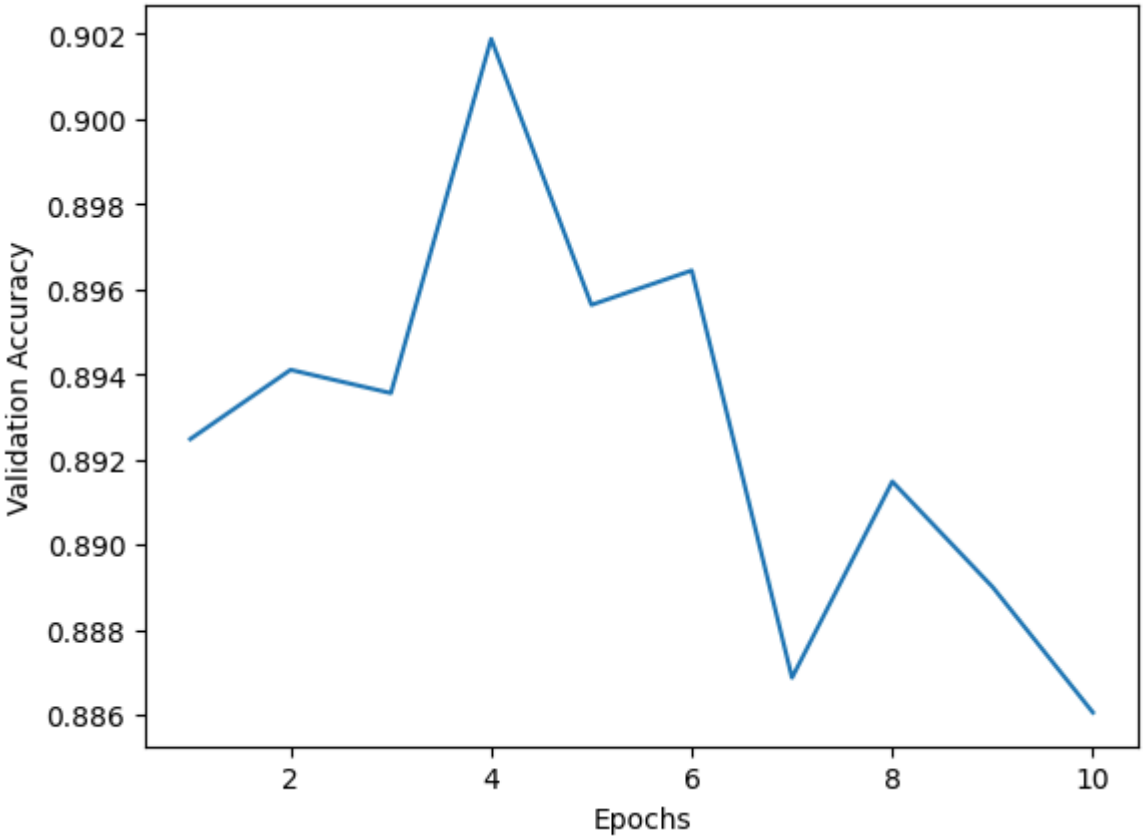

Train F1 Score
0.75
0.70
0.65
0.60
1
2
3
4
5
6
Epochs

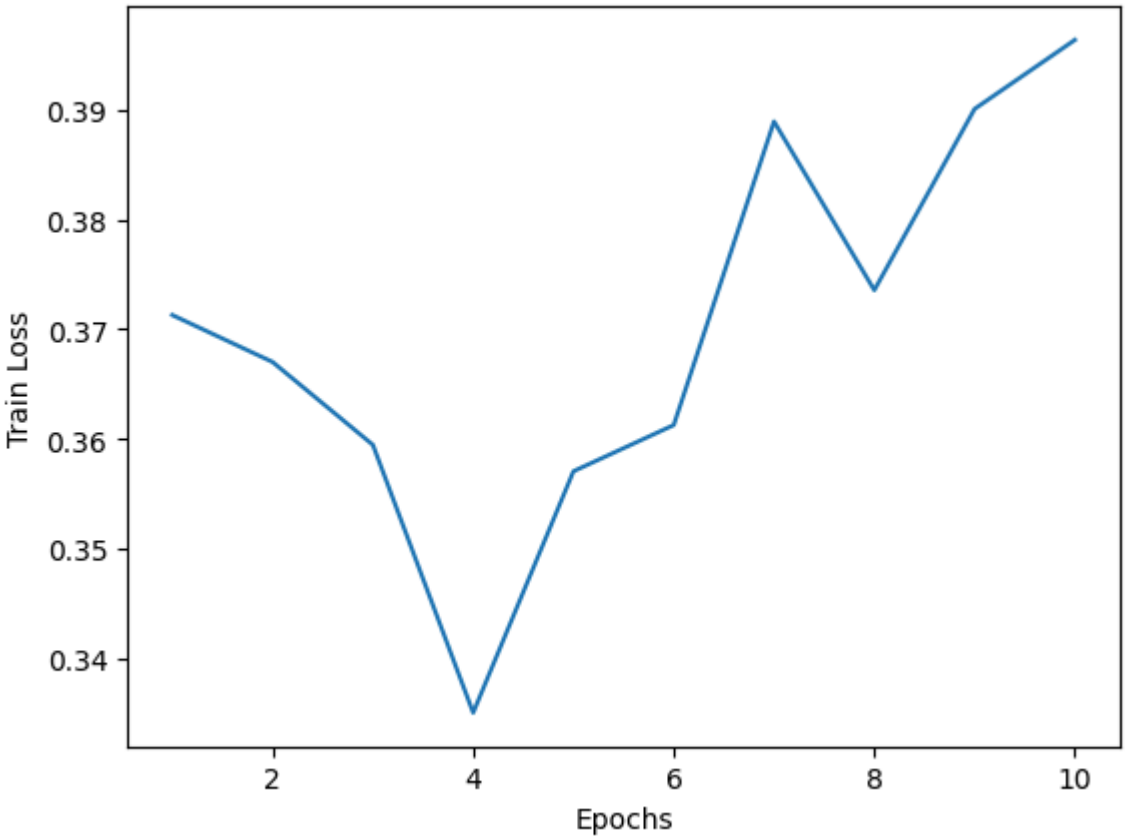


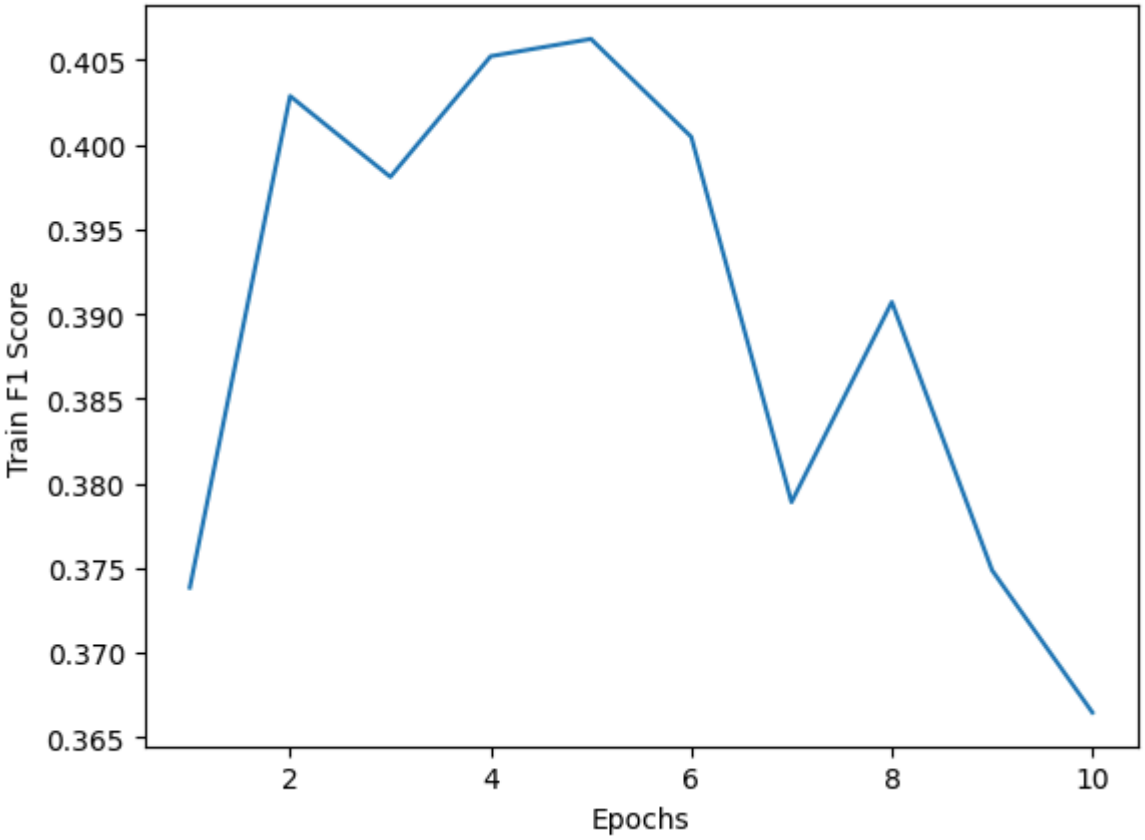
4) RNN with GloVe:

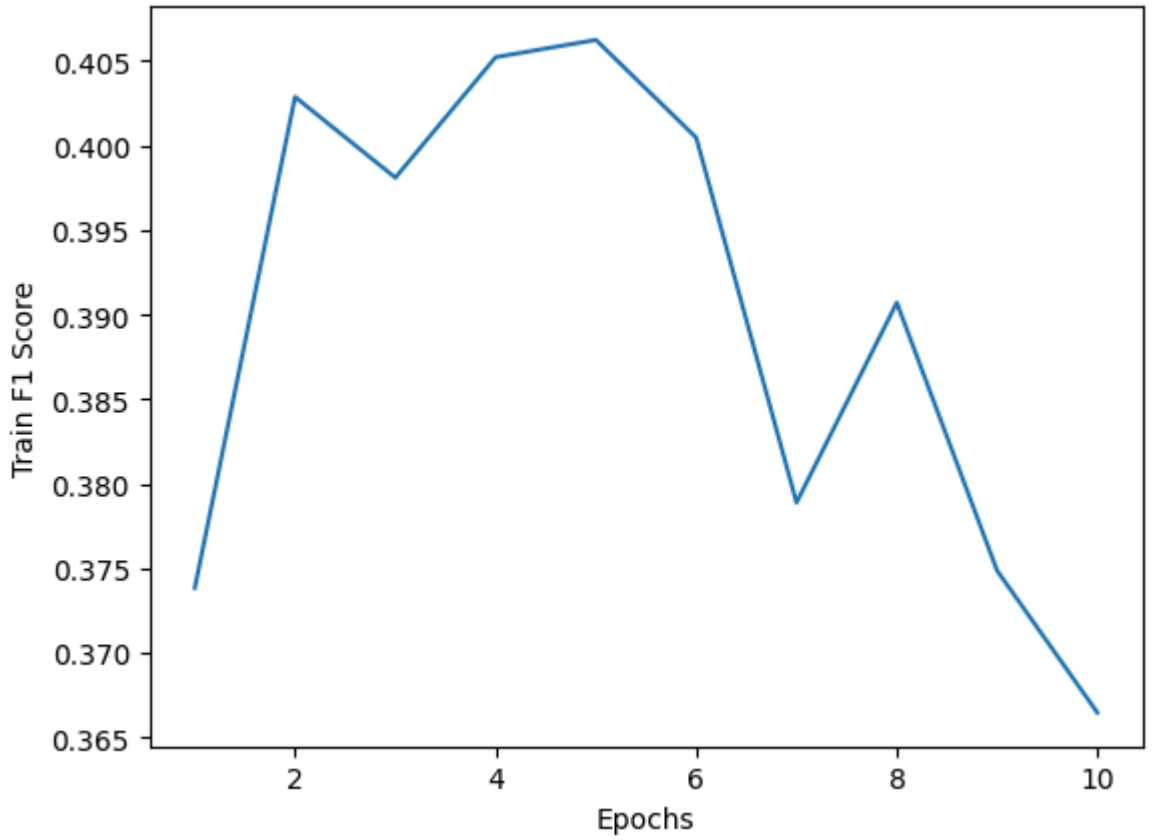



0.40
0.39
0.38
0.37
0.36
0.35
Validation F1 Score
2
4
6
8
10
Epochs

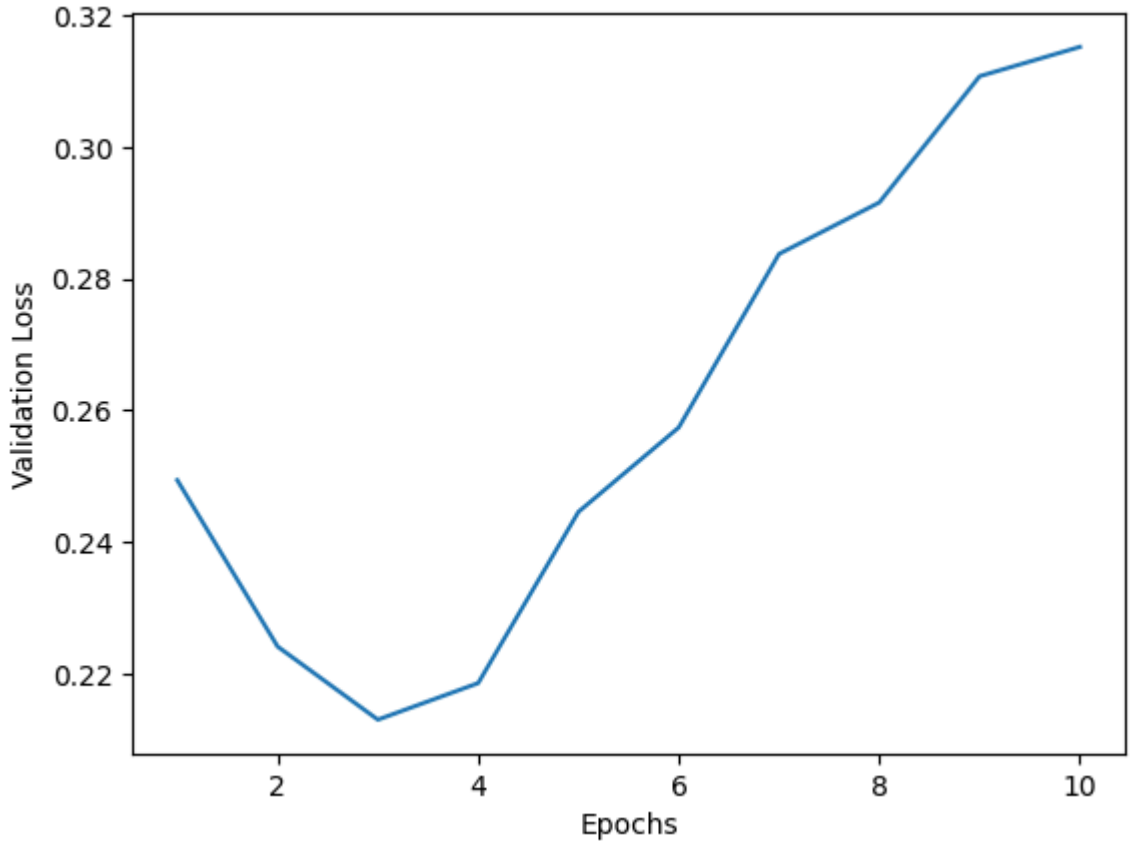

0.902
0.900
0.898
0.896
0.894
0.892
0.890
0.888
0.886
Validation Accuracy
2
4
6
8
10
Epochs

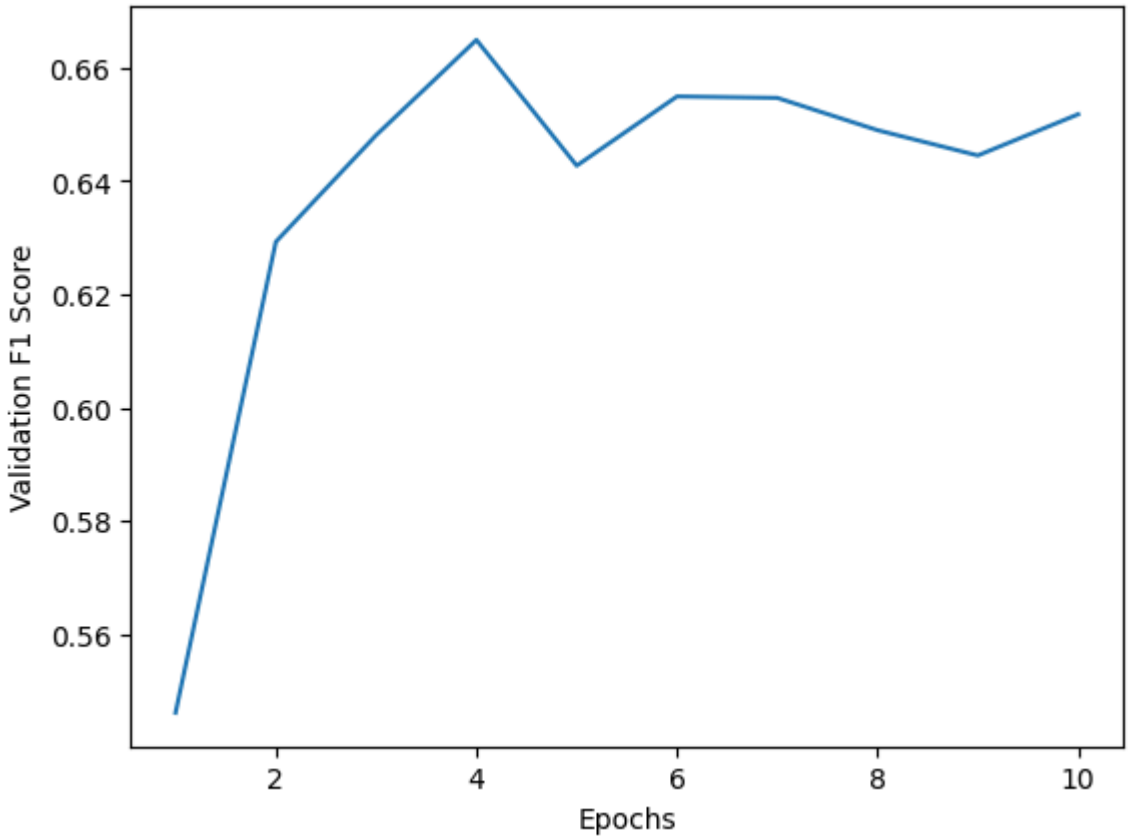

Train Loss
0.39
0.38
0.37
0.36
0.35
0.34
2
4
6
8
10
Epochs

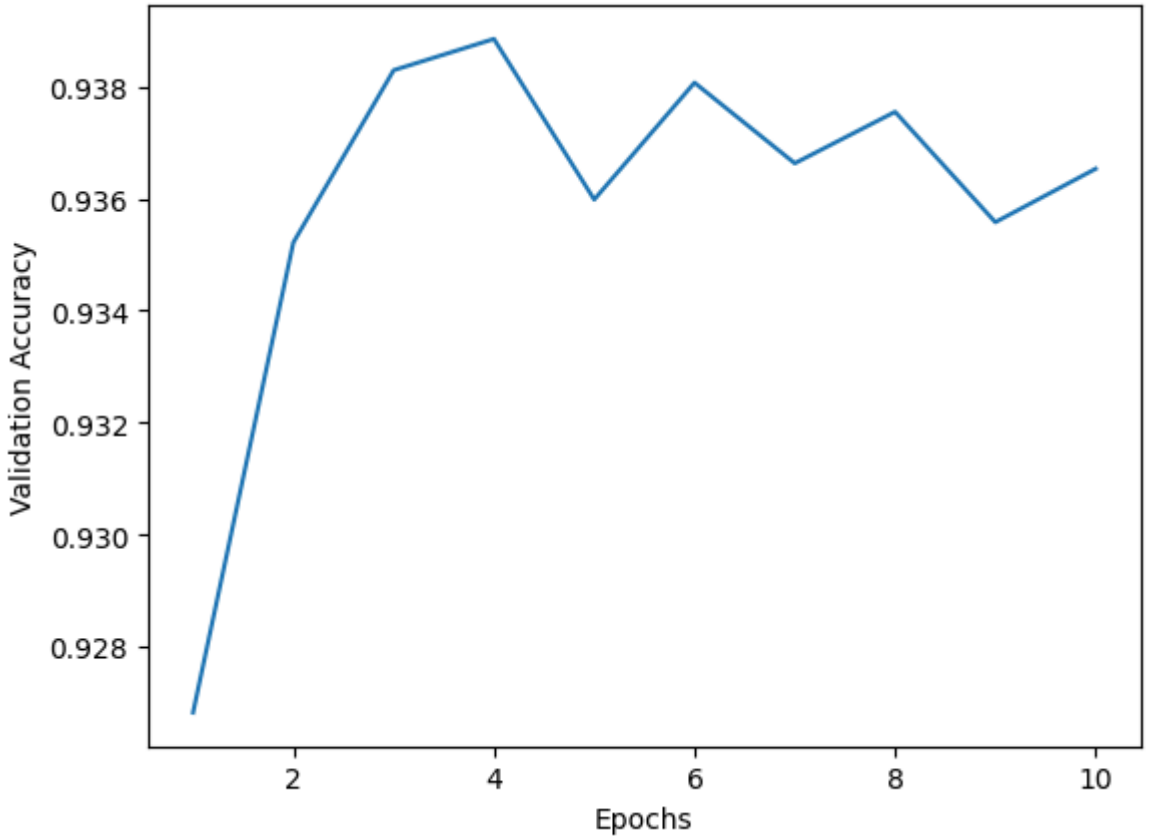

Train F1 Score
0.405
0.400
0.395
0.390
0.385
0.380
0.375
0.370
0.365
2
4
6
8
10
Epochs

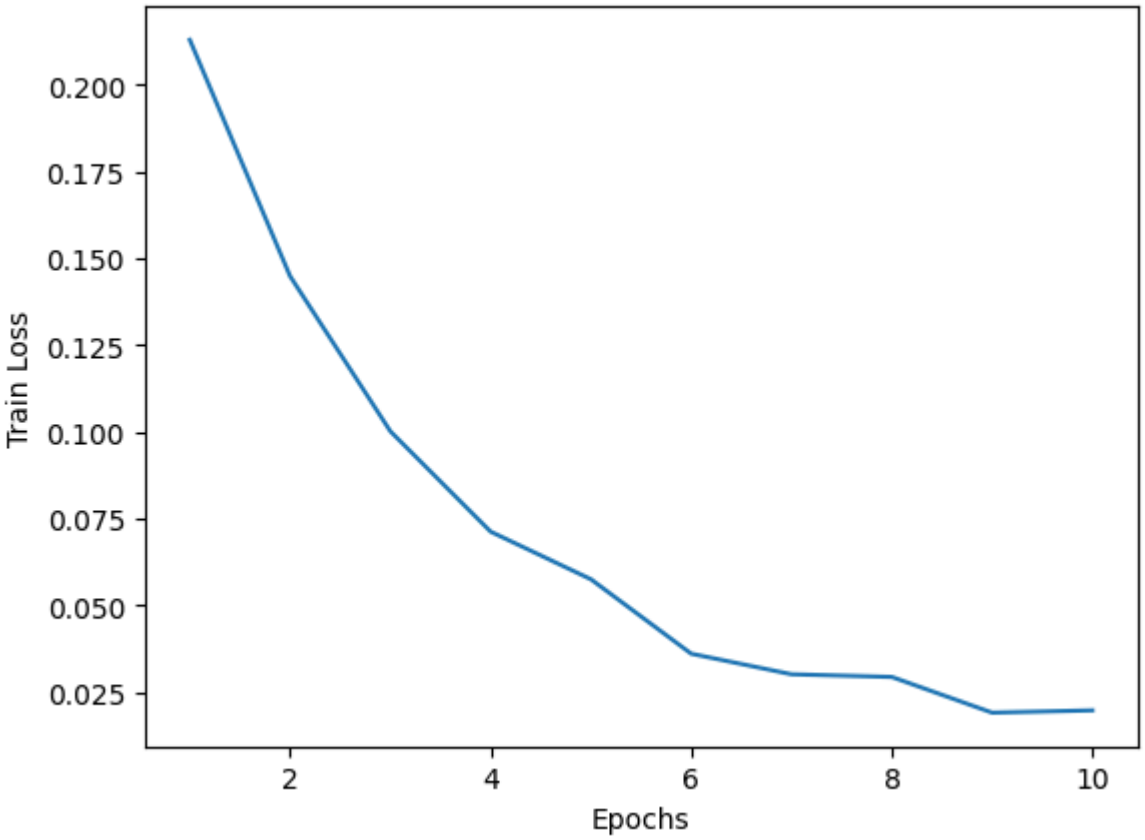


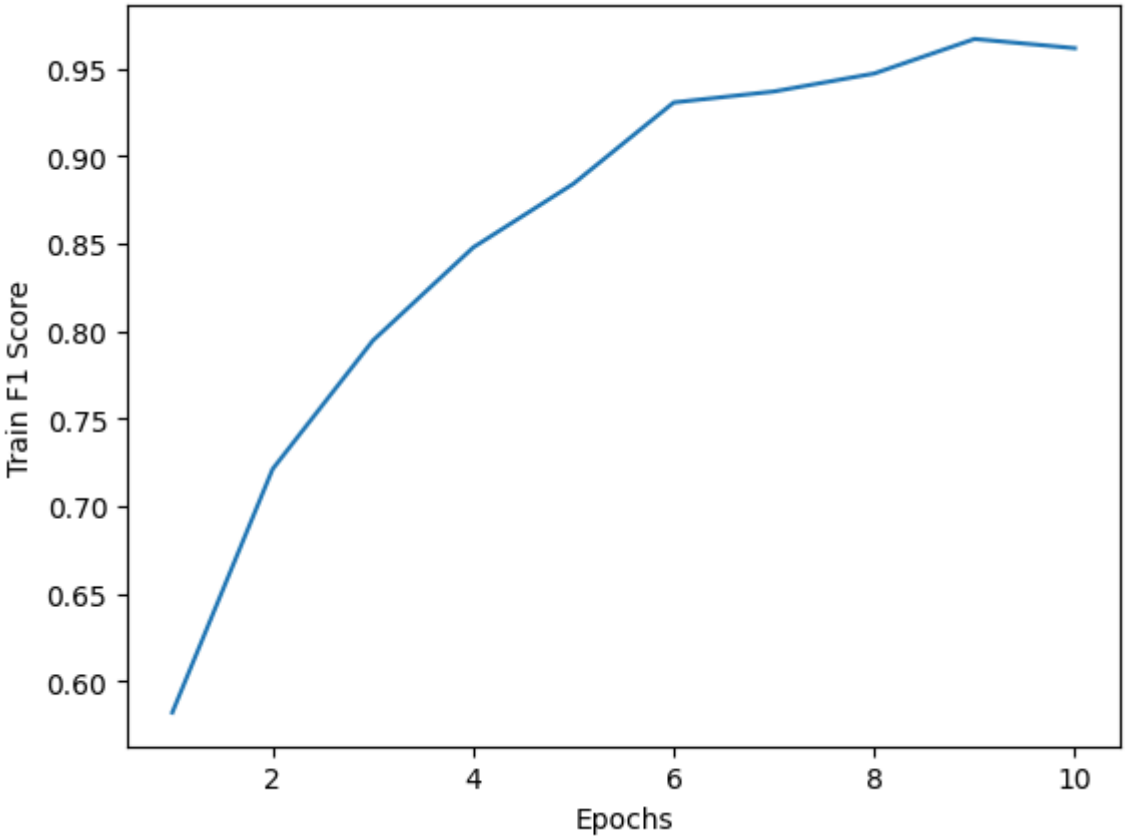
5) LSTM with GloVe:

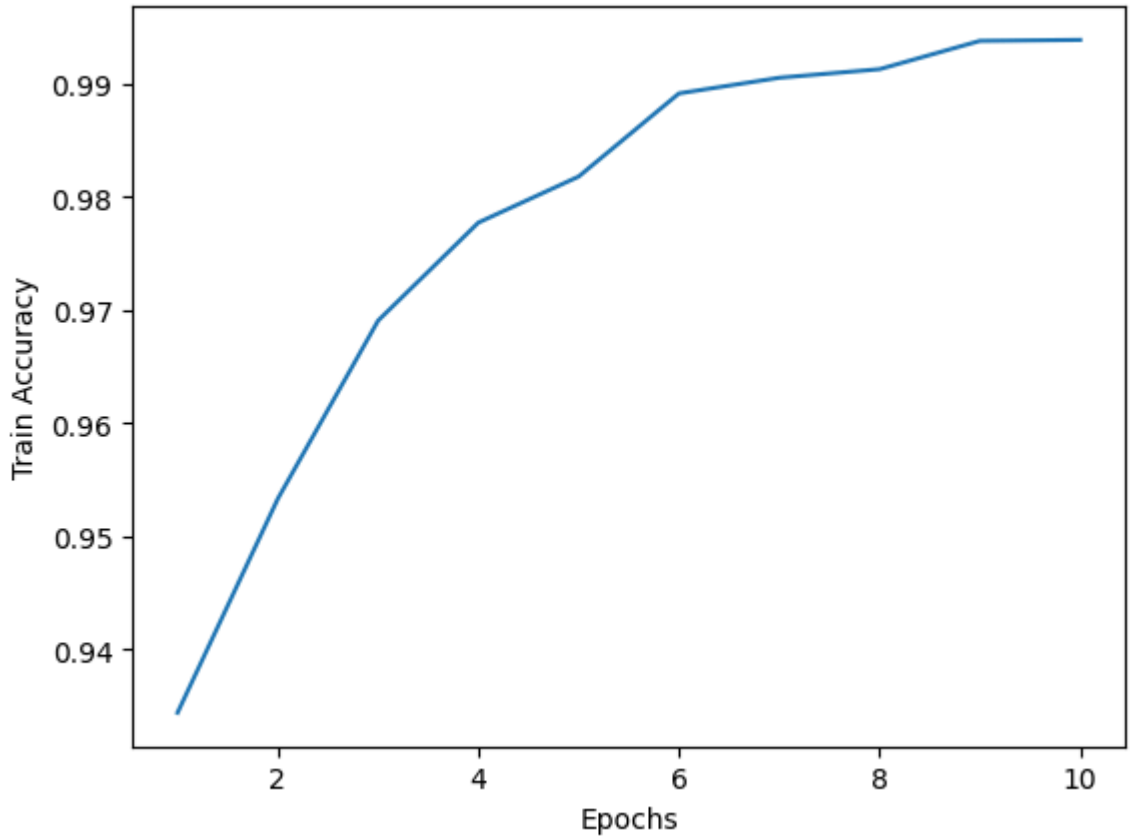



0.66
0.64
0.62
0.60
0.58
0.56
Validation F1 Score
2
4
6
8
10
Epochs

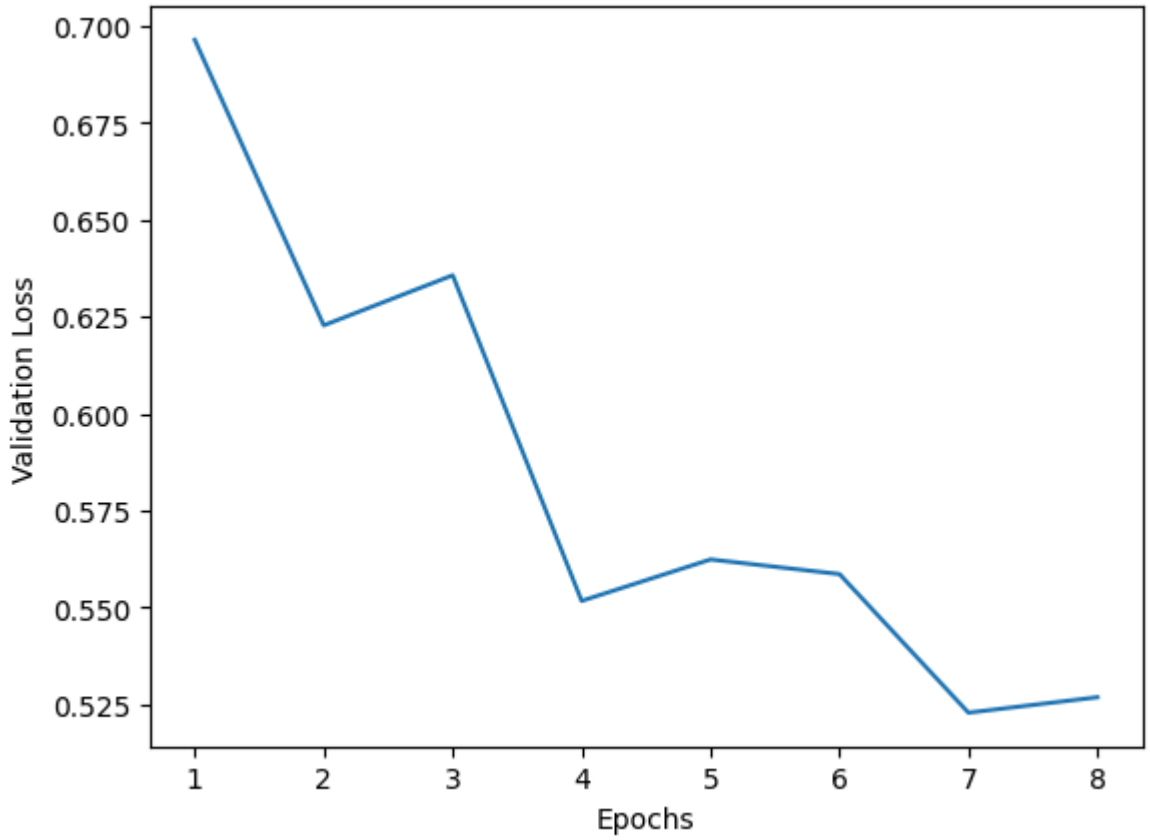

0.938
0.936
0.934
0.932
0.930
0.928
Validation Accuracy
2
4
6
8
10
Epochs

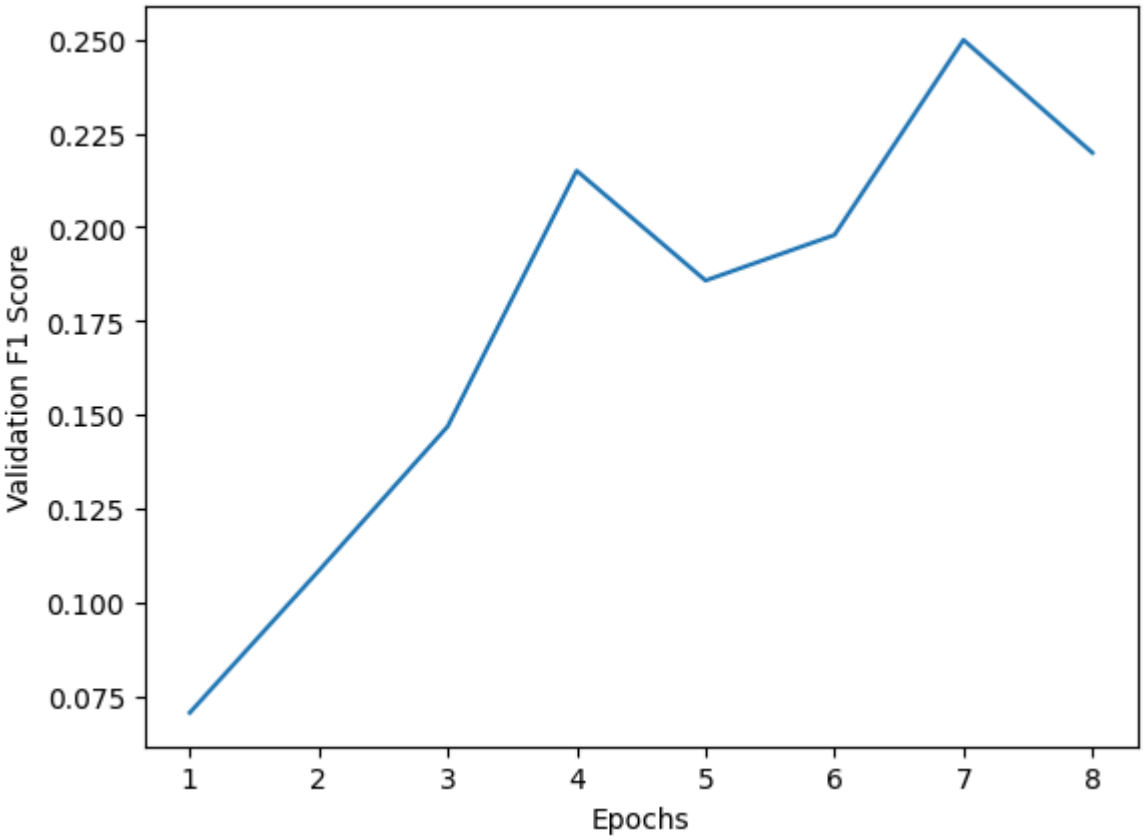

Train Loss
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
2
4
6
8
10
Epochs

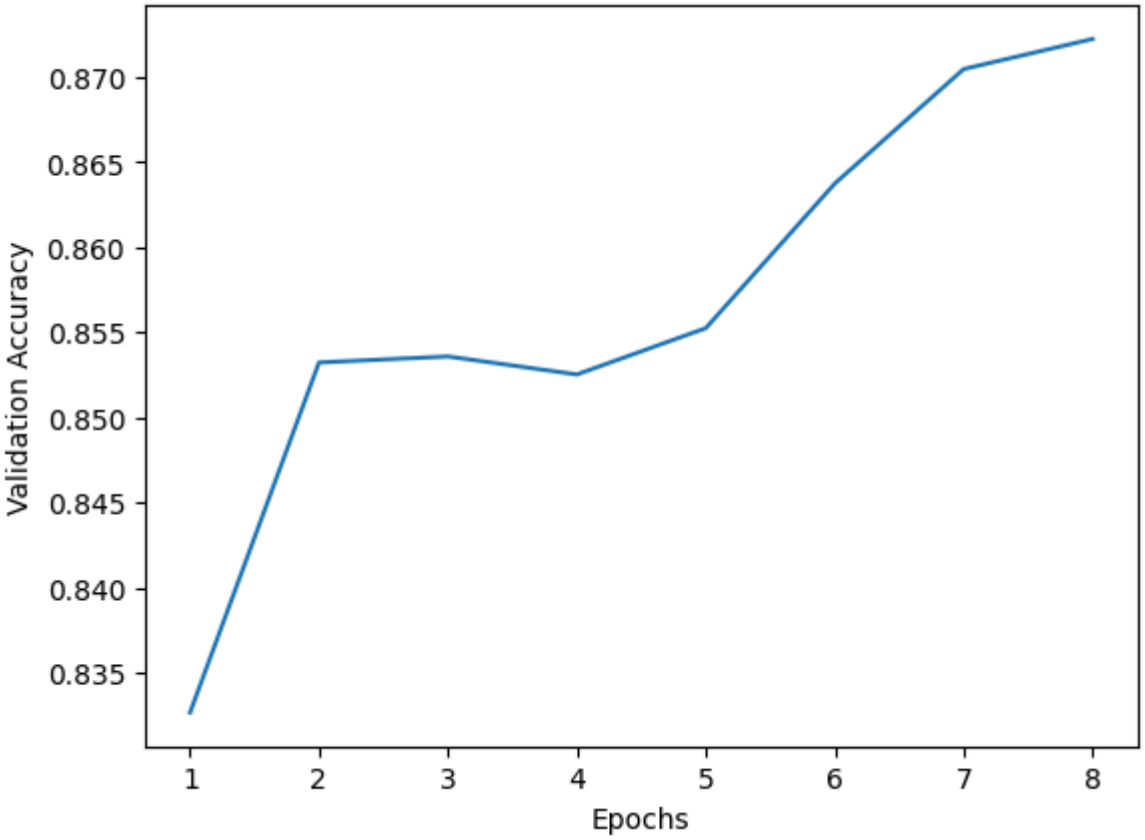

Train F1 Score
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
2
4
6
8
10
Epochs

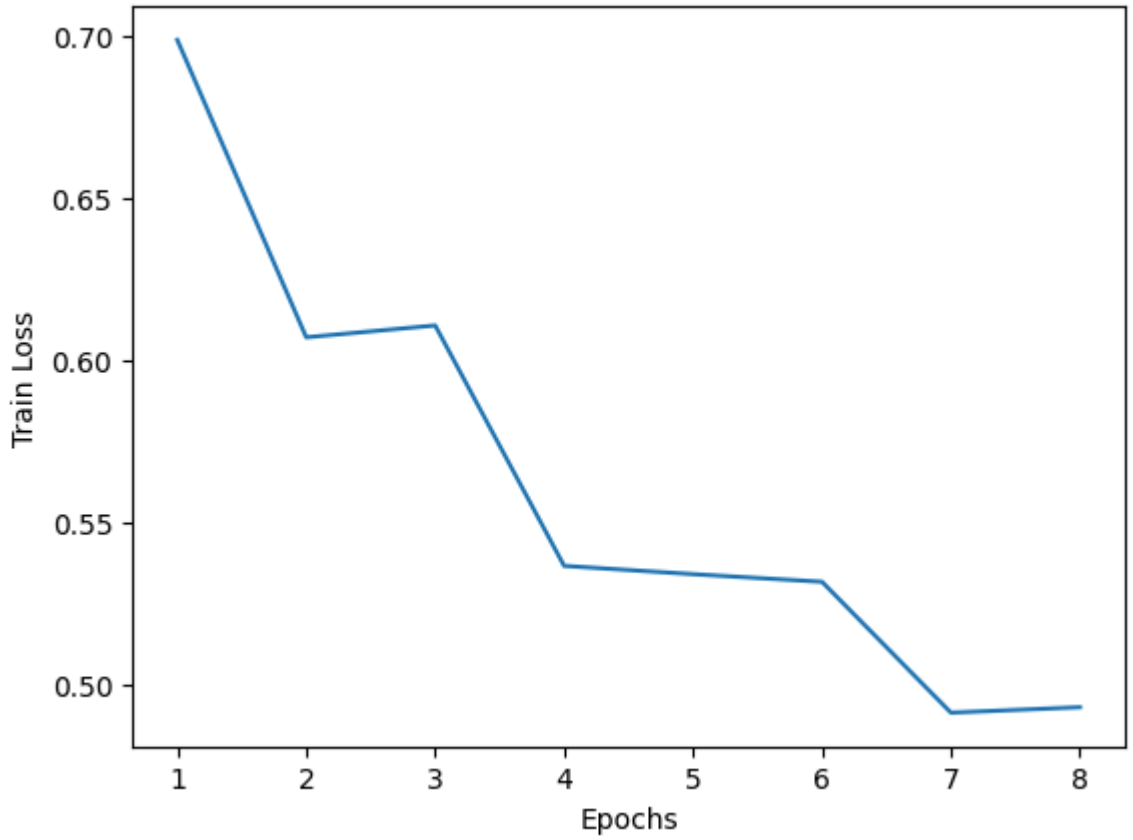


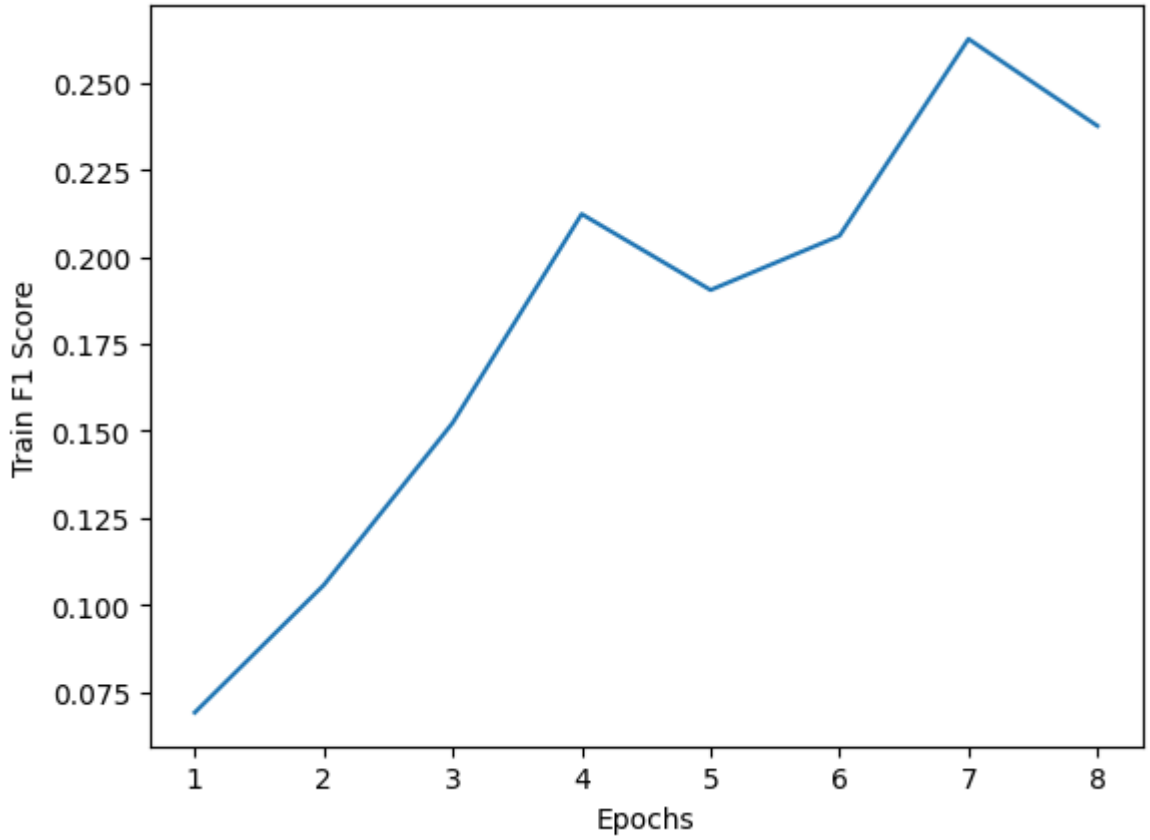
6) RNN with FastText:

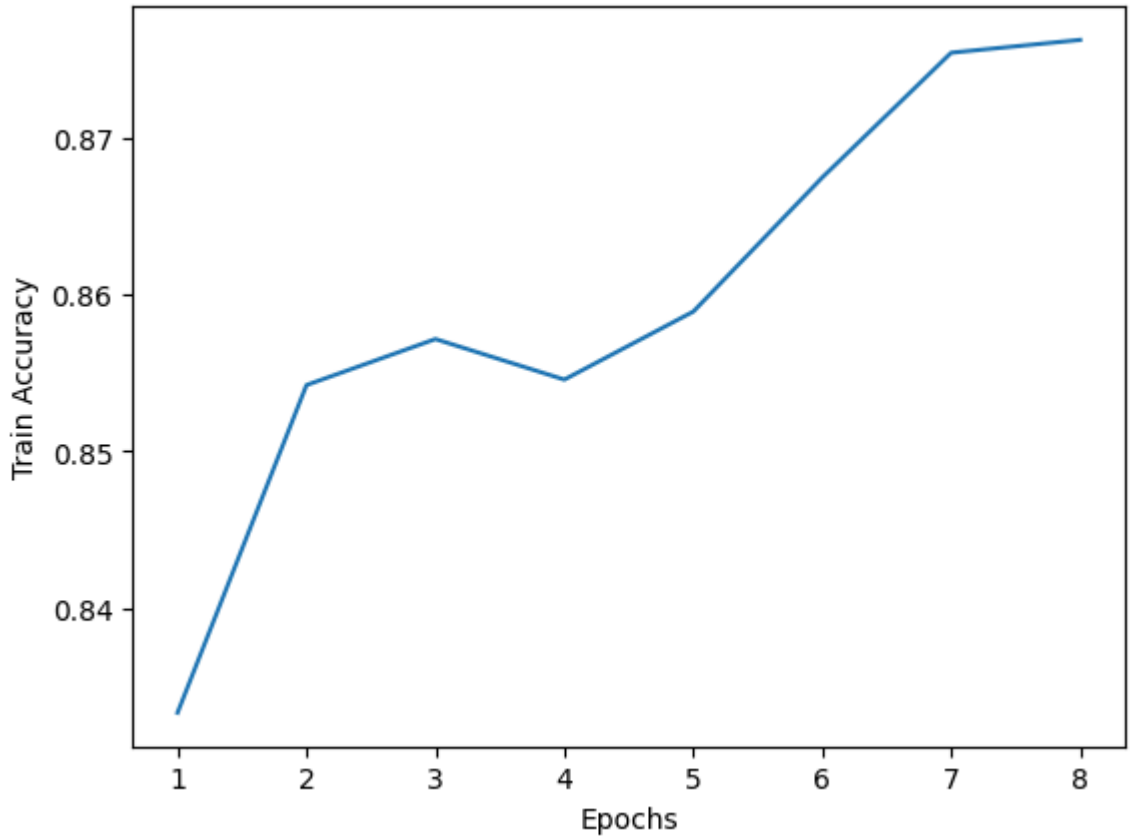



Validation F1 Score
0.250
0.225
0.200
0.175
0.150
0.125
0.100
0.075
1
2
3
4
5
6
7
8
Epochs

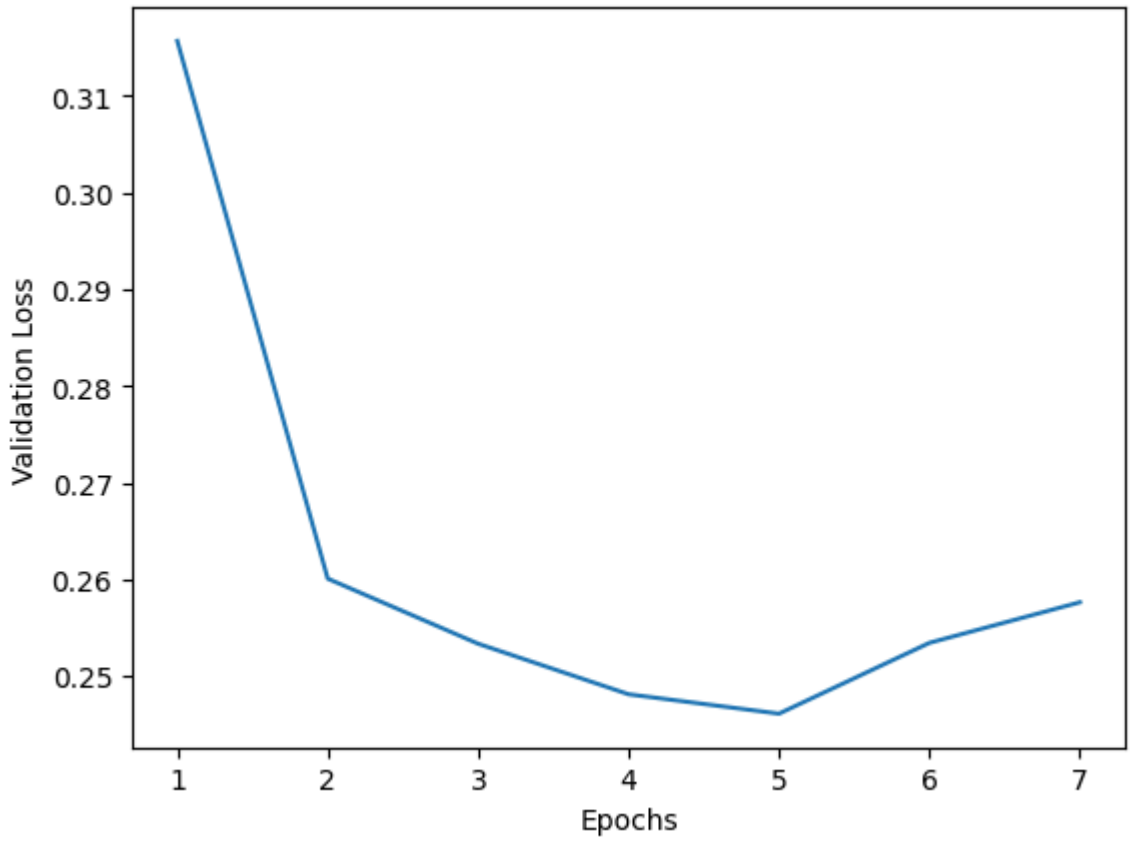

Validation Accuracy
0.870
0.865
0.860
0.855
0.850
0.845
0.840
0.835
1
2
3
4
5
6
7
8
Epochs

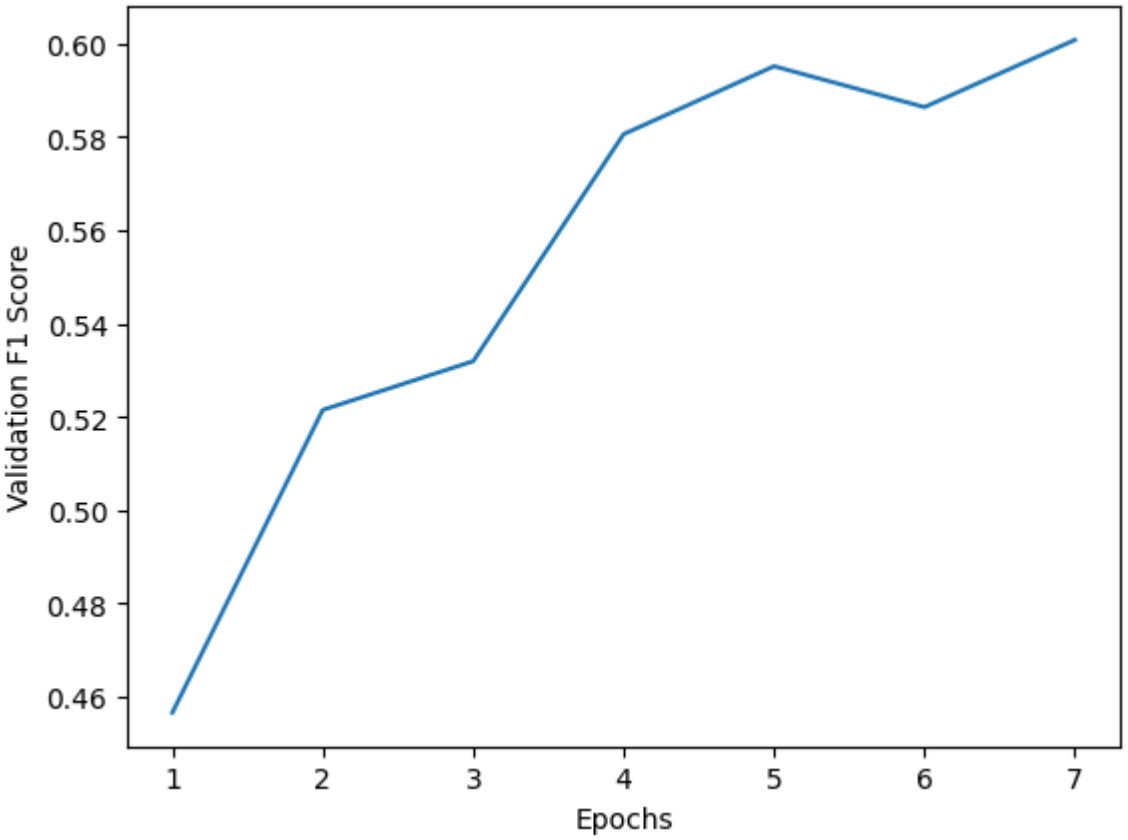

0.70
0.65
0.60
0.55
0.50
Train Loss
1
2
3
4
5
6
7
8
Epochs

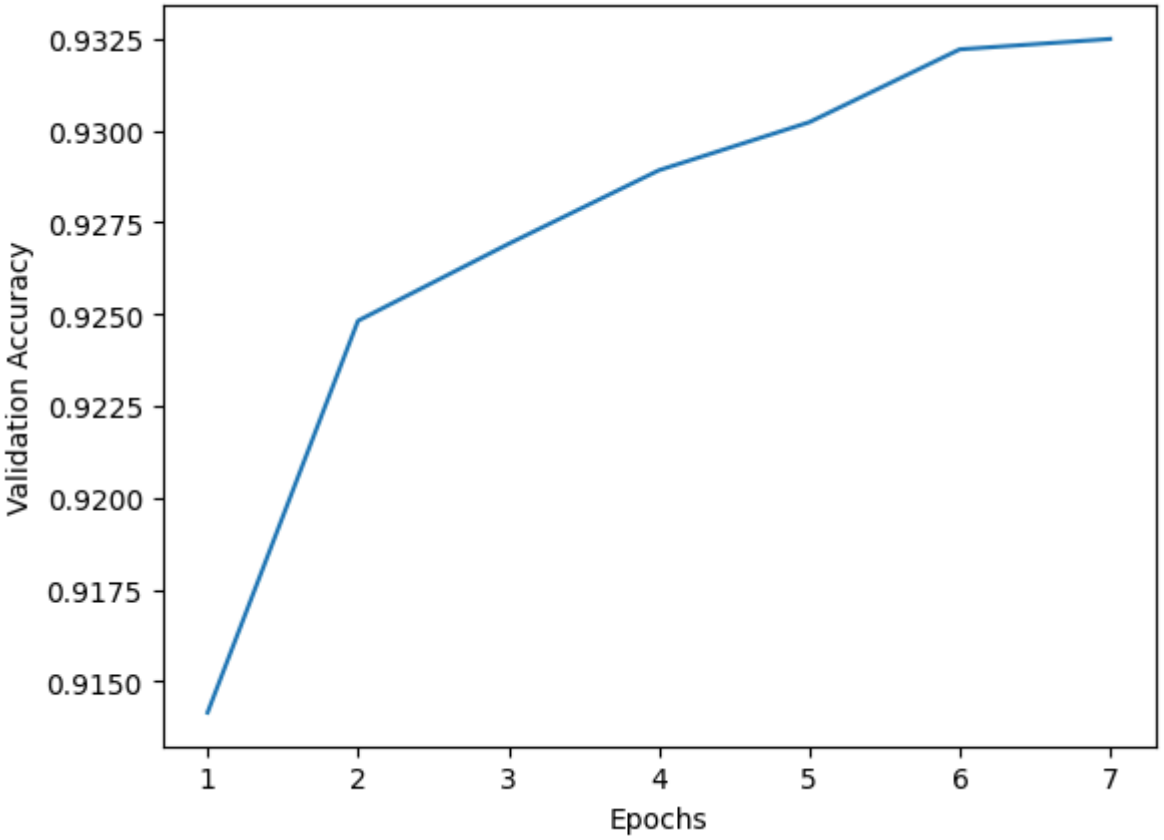

0.250
0.225
0.200
0.175
0.150
0.125
0.100
0.075
Train F1 Score
1
2
3
4
5
6
7
8
Epochs

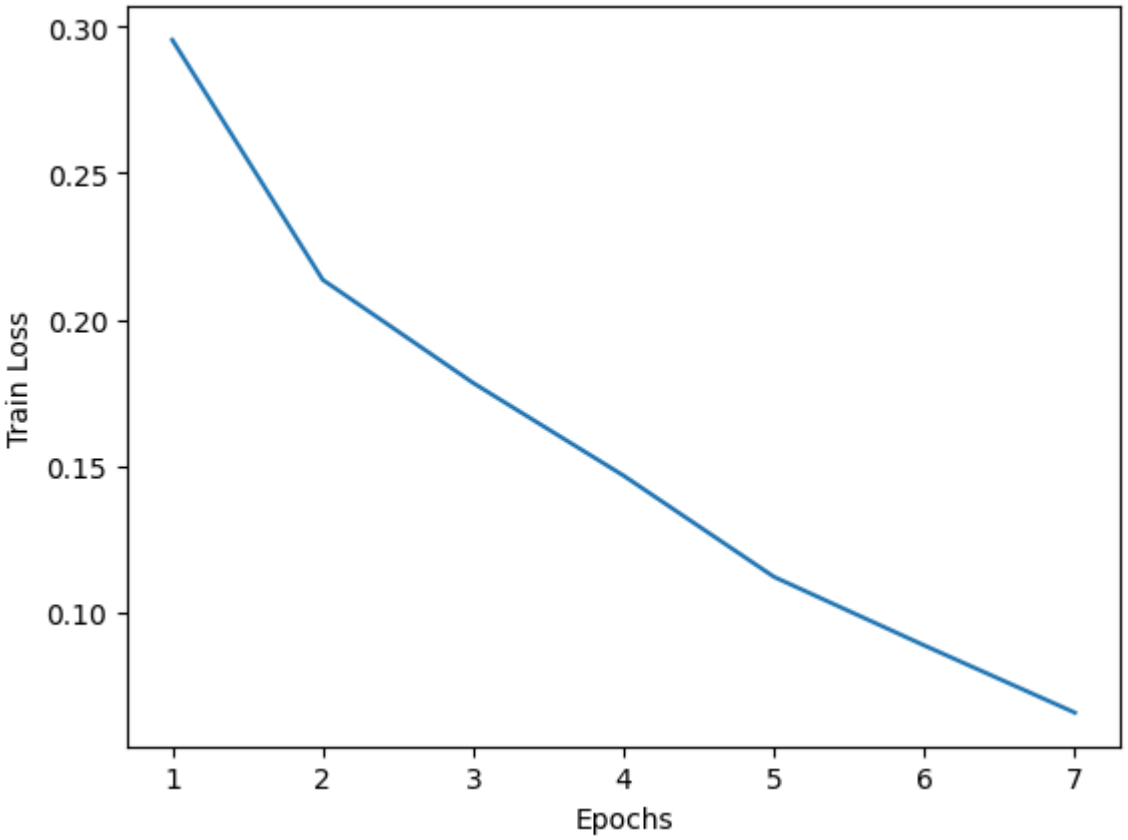


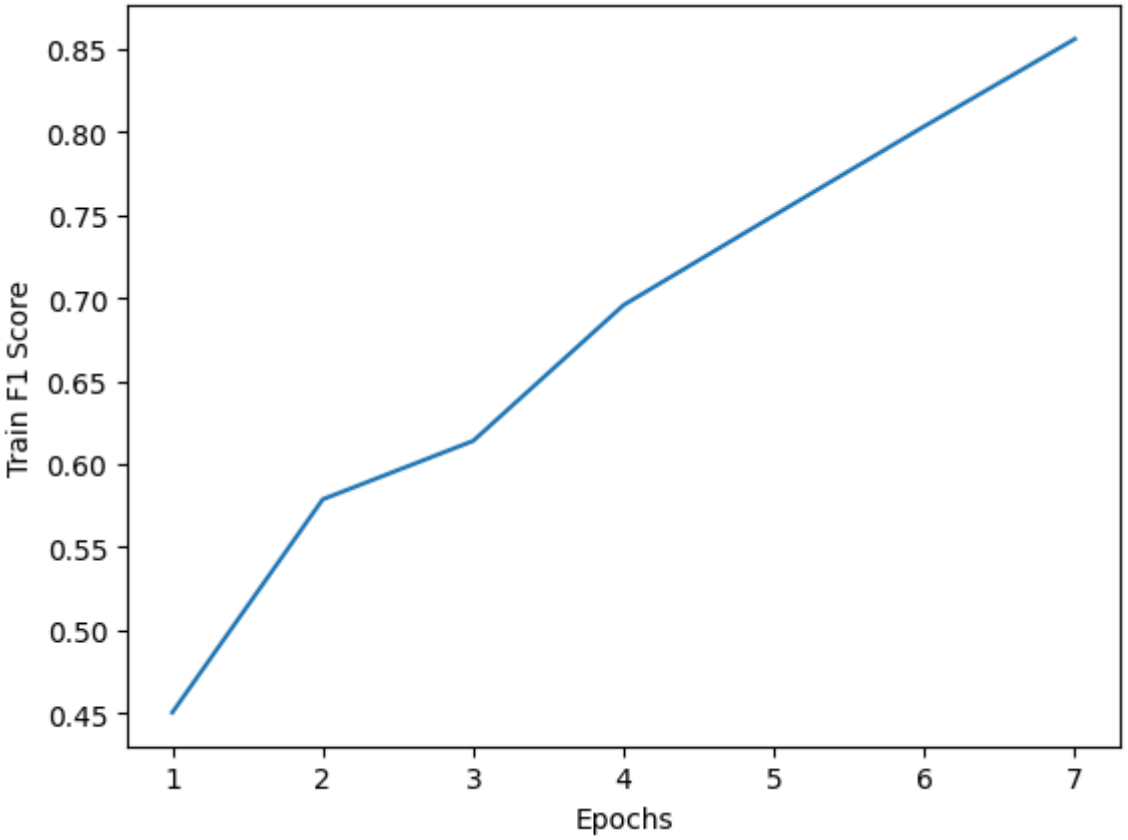
7) LSTM with FastText:

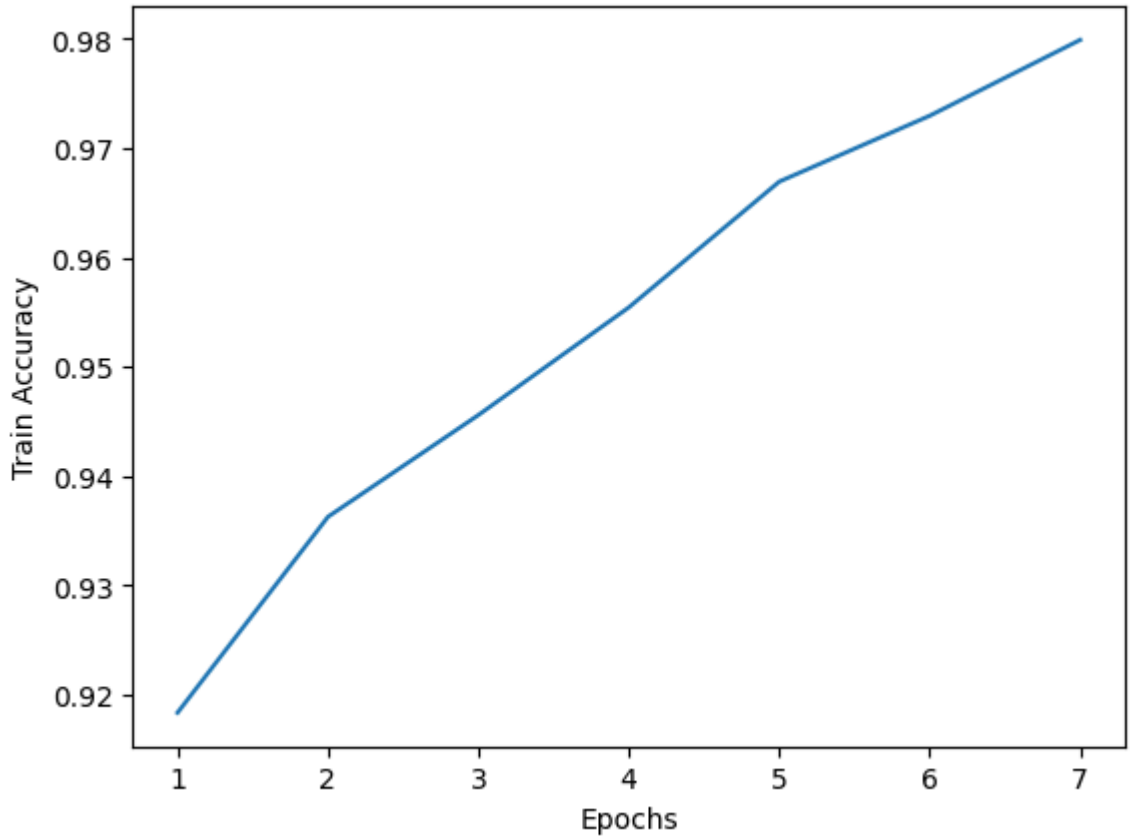



Validation F1 Score
0.60
0.58
0.56
0.54
0.52
0.50
0.48
0.46
1
2
3
4
5
6
7
Epochs

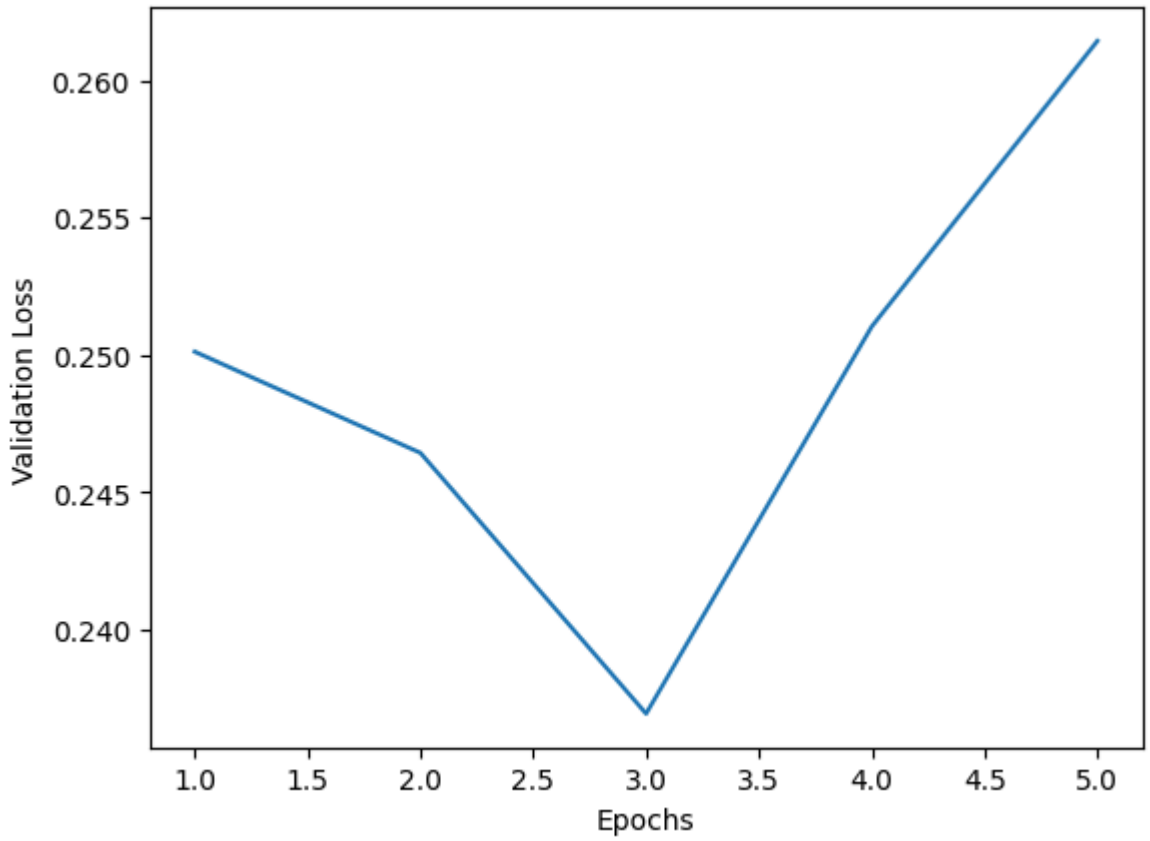

Validation Accuracy
0.9325
0.9300
0.9275
0.9250
0.9225
0.9200
0.9175
0.9150
1
2
3
4
5
6
7
Epochs

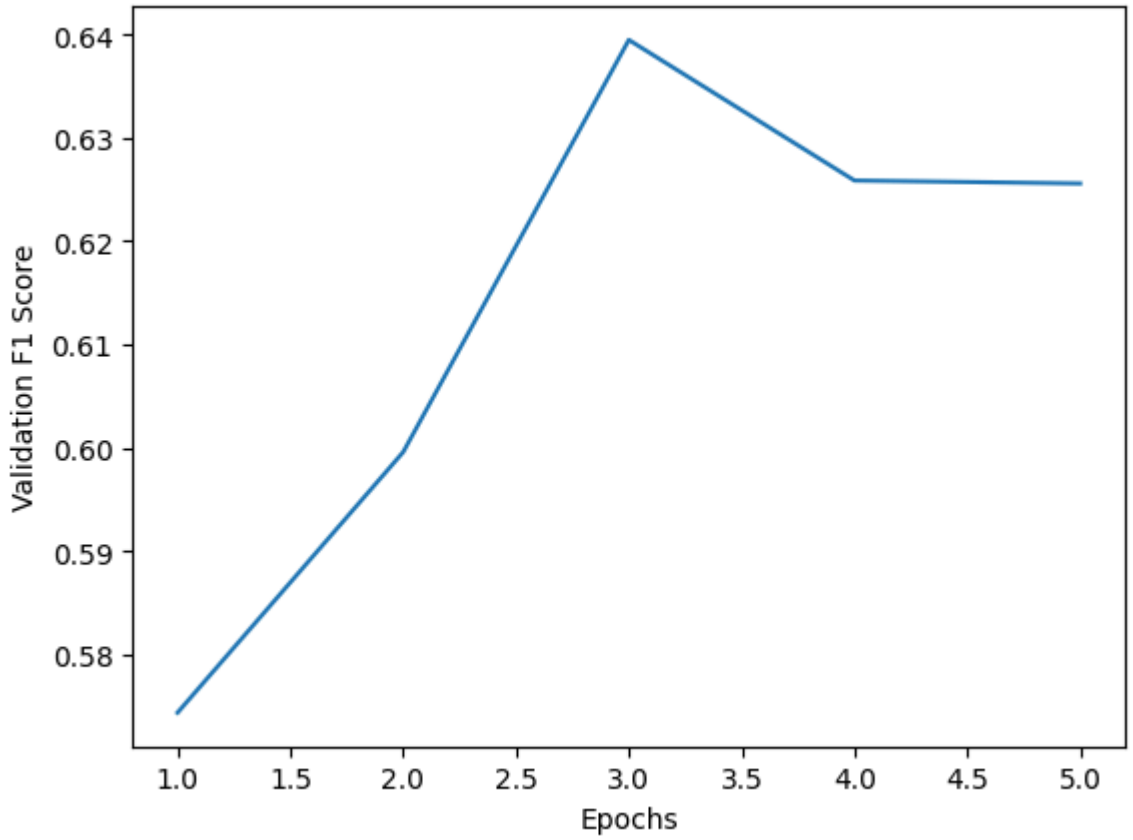

0.30
0.25
0.20
0.15
0.10
Train Loss
1
2
3
4
5
6
7
Epochs

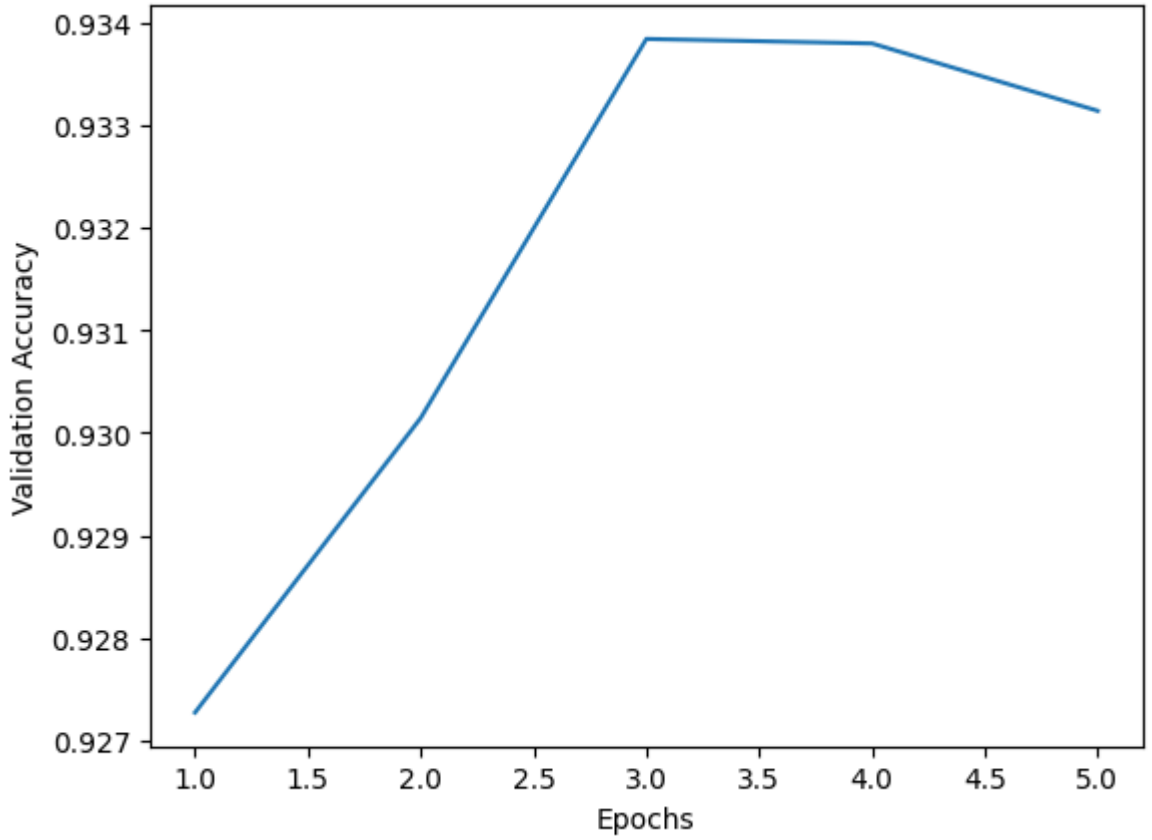

0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
Train F1 Score
1
2
3
4
5
6
7
Epochs

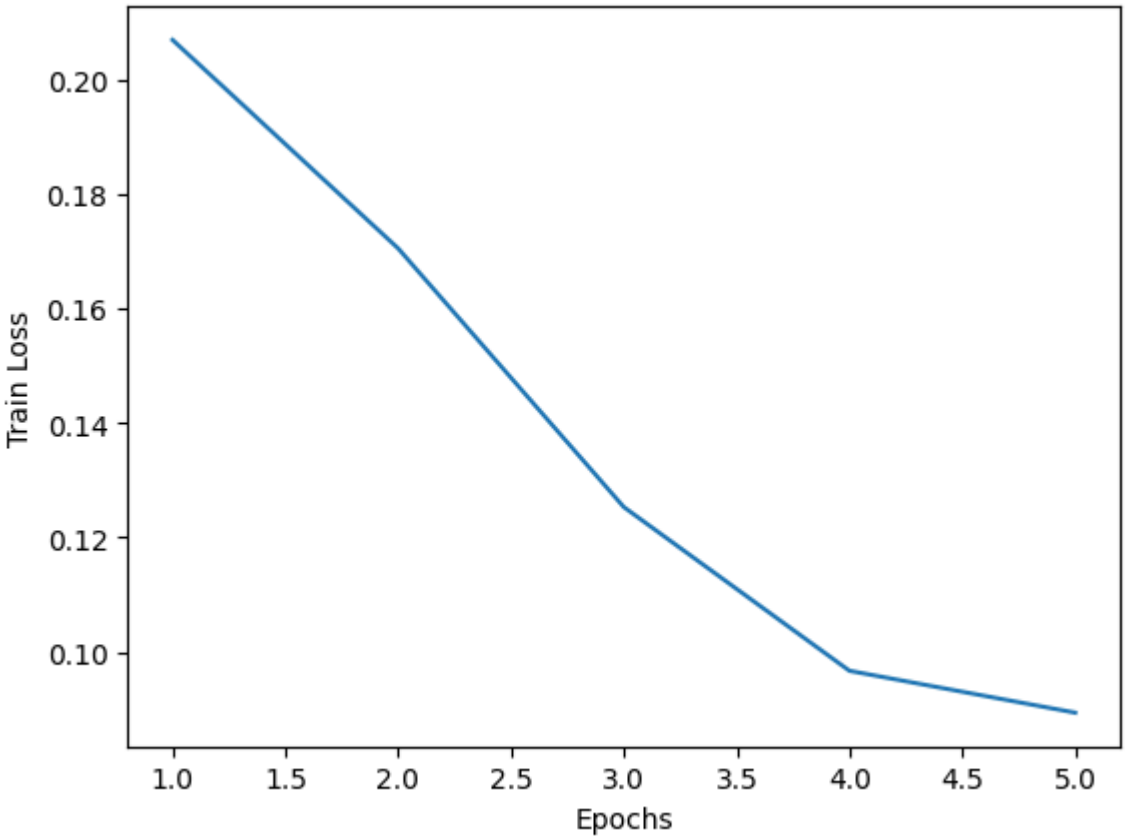


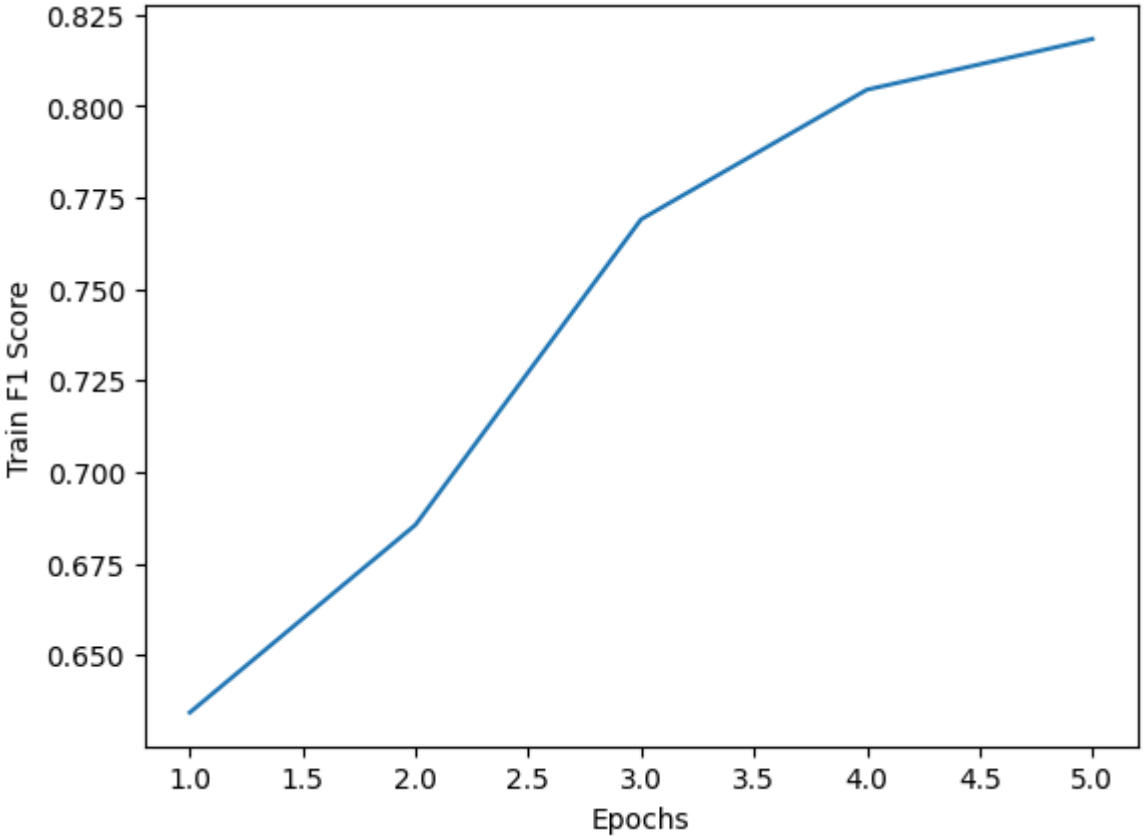
8) GRU with GloVe:

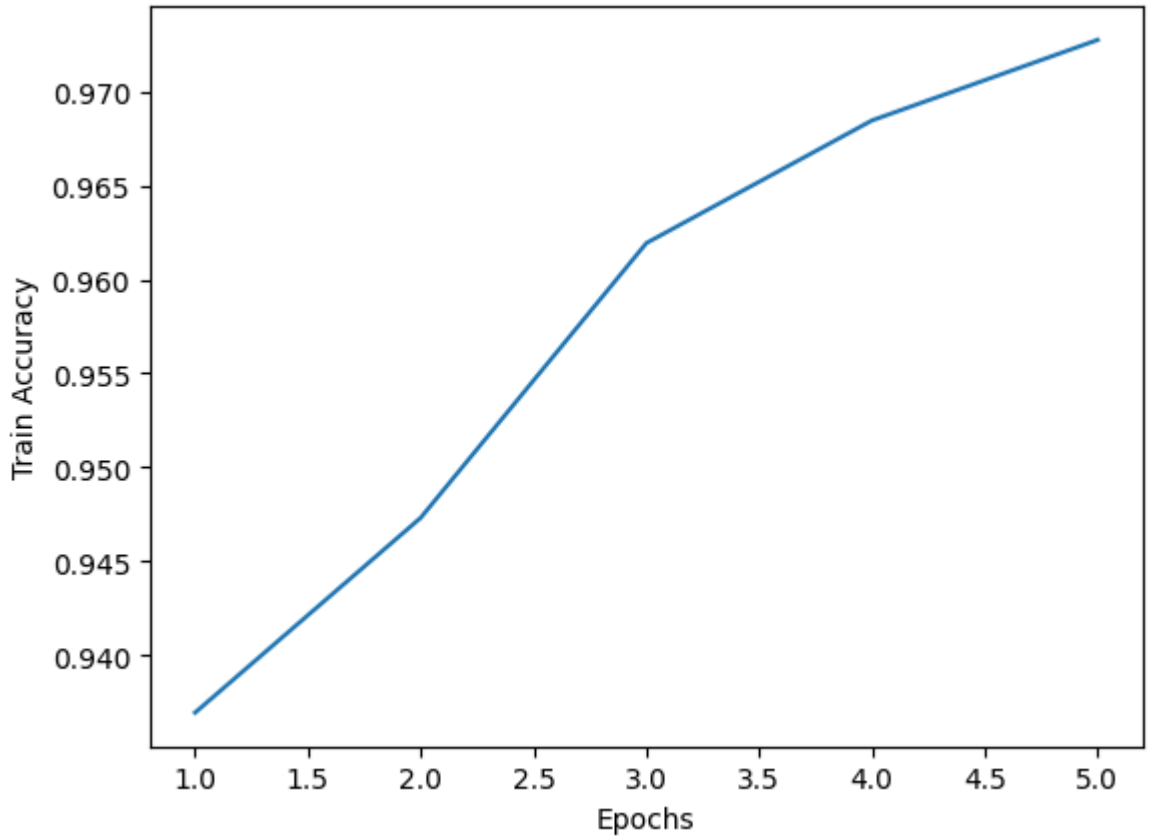



0.64
0.63
0.62
0.61
0.60
0.59
0.58
Validation F1 Score
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs

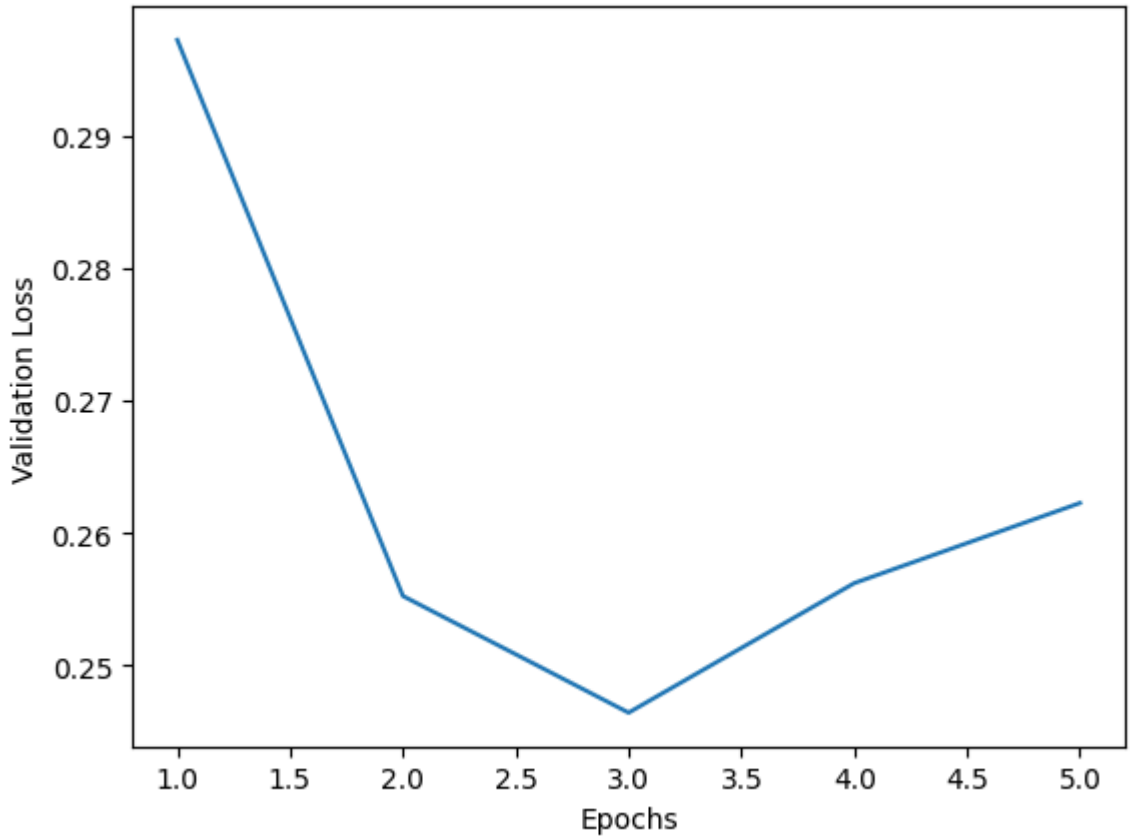

0.934
0.933
0.932
0.931
0.930
0.929
0.928
0.927
Validation Accuracy
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs

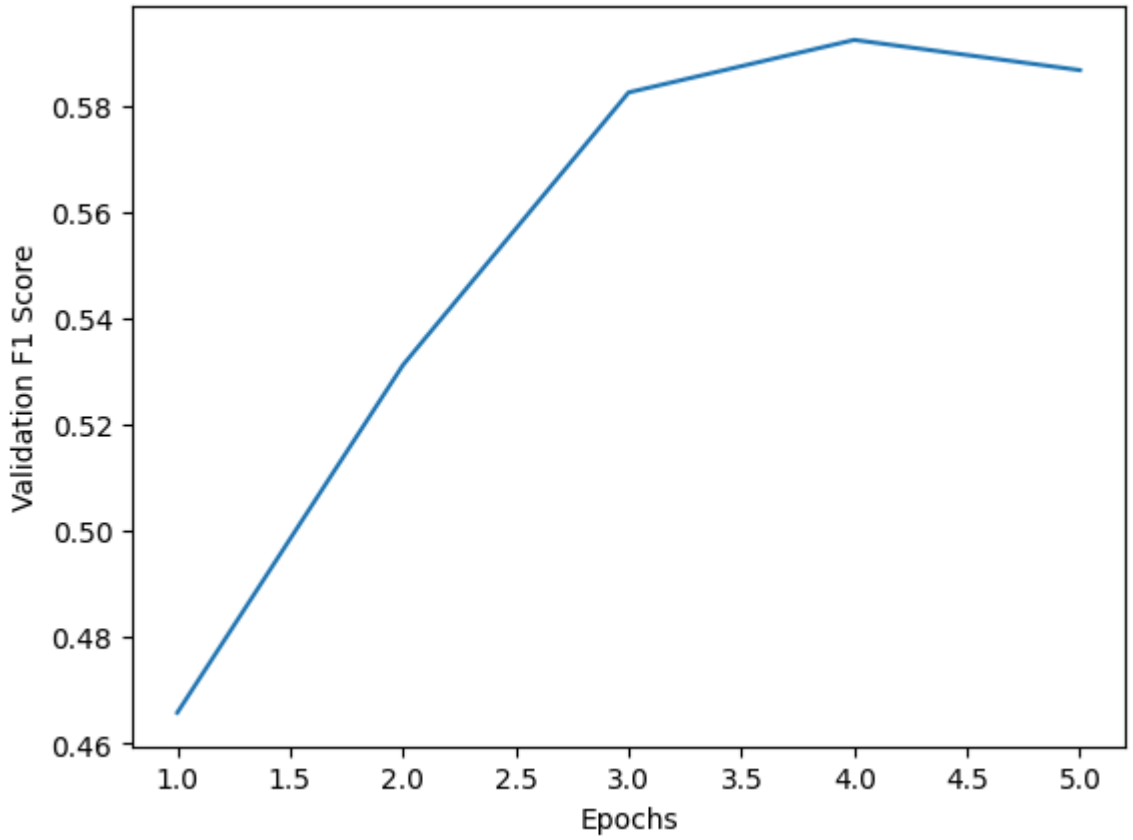

Train Loss
0.20
0.18
0.16
0.14
0.12
0.10
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs

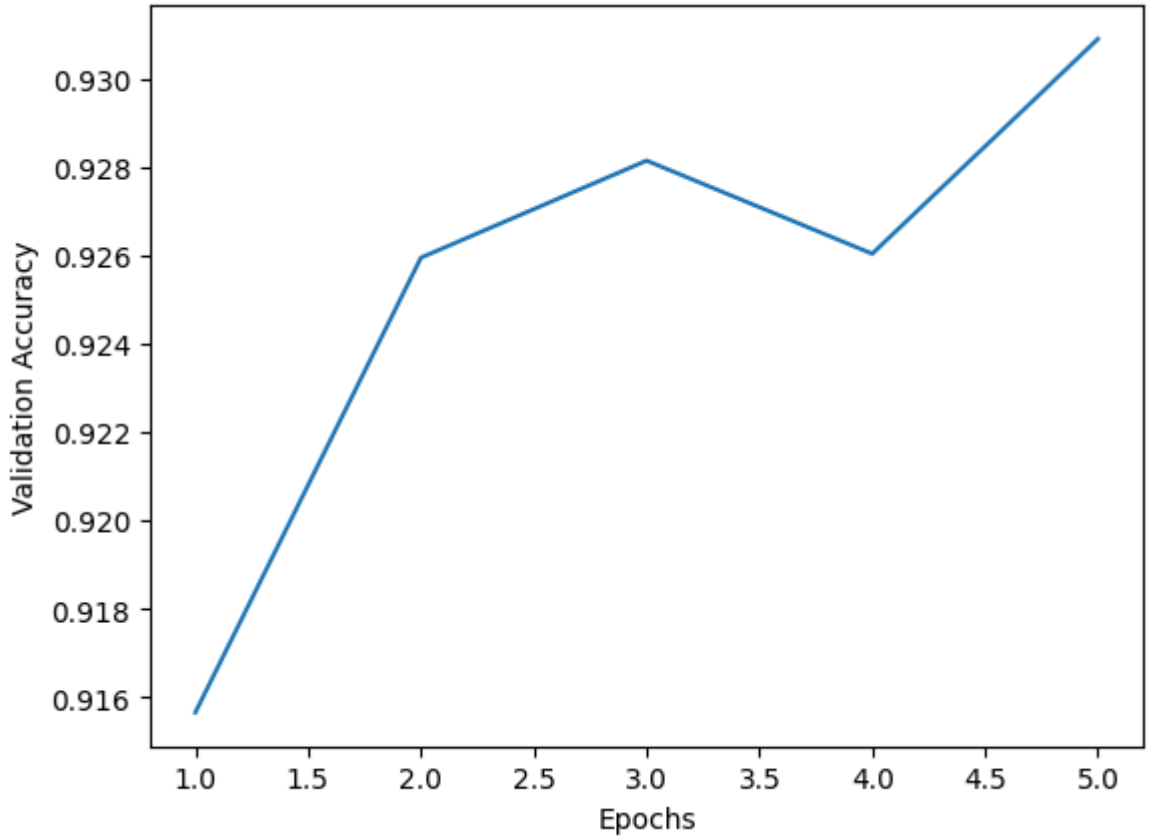

Train F1 Score
0.825
0.800
0.775
0.750
0.725
0.700
0.675
0.650
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs

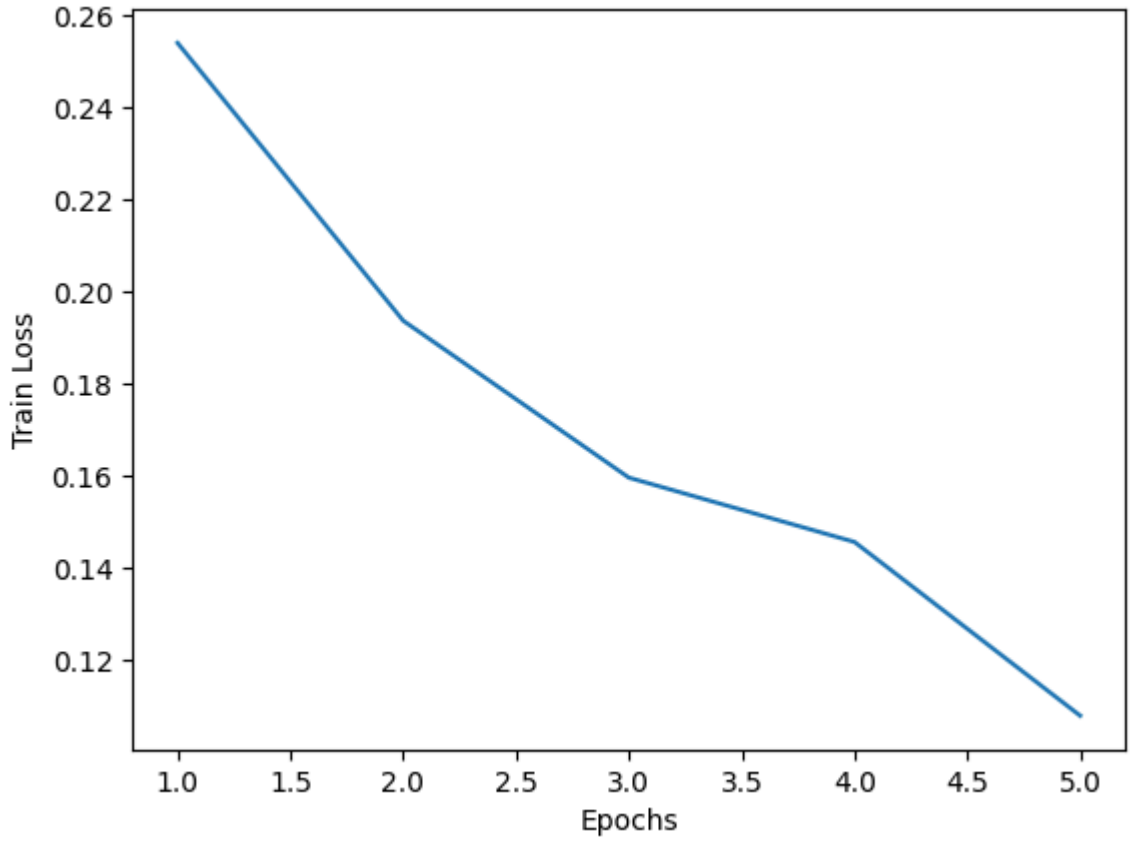


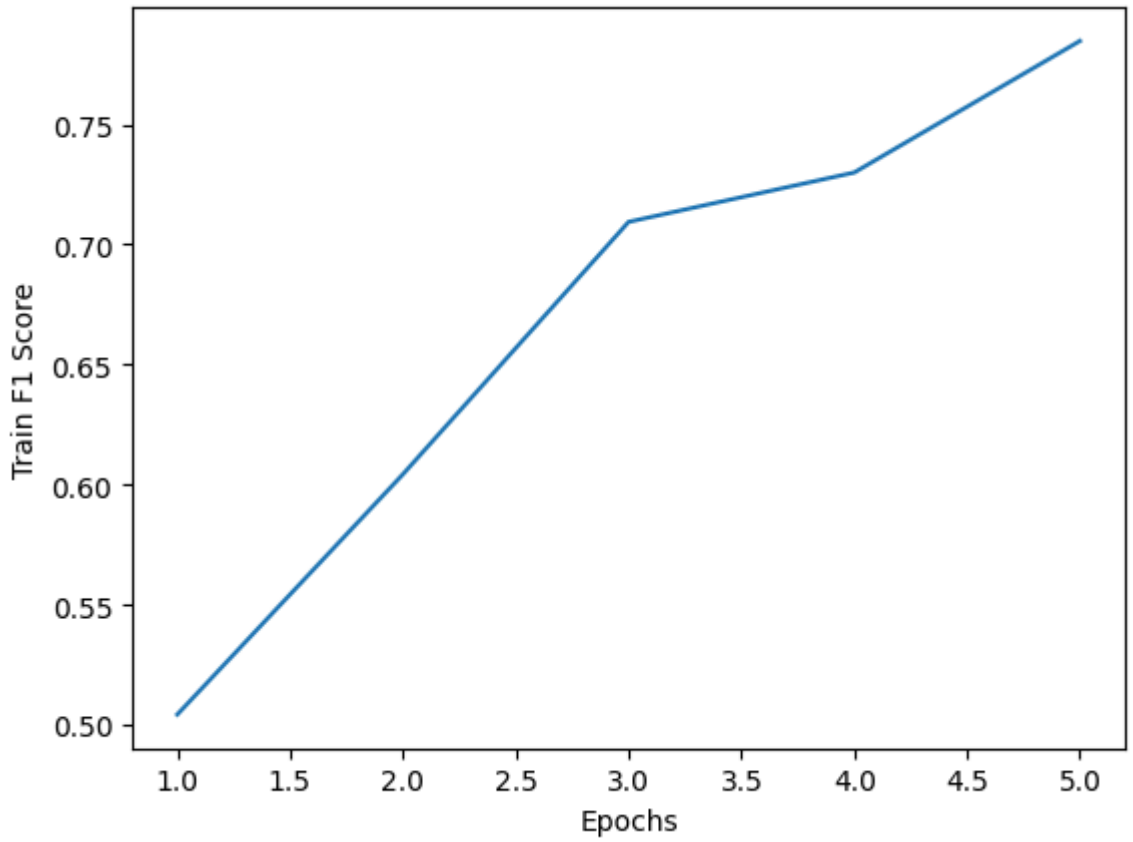
9) GRU with FastText:

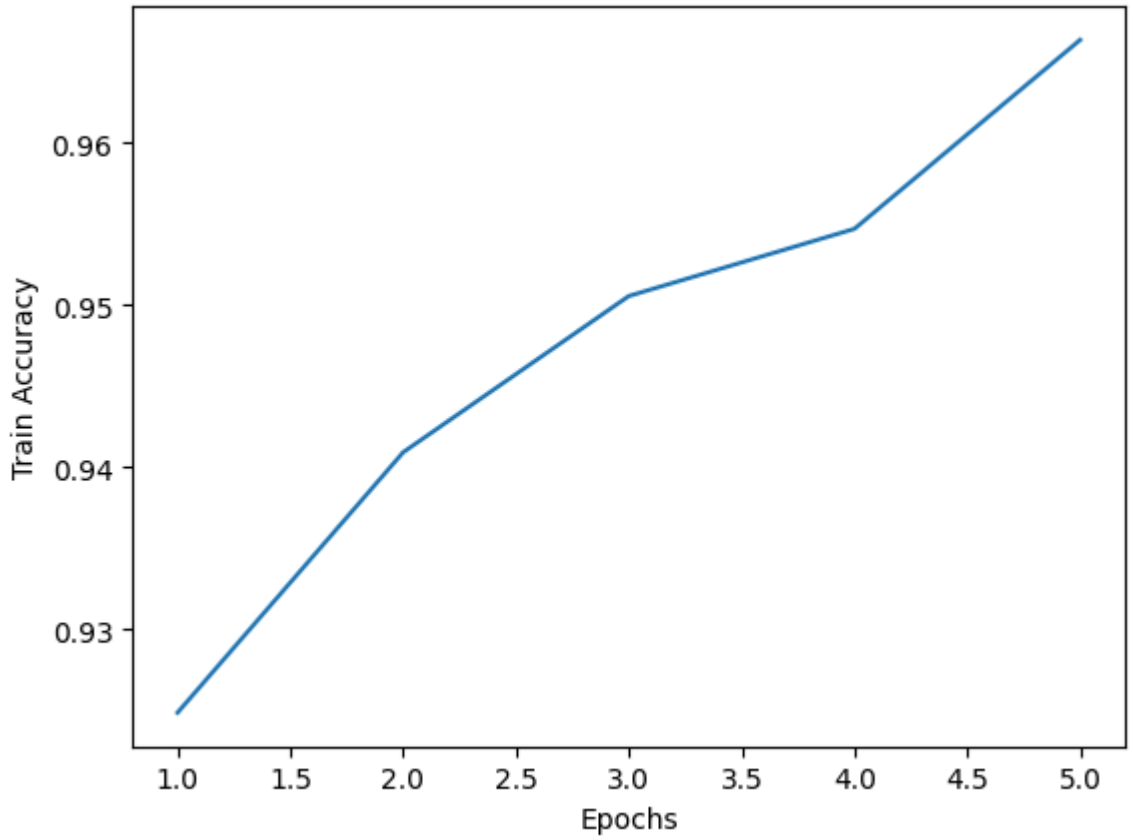



Validation F1 Score
0.58
0.56
0.54
0.52
0.50
0.48
0.46
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs


Validation Accuracy
0.930
0.928
0.926
0.924
0.922
0.920
0.918
0.916
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs


0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
Train Loss
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs


0.75
0.70
0.65
0.60
0.55
0.50
Train F1 Score
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Epochs



Analysis of Graphs for Task 1: In all the graphs we see that the loss or f1 score is fluctuating a lot for Vanilla RNN's while those of LSTM's and GRU's are smooth. Also in RNN's the graph is not monotonic while in LSTM's and GRU's the graphs in most cases are monotonic. This is probably because of vanishing/exploding gradient problems associated with RNN's, as possible sequences are very long. Because of this the loss is higher and f1 score lower in case of RNN's in comparison to LSTM's and GRU's.

Analysis of Graphs for Task-2: The shapes of the graphs are more or less consistent throughout different models and word embeddings, and work as per expectations. The training loss decreases with the number of epochs, while training F1 score increases with epochs. The initial increase/decrease is higher, and this rate keeps decreasing as the loss approaches a minimum. As the number of epochs keep increasing, the model keeps getting better fit on the training data. However, this may cause the problem of overfitting, for which we plot graphs on the validation dataset. We see that the validation loss increases after 5-6 epochs of training, suggesting that after that the model is getting overfit on the training data. However, since our test F1 score was above the required minimum, we let this be. Similarly, the validation F1 score increases till 5-6 epochs, then approaches a constant value, as the learning has more or less completed. To distinguish between the models, we consider the values of the losses in the graphs. The values indicate that the training loss does not reduce with epochs in case of RNN as compared to LSTM and GRU. Also, the validation loss does not increase continuously after 5-6 epochs in RNN, unlike LSTM and GRU. This indicates that overfitting does not happen in case of RNN to the extent that it does in case of

GRU/LSTM. The graphs of RNN are more volatile in general, and this difference can probably be put down to the vanishing/exploding gradient problems in RNN. In BiLSTM-CRF, the values of the losses are much higher in general, which is probably due to the presence of the extra CRF layer which might compound the losses.

JSON File Code Snippets:

TASK-2:

a) Train:


```
"1": {
      "text": "I charge it at night and skip taking the cord with me because of the good battery life .",
      "labels": [
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "B",
        "I",
        "O"
      ]
    },
    "2": {
      "text": "it is of high quality , has a killer GUI , is extremely stable , is highly expandable , is bundled with lots of very good applications , is easy to use , and is absolutely gorgeous .",
      "labels": [
        "O",
        "O",
        "O",
        "O",
        "B",
```

```
            "O",
            "O",
            "O",
            "O",
            "B",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "B",
            "O",
            "O",
            "O",
            "O",
            "B",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O"
        ]
    }
```


b) VALIDATION:


```
    "1": {
        "text": "In the shop , these MacBooks are encased in a soft rubber enclosure - so you will never know about the razor edge until you buy it , get it home , break the seal and use it ( very clever con ) .",
        "labels": [
            "O",
            "O",
            "O",
```

```
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "B",
            "I",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "O"
        ]
    },
    "2": {
```

```
        "text": "This laptop meets every expectation and Windows 7 is great !",
        "labels": [
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "B",
          "I",
          "O",
          "O",
          "O"
        ]
      }
```


c) TEST:


```
    "1": {
        "text": "Boot time is super fast , around anywhere from 35 seconds to 1 minute
    .",
        "labels": [
          "B",
          "I",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O",
          "O"
        ]
      },
      "2": {
        "text": "tech support would not fix the problem unless I bought your plan for $
    150 plus .",
        "labels": [
          "B",
          "I",
          "O",
```

"O",
"O",
"O",
"O",
"O",
"O",
"O",
"O",
"O",
"O",
"O",
"O",
"O",
"O"
]
}

TASK-1:

a) TRAIN:
"bf0edaaf64384792ae08beade375cd39": {"text": "In other words, on the delaying of payment of compensation on unjustifiable ground, the employer does not earn any penalty, if the claimant is being compensated with higher rate of interest.", "labels": ["O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O"]}, "39afb0578cba4848a4749c1e65e8bea3": {"text": "This in the view of the High Court would impede the freedom of inter-State trade, com- merce and intercourse within the meaning of Art. 301 of the Constitution and was not saved by Art. 303.", "labels": ["O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "B_PROVISION", "I_PROVISION", "O", "O", "B_STATUTE", "O", "O", "O", "O", "O", "B_PROVISION", "I_PROVISION"]}

b) VALIDATION:
"46ca6551ca8747f59de909b4d2ecc253": {"text": "In the said reply, it has been denied that respondent No. 2 (the Collector, Bikaner) has gone to the spot or that any police force had been sent and it has been asserted that the possession of the property was demanded, delivered and taken over in a peaceful manner after preparing a panchnama at the spot.", "labels": ["O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "B_GPE", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O"]}, "87dc90d9302648f489430b82ef9bb082": {"text": "The Respondent shall intimate to the claimants/ petitioners about it having deposited the cheques in

favour of petitioner in terms of the award, at the Suit No.838/11, Ram Lalli & ors. Vs Ratan Singh & anr. (pg- 14 of 15) address of the petitioner mentioned at the title of the award, so as to facilitate them to withdraw the same.", "labels": ["O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "B_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O"]}

c) TEST:
"03f3901e95ed493b866bd7807f623bc0": {"text": "True, our Constitution has no 'due process' clause or the VIII Amendment; but, in this branch of law, after R.C. Cooper v. Union of India, (1970) 1 SCC 248 and Maneka Gandhi v. Union of India, (1978) 1 SCC 248, the consequence is the same.", "labels": ["O", "O", "B_STATUTE", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "B_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "O", "B_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "I_PRECEDENT", "O", "O", "O", "O", "O"]}, "b0311cba3aac4d909eec6e156c059617": {"text": "(See Principles of Statutory Interpretation by Justice G.P. Singh, 9th Edn., 2004 at p. 438.).", "labels": ["O", "O", "O", "O", "O", "O", "O", "B_JUDGE", "I_JUDGE", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O"]}, "0215658407ec479fa4e0e73ca0bf98b6": {"text": "Their Lordships have said -- \"It is a sound rule of construction of a statute firmly established in England as far back as 1584 when Heydon's case was decided that --\".....", "labels": ["O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "O", "B_GPE", "O", "O", "O", "O", "O", "O", "B_OTHER_PERSON", "O", "O", "O", "O", "O"]}

TEAM MEMBER CONTRIBUTIONS:

Task-1, Part-1: Vatsal
Task-1, Part-2: Vatsal
Task-2, Part-1: Jay
Task-2, Part-2: Rohan (RNN, LSTM), Jay (GRU)
Task-2, Part-3: Harsh

GOOGLE DRIVE LINK FOR SAVED PYTORCH MODELS FOR TASK-1
(included like this due to large size):
https://drive.google.com/drive/folders/19tMhWJYWiyPbH_14rI7zckBHSbSiB_7c?usp=drive_link

Team Members:

1) Harsh (2020061)
2) Jay Saraf (2020438)
3) Rohan Gupta (2020113)
4) Vatsal Chaudhary (2020549)