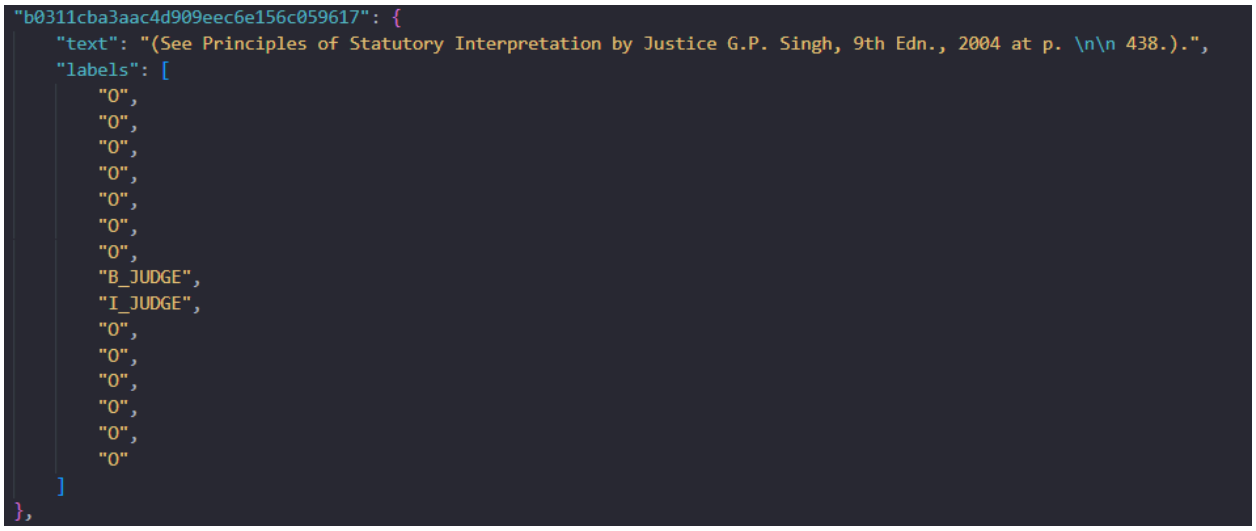
NLP Assignment 2 Report

Saumil Lakra *Sanskar Ranjan* *Jeremiah Rokhum* *Vishal Singh*

Task 1: NER Dataset

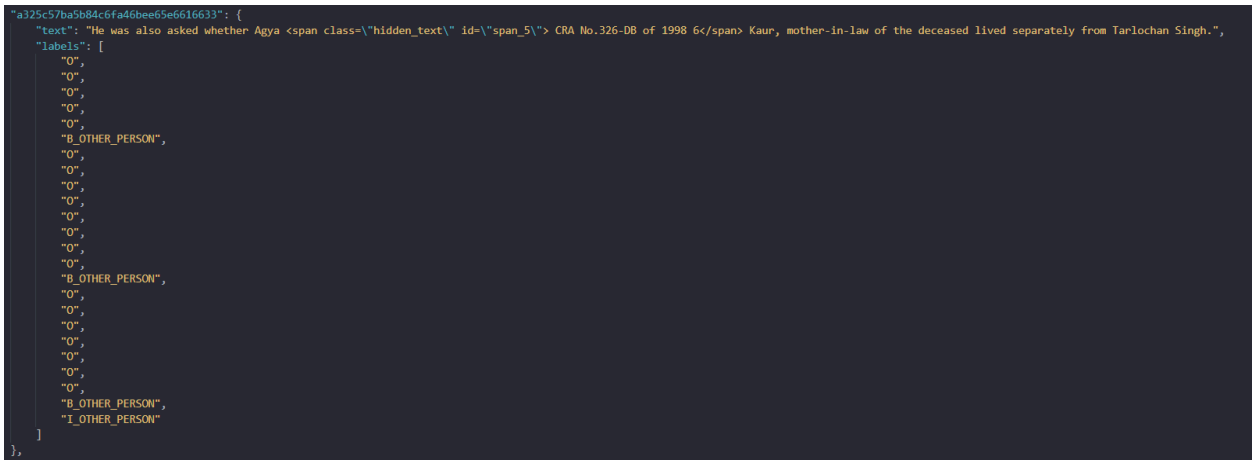
The final processed data looks like for the NER Dataset.

This is a screenshot of one of the testing data.



```
"b0311cba3aac4d909eec6e156c059617": {
    "text": "(See Principles of Statutory Interpretation by Justice G.P. Singh, 9th Edn., 2004 at p. \n\n 438.).",
    "labels": [
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "B_JUDGE",
        "I_JUDGE",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O"
    ]
},
```

This is a screenshot of an instance of the training data.



```
"a325c57ba5b84c6fa46bee65e6616633": {
    "text": "He was also asked whether Agya <span class=\"hidden_text\" id=\"span_5\"> CRA No.326-DB of 1998 6</span> Kaur, mother-in-law of the deceased lived separately from Tarlochan Singh.",
    "labels": [
        "O",
        "O",
        "O",
        "O",
        "O",
        "B_OTHER_PERSON",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "B_OTHER_PERSON",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "B_OTHER_PERSON",
        "I_OTHER_PERSON"
    ]
},
```

This is a screenshot of an instance of the validation.json file.

```
"9945491a6a894d8583ddba684548ce91": {
    "text": "It is contended on behalf of the petitioners that petitioner no.3, Ms. Shraddha Bhargava is a practicing lawyer at District Courts and High Court of Delhi.",
    "labels": [
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "B_PETITIONER",
        "I_PETITIONER",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "B_COURT",
        "I_COURT",
        "I_COURT",
        "I_COURT"
    ]
},
```


Graphs:

Points to note: Training and Validation losses usually decrease until an epoch, and then, training loss drops and tends to become constant while validation losses keep growing. This is basically the point where overfitting happens.

Underfitting: When the training dataset is too complex for our model. It can't capture them. The gap between the curve basically means that, the model is missing something

Task 2: ATE Dataset


```
{
    "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",
            "O"
        ]
    },
```


This is a screenshot of one of the instances of the testing data.

```
{
    "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"
        ]
    },
```


This is a screenshot of an instance of the training data.

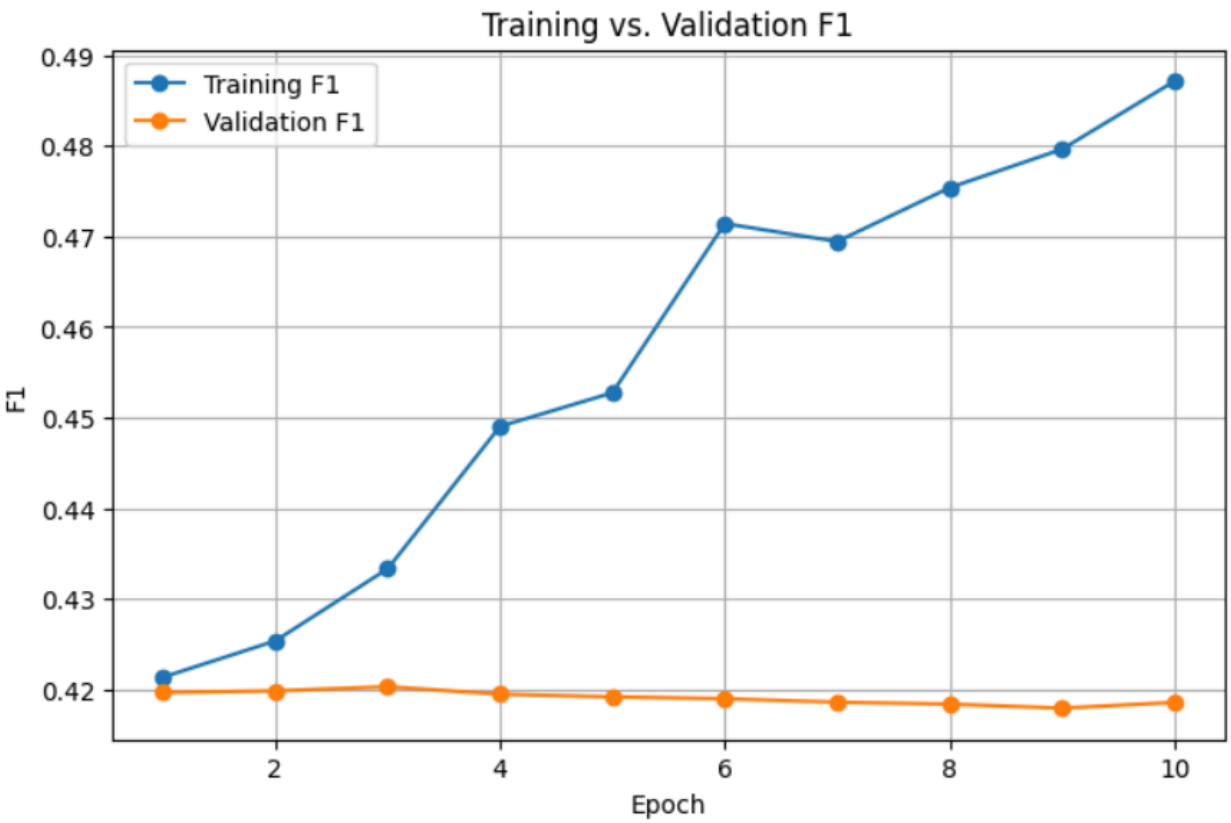

```
    "2": {
        "text": "This laptop meets every expectation and Windows 7 is great !",
        "labels": [
            "O",
            "O",
            "O",
            "O",
            "O",
            "O",
            "B",
            "I",
            "O",
            "O",
            "O"
        ]
```


This is a screenshot of an instance of the validation data.

Word2Vec

Vanilla Rnn:

The training loss (blue line) steadily decreases as the number of epochs increases, indicating the model is effectively learning from the training data. However, the validation loss (orange line) remains flat, suggesting the model might be overfitting to the training data. While the model exhibits a decreasing trend in training loss, there is a risk of overfitting as the validation loss fails to show a similar improvement.

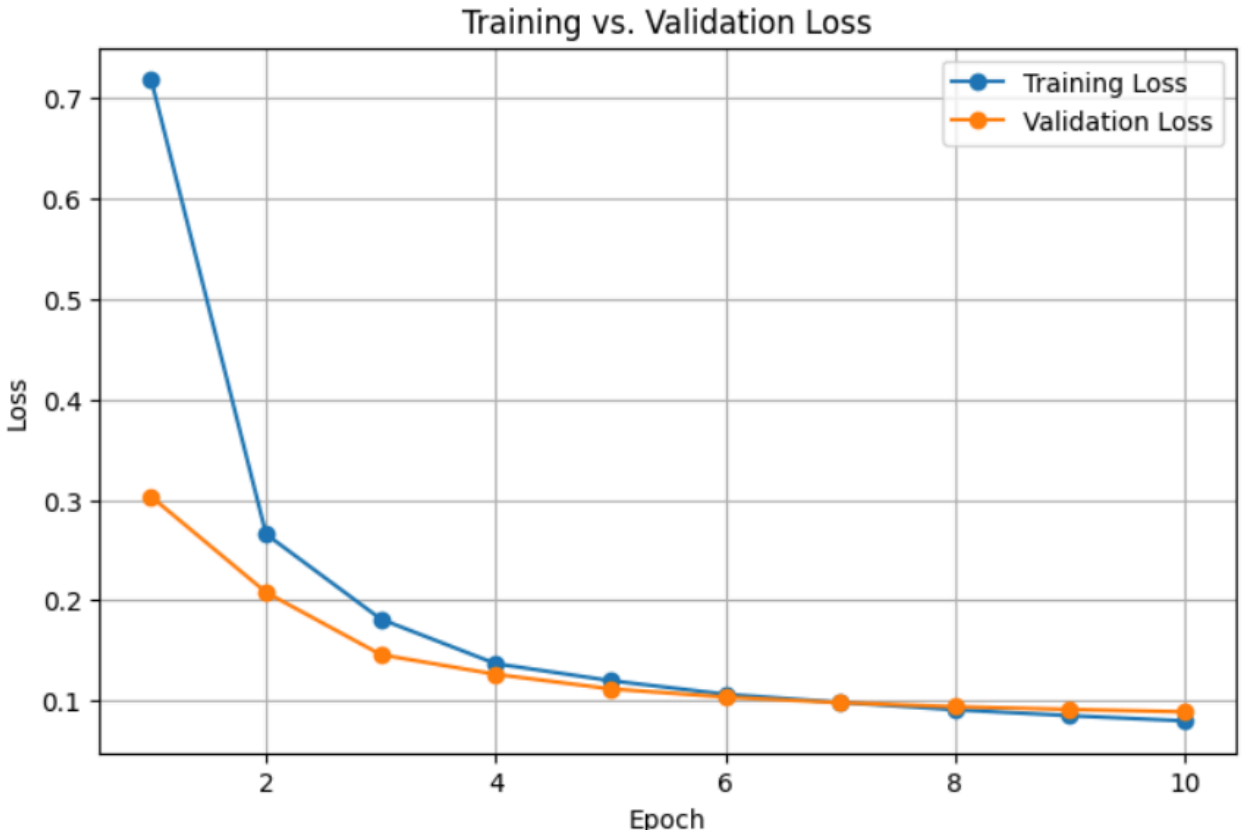


The graph shows encouraging signs of model improvement on the training data, as the training macro F1 score (blue line) increases with each epoch. However, the validation macro F1 score (orange line) stagnates, which suggests the model might be overfitting to the training data and failing to generalize well to unseen data. This graph visualizes the model's performance during training. While the training macro F1 score increases (blue line), indicating the model is learning the training data, the validation macro F1 score shows minimal improvement (orange line). This implies a risk of overfitting.

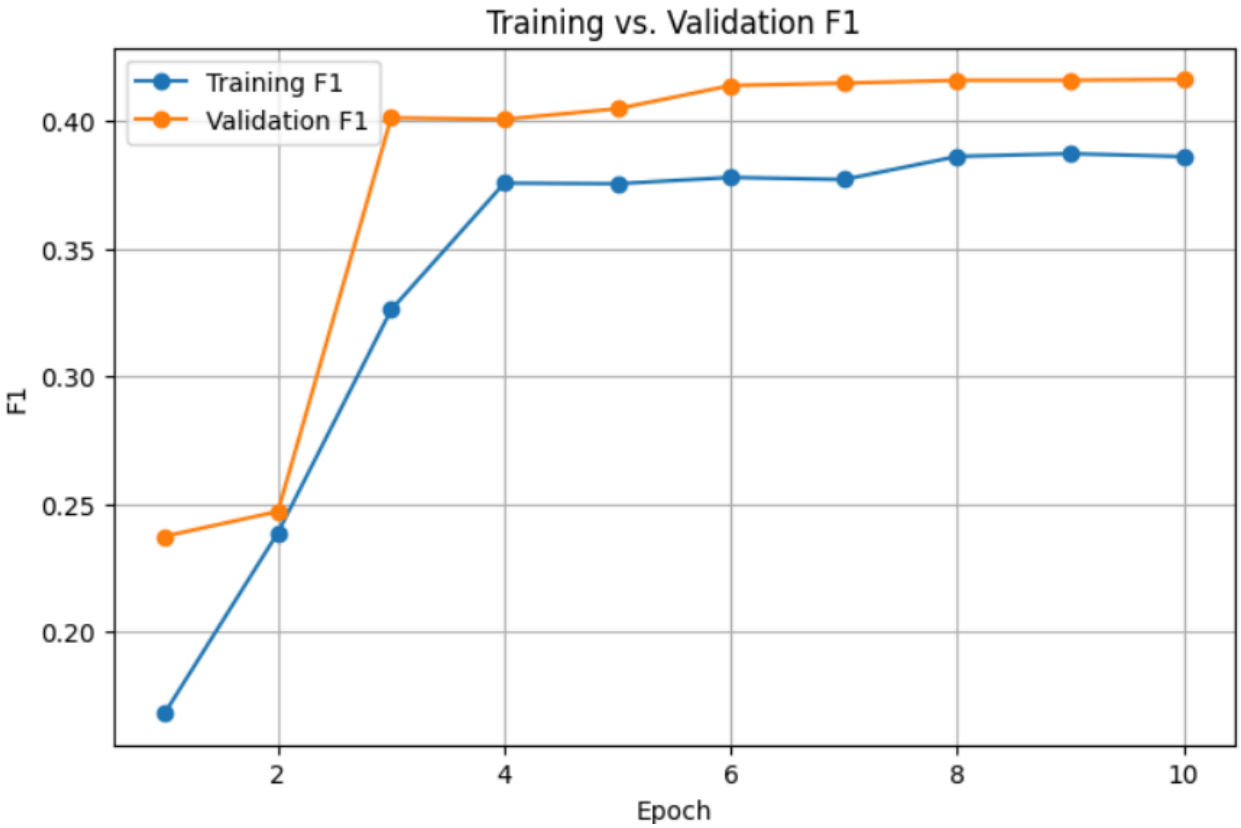
```
Loss: 0.12434762716293335
Metrics: [0.9681832194328308, 0.36756810545921326, 0.9711622595787048, 0.96146559715271]
```

F1 = `0.4869181753317442`

LSTM



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.

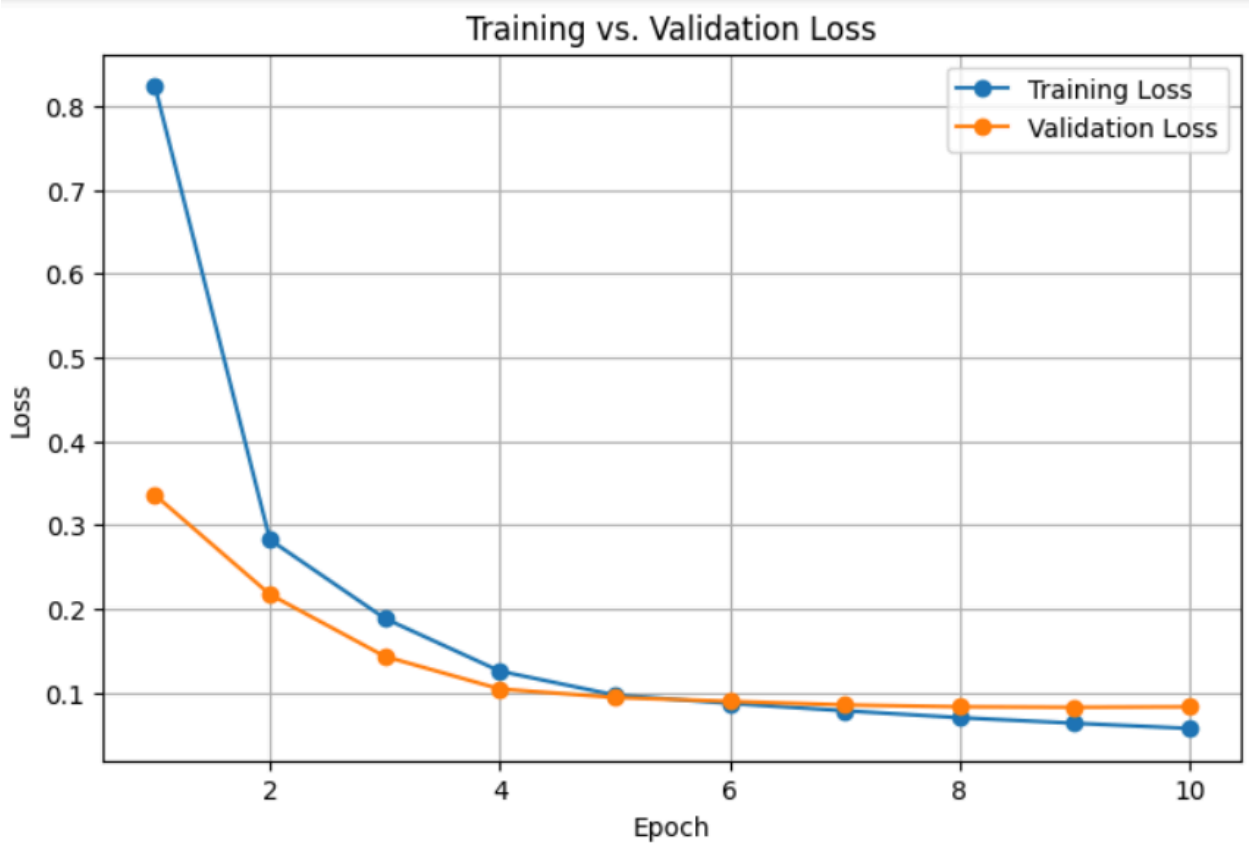


The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

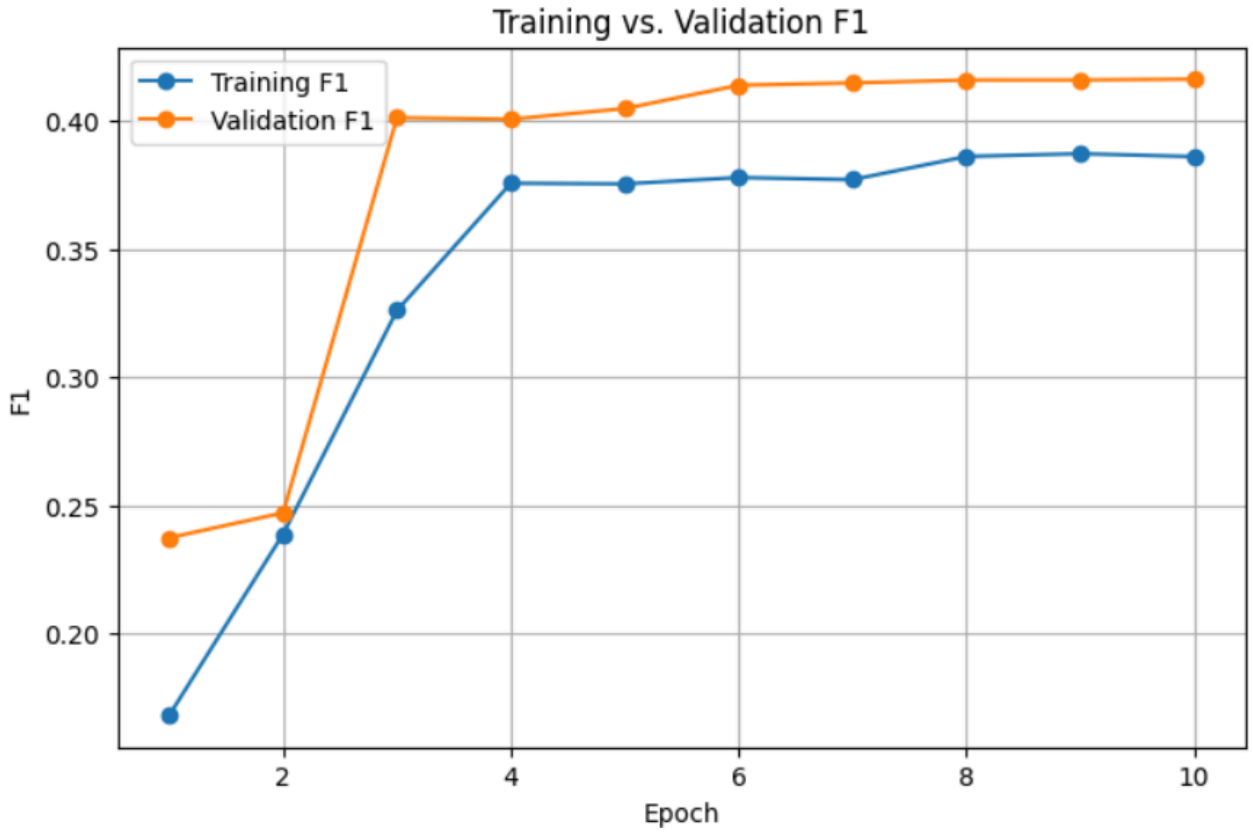
```
Loss: 0.11289666593074799
Metrics: [0.971768319606781, 0.3421818017959595, 0.9746668338775635, 0.9673296213150024]
```

F1 = `0.47194246402119094`

GRU



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



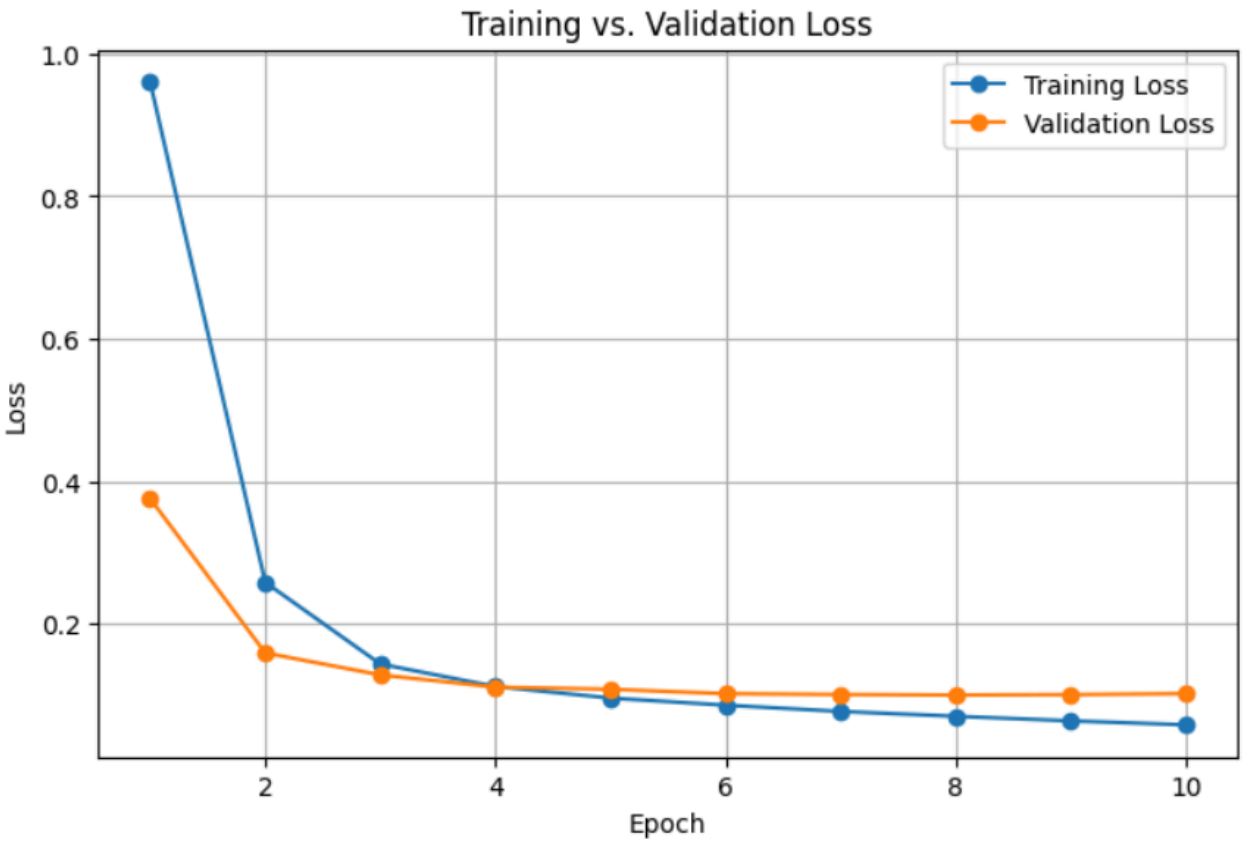
The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Loss: 0.11289666593074799
Metrics: [0.971768319606781, 0.3421818017959595, 0.9746668338775635, 0.9673296213150024]
```

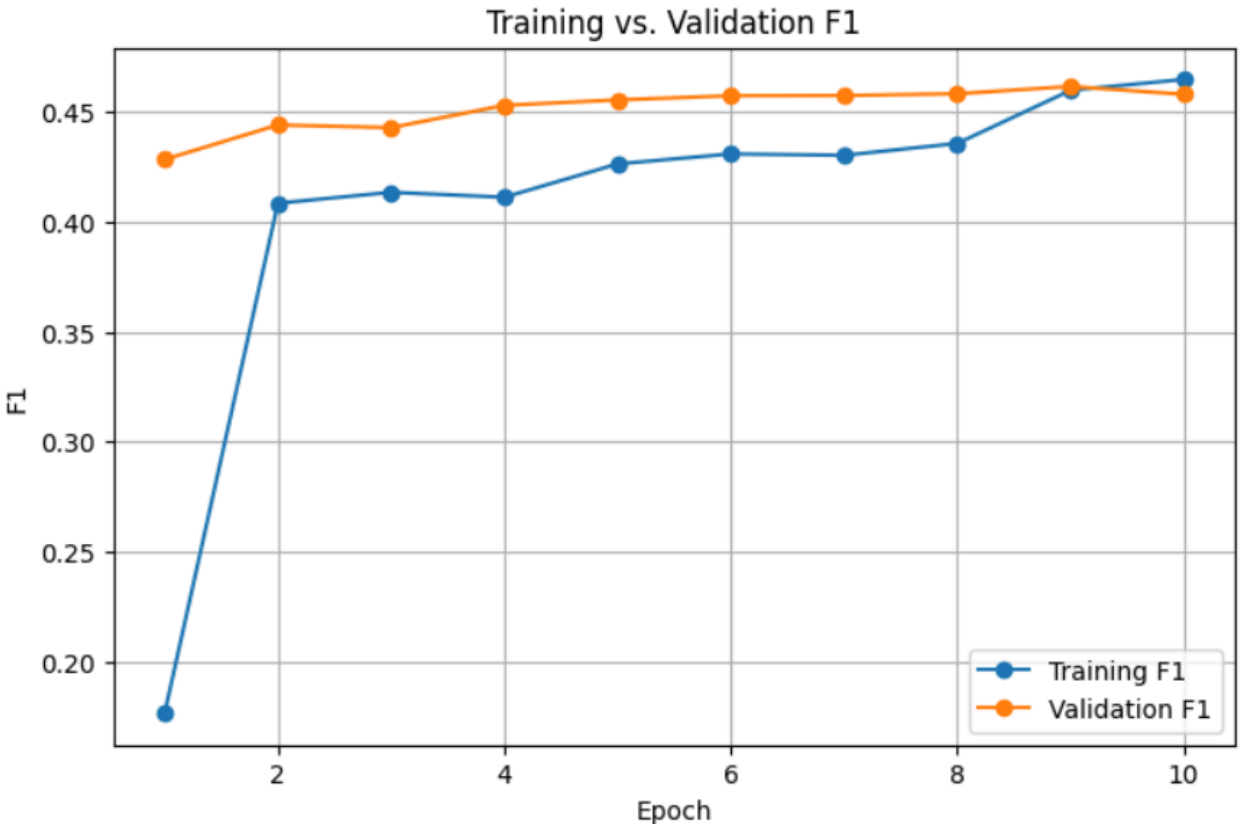
F1 = `0.4885332363193045`

GloVe

Vanilla RNN



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.

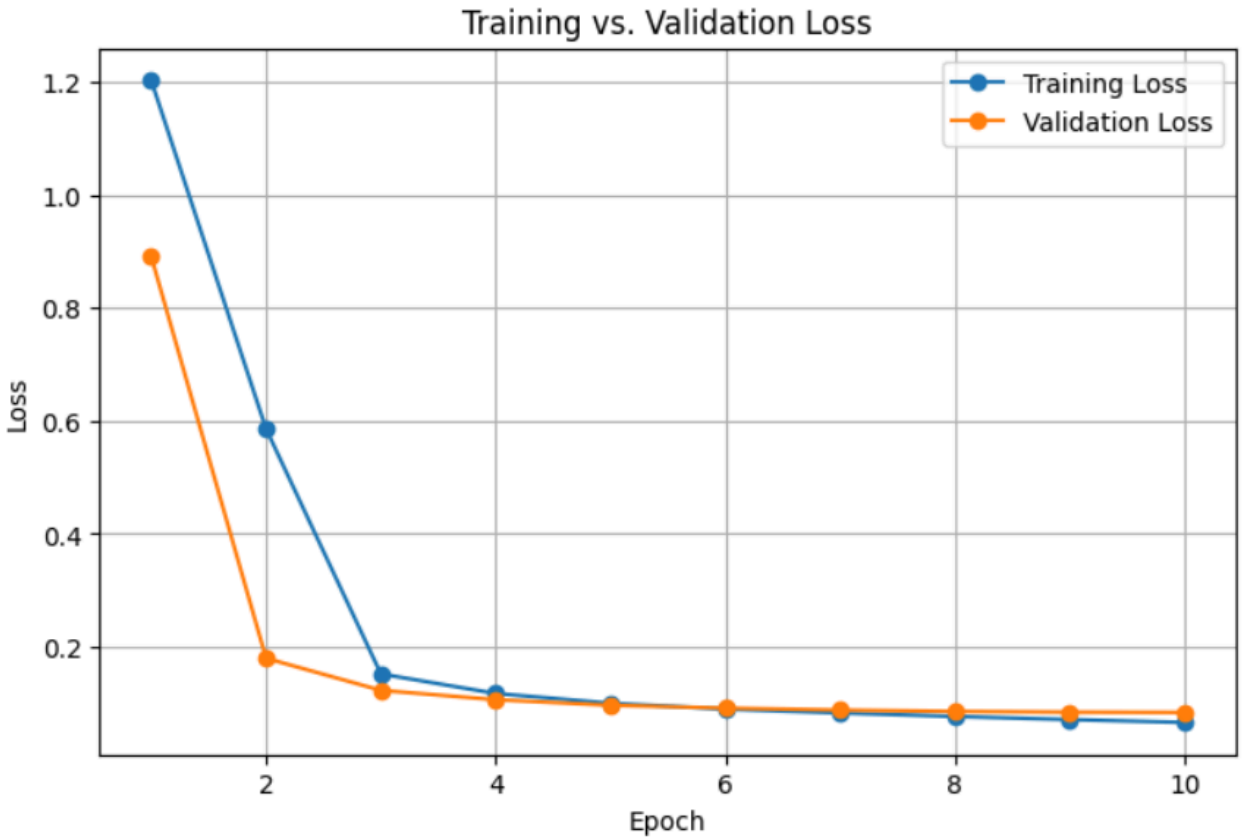


The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

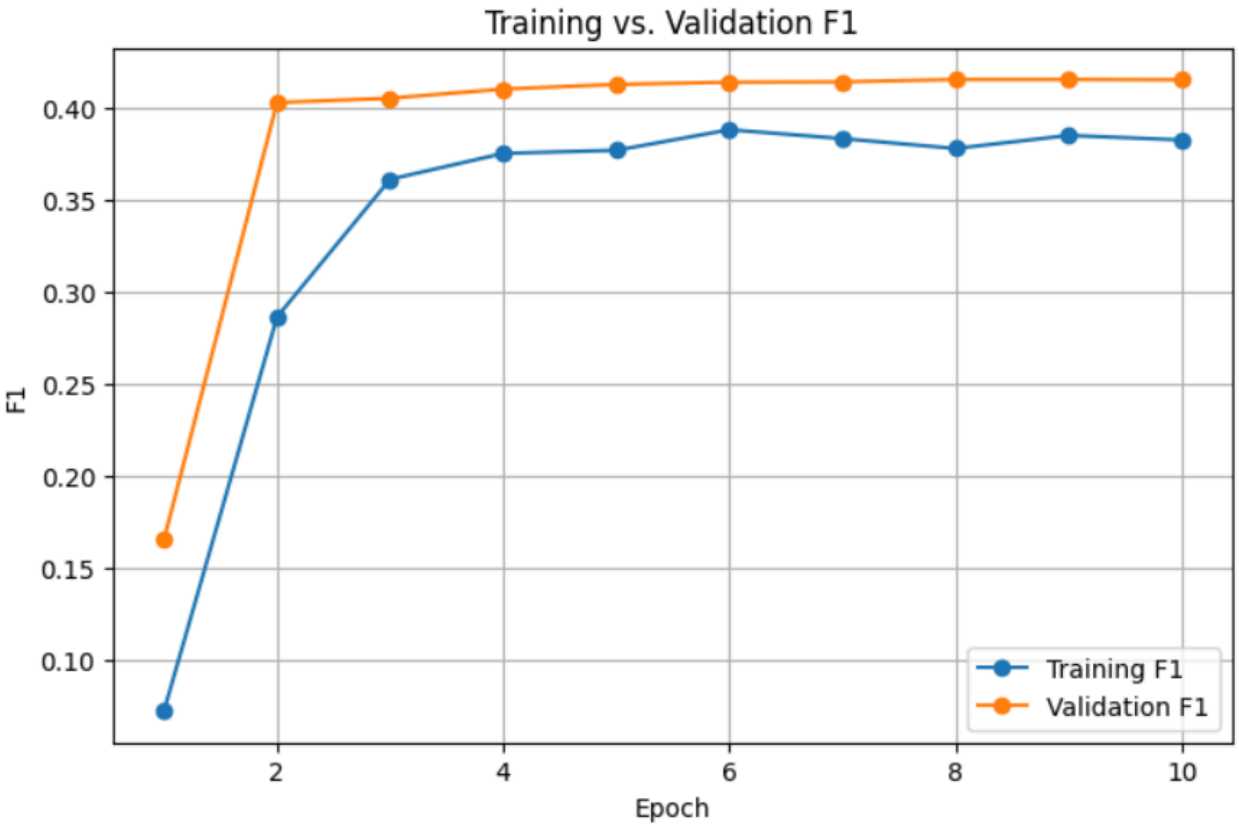
```
Loss: 0.18089483678340912
Metrics: [0.9433239698410034, 0.3489941358566284, 0.9559844732284546, 0.9364802837371826]
```

F1 = `0.47186878375164776`

LSTM



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.

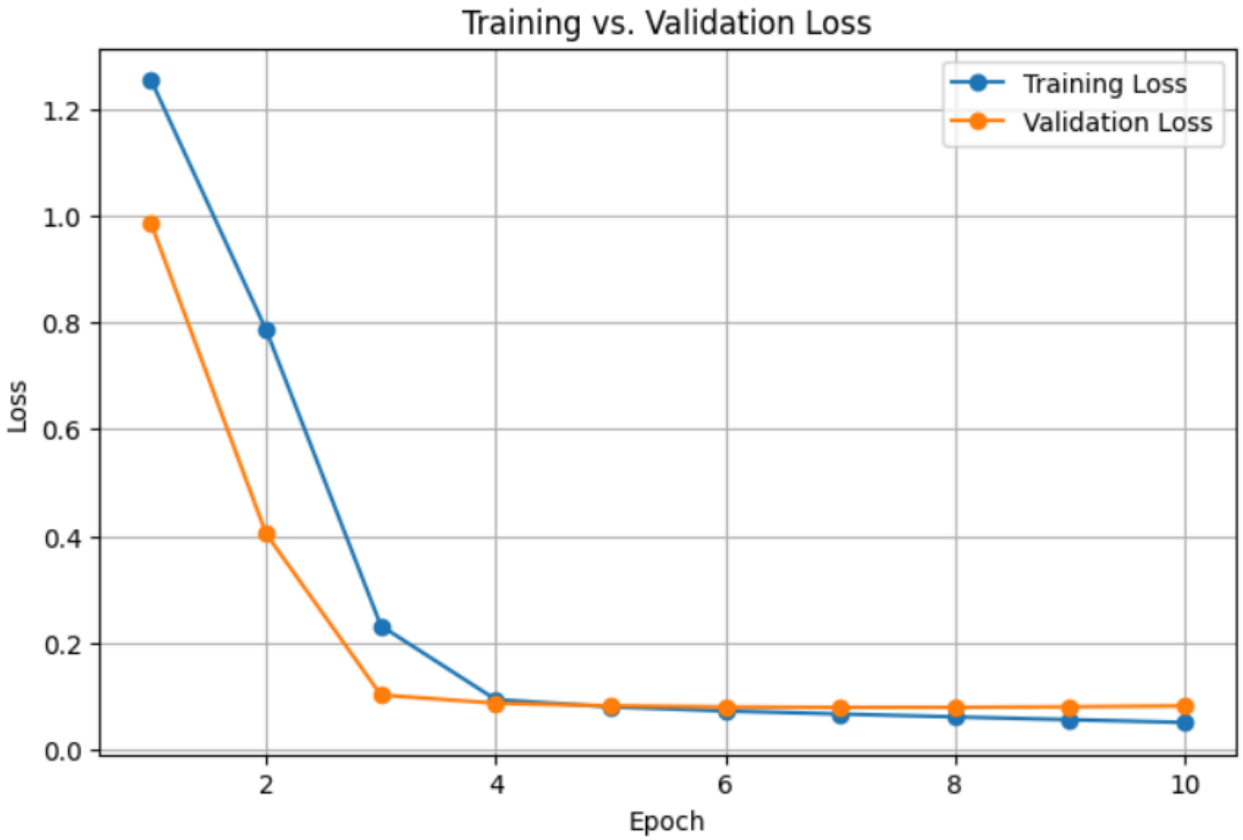


The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

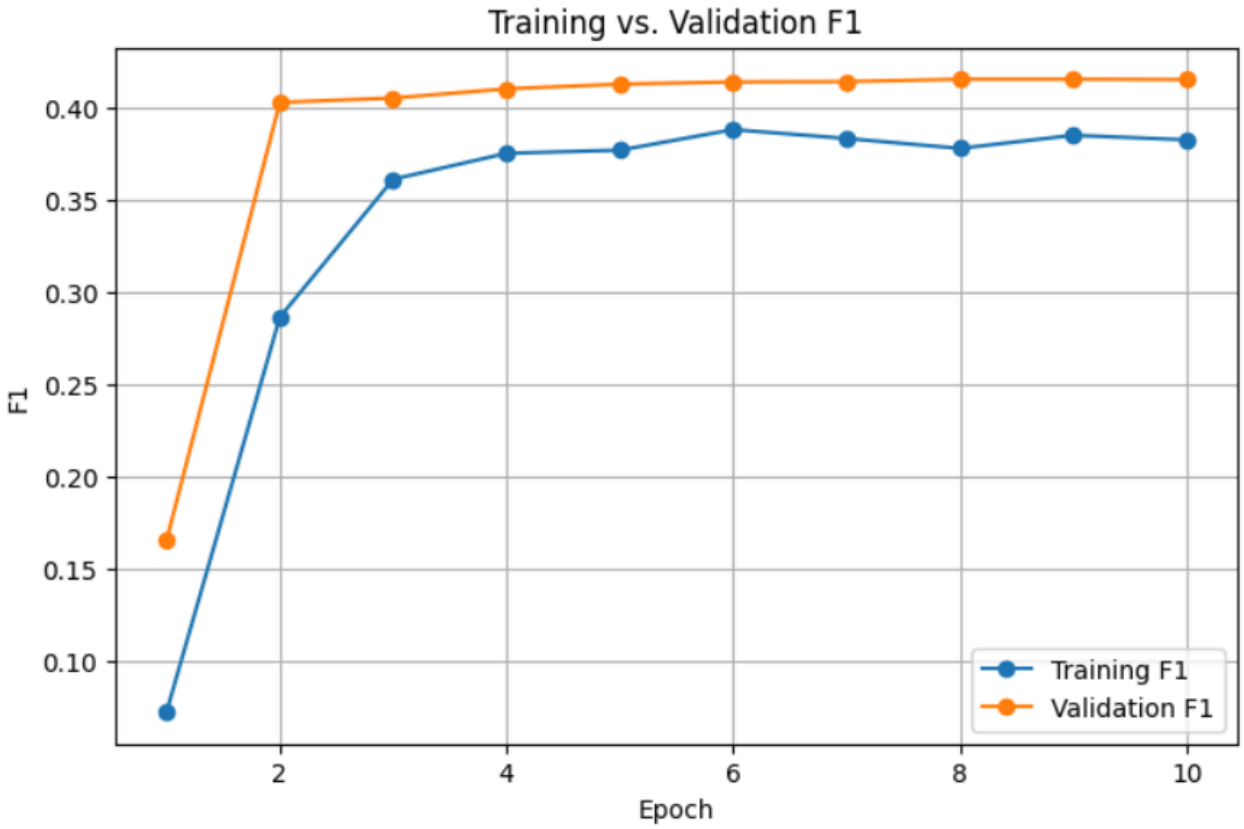
```
Loss: 0.1414303332567215
Metrics: [0.9557926654815674, 0.3272154629230499, 0.9658235311508179, 0.9504544734954834]
```

F1 = `0.46707767936787 69`

GRU



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



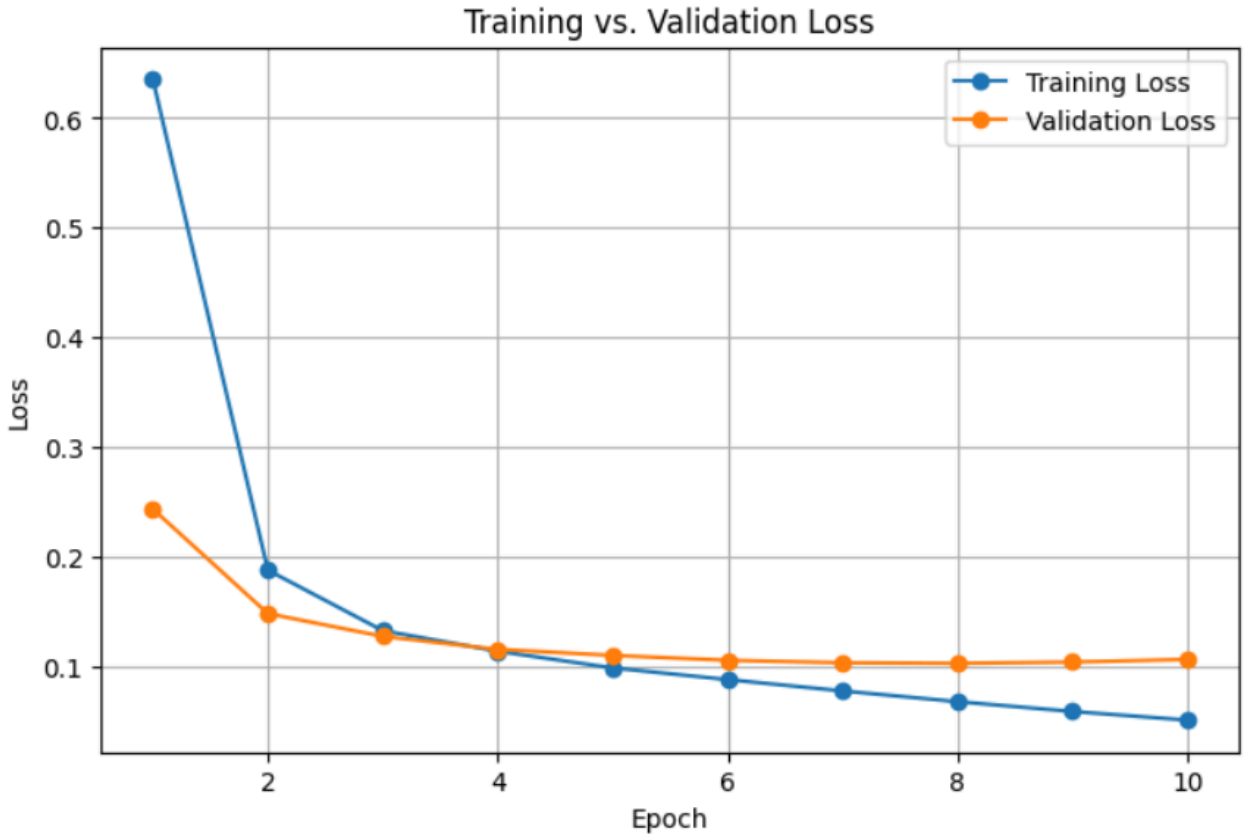
The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Loss: 0.1414303332567215
Metrics: [0.9557926654815674, 0.3272154629230499, 0.9658235311508179, 0.9504544734954834]
```

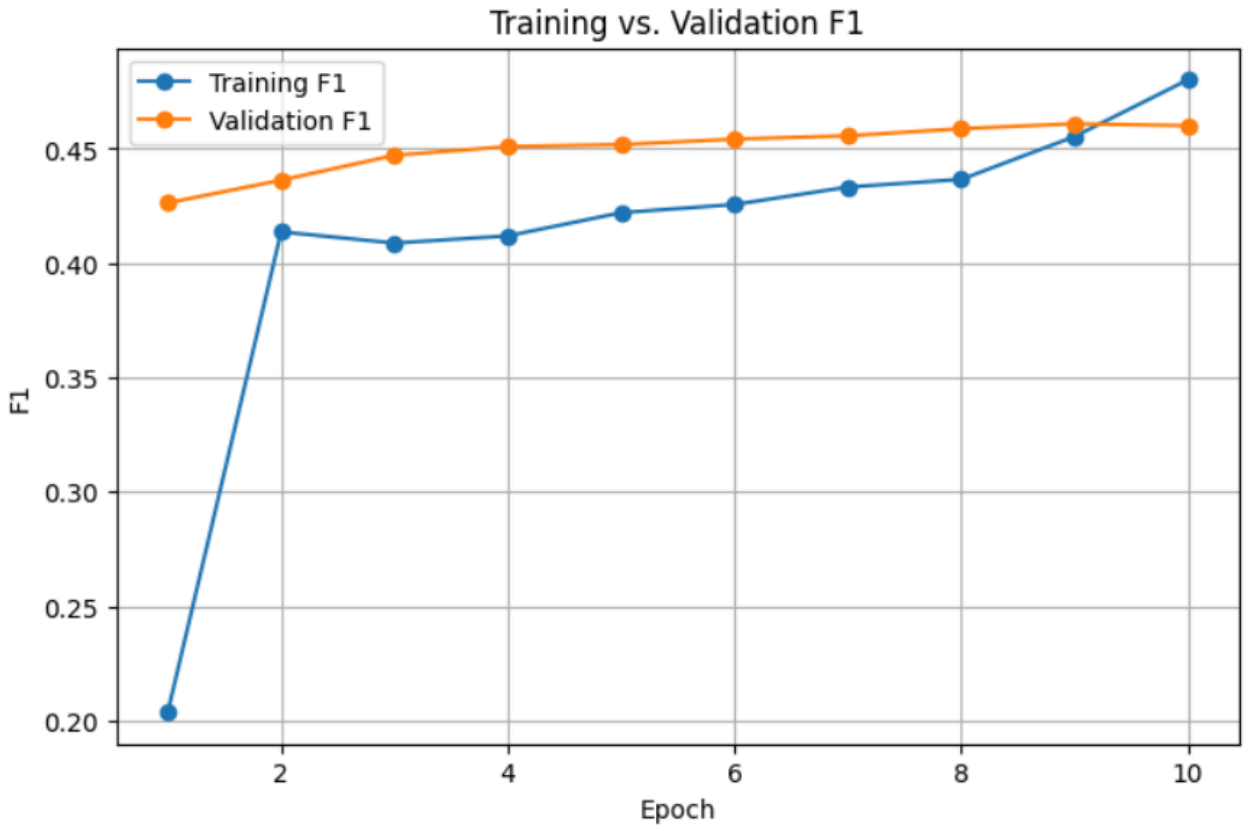
F1 = `0.4749094833689165`

FastText

Vanilla RNN



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.

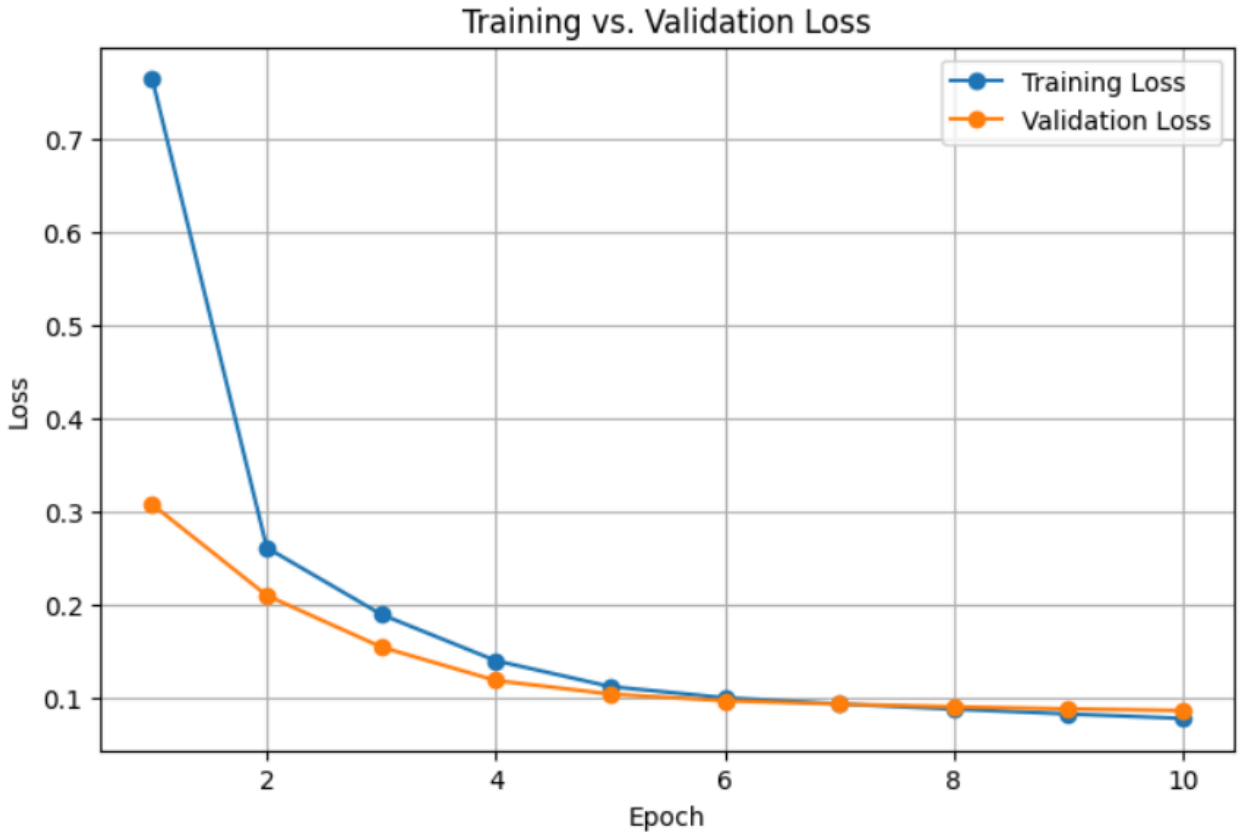


The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

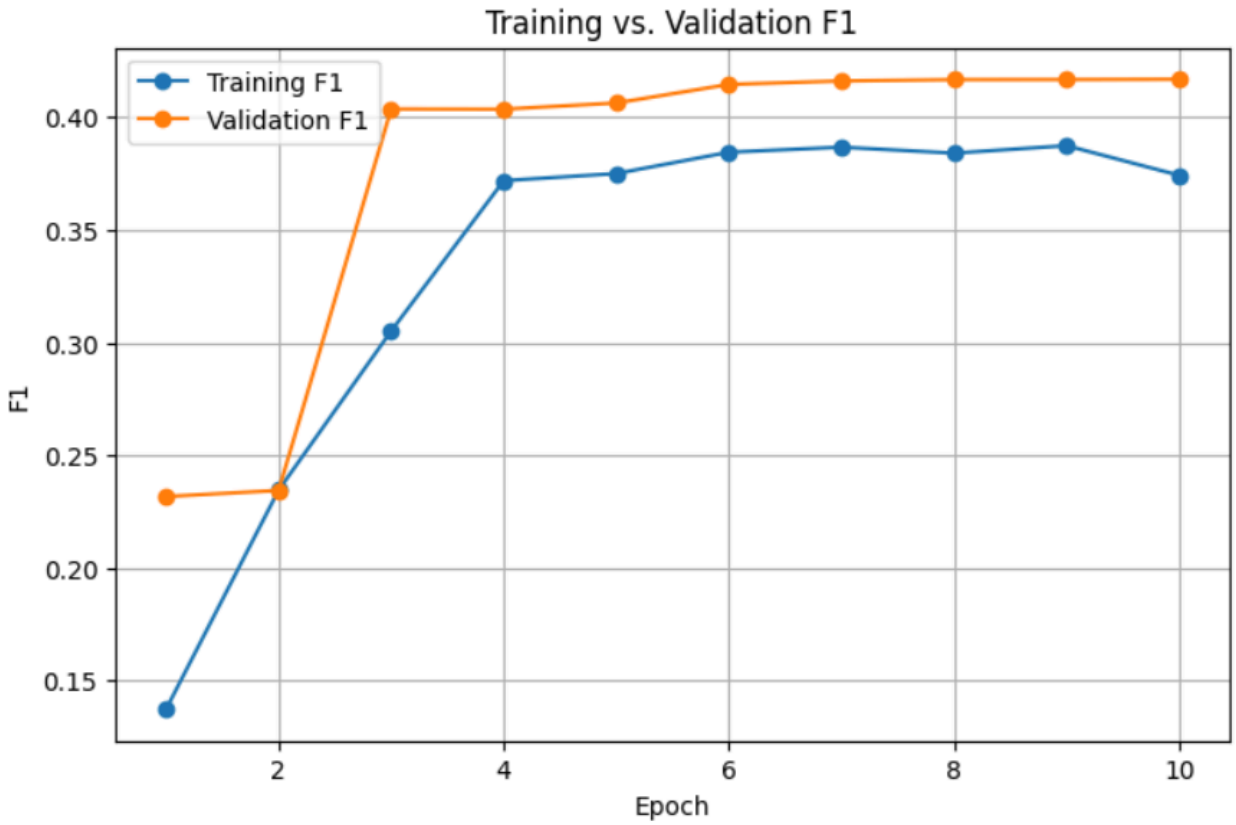
```
Loss: 0.13091148436069489
Metrics: [0.9618120789527893, 0.3651304543018341, 0.9679449200630188, 0.9545453786849976]
```

F1 = `0.4912289392945629`

LSTM



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.

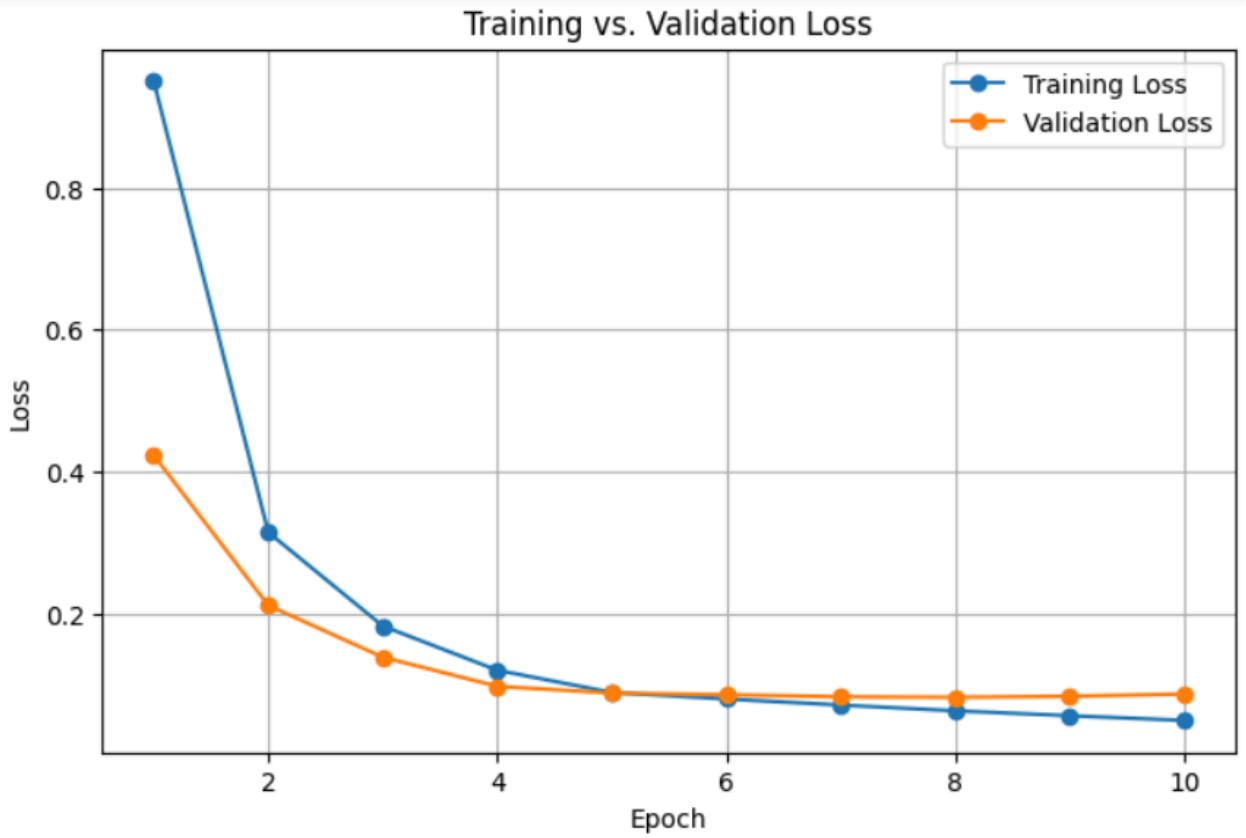


The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

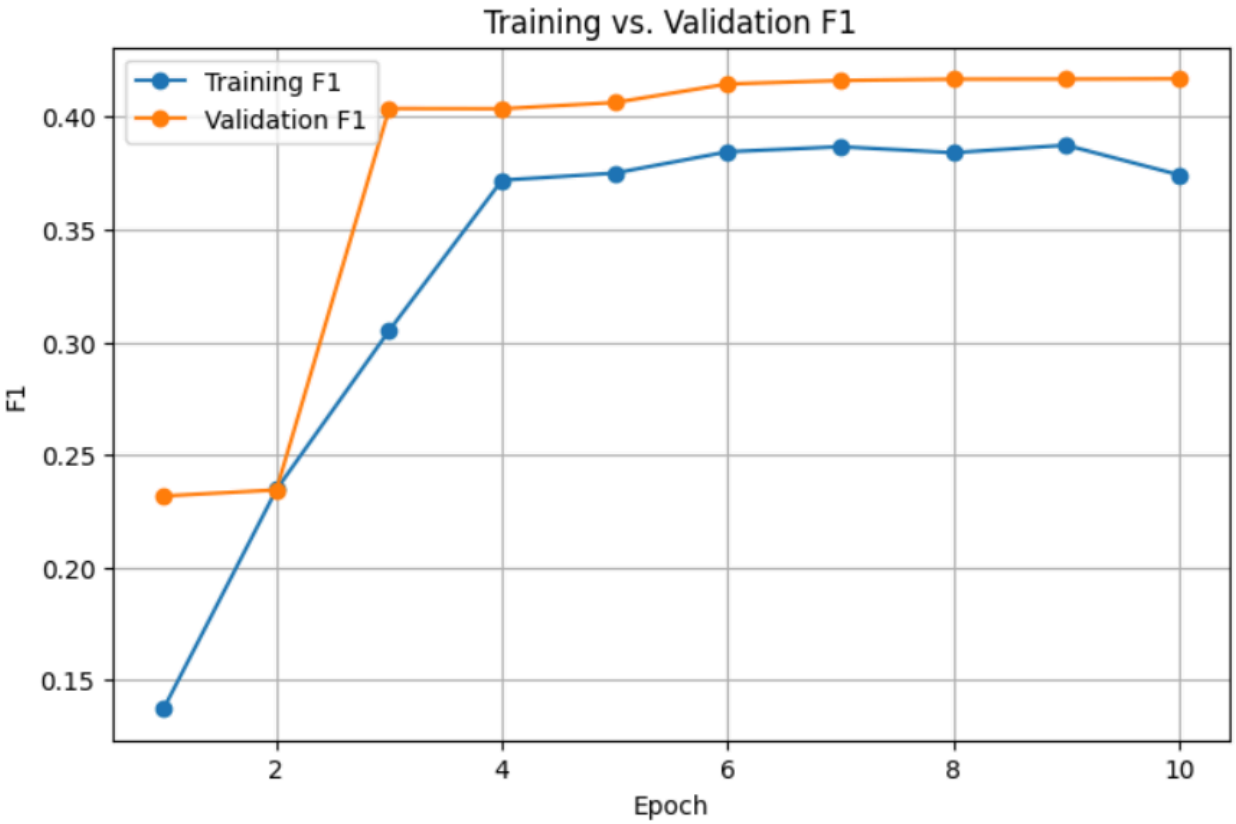
```
Loss: 0.10233476758003235
Metrics: [0.9702134132385254, 0.33980175852775574, 0.9750735759735107, 0.964545488357544]
```

F1 = `0.4738603646472841`

GRU



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



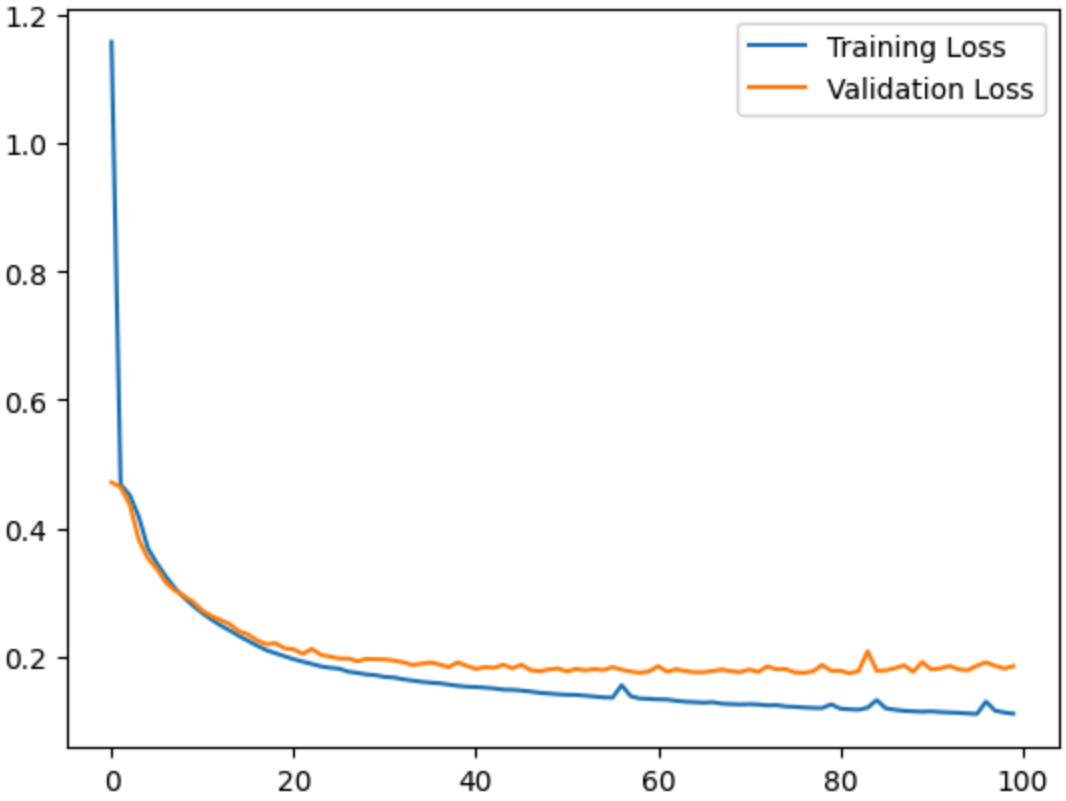
The macro F1 score of the training set (blue line) fluctuates slightly but shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Loss: 0.10233476758003235
Metrics: [0.9702134132385254, 0.33980175852775574, 0.9750735759735107, 0.964545488357544]
```

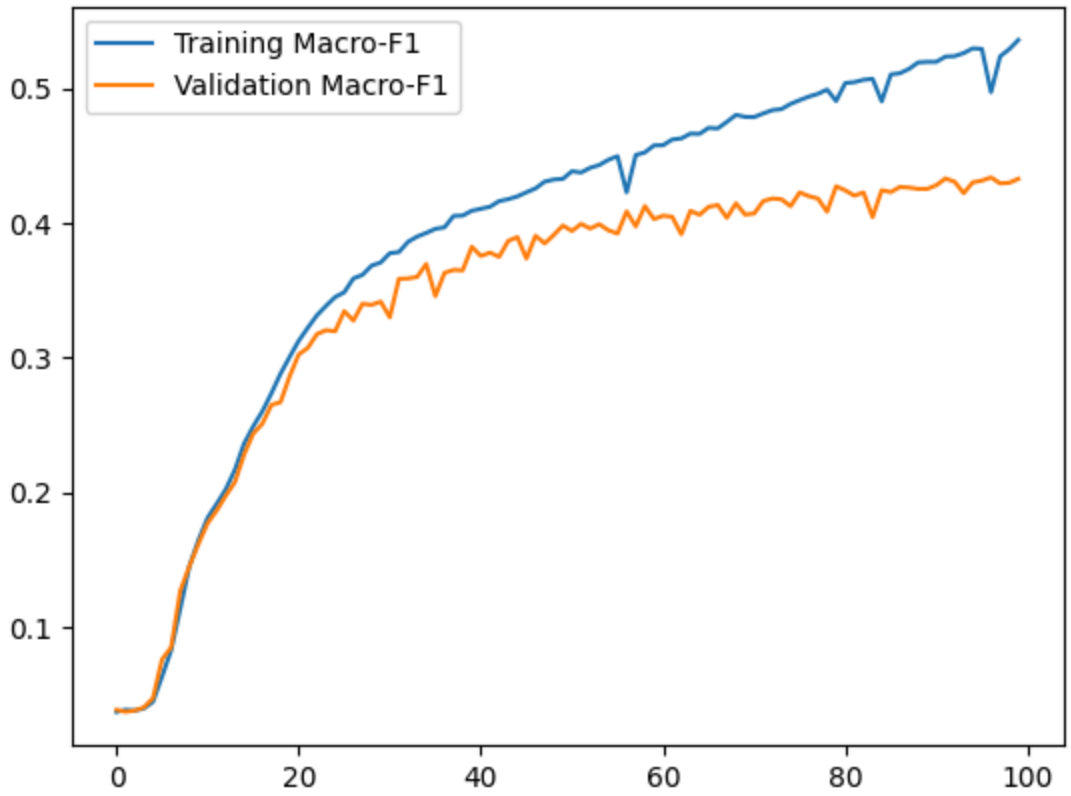
F1 = `0.4905321483869017`

Task 1:

Word2Vect
Vanilla RNN



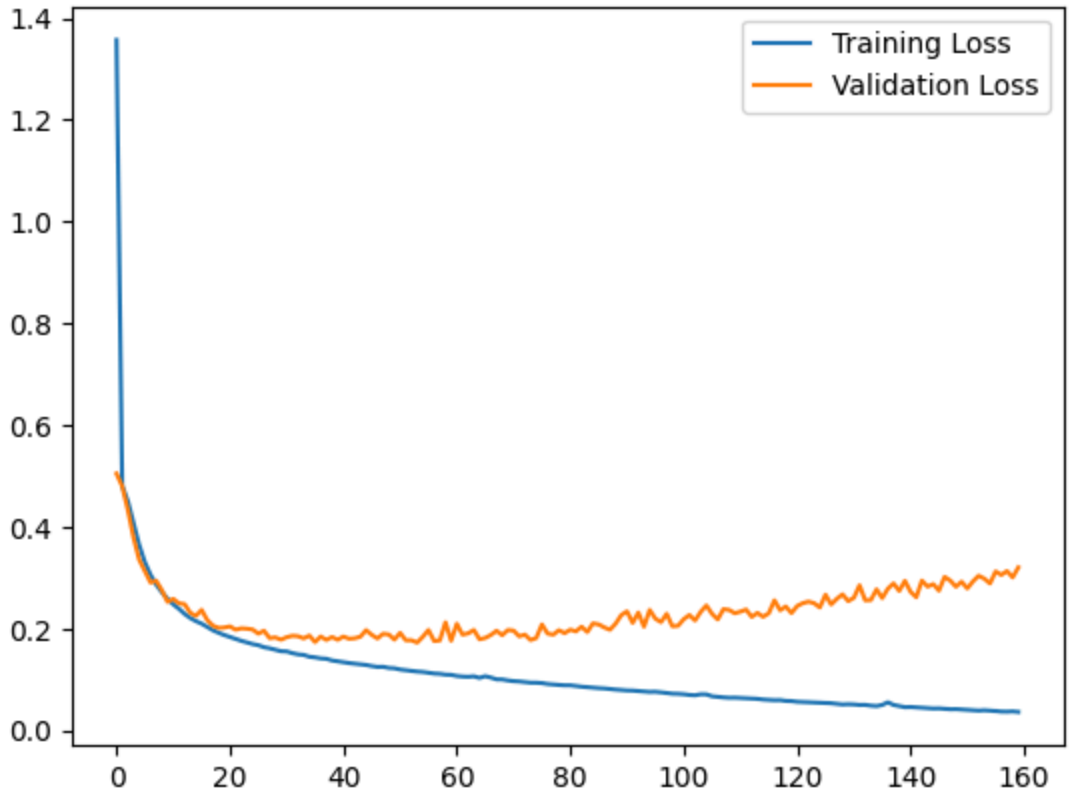
The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



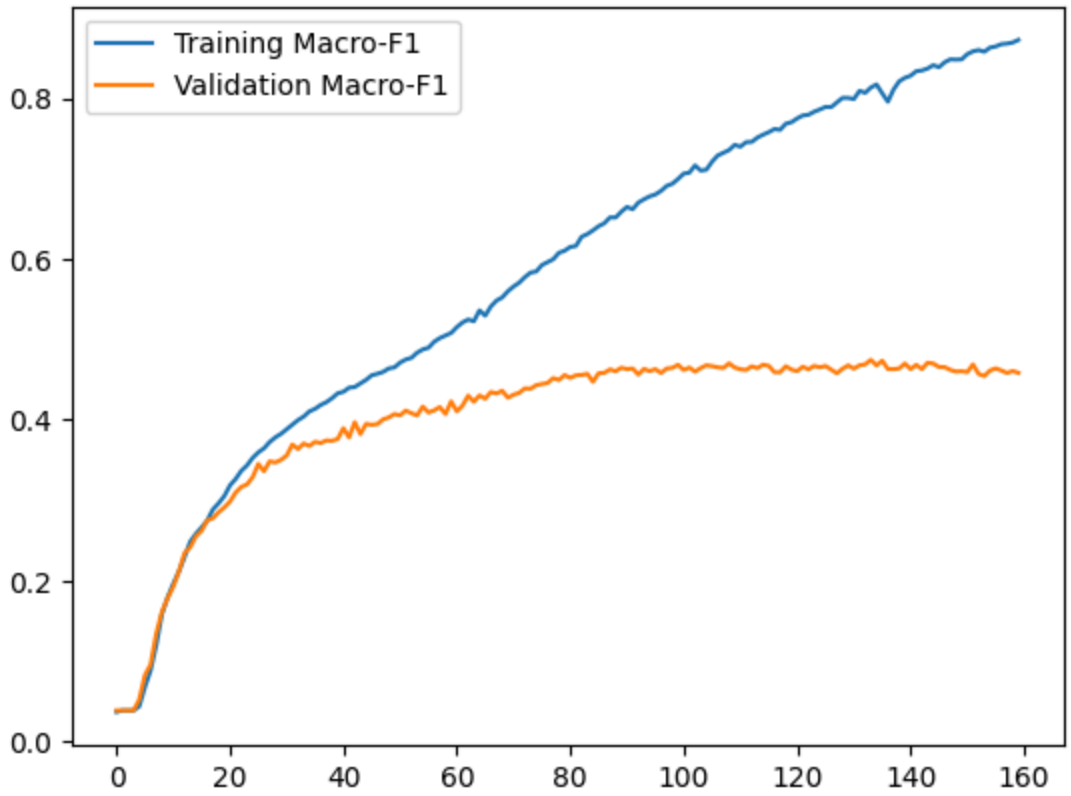
The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

Test Loss: 0.1963, Test Macro-F1: 0.4264

LSTM



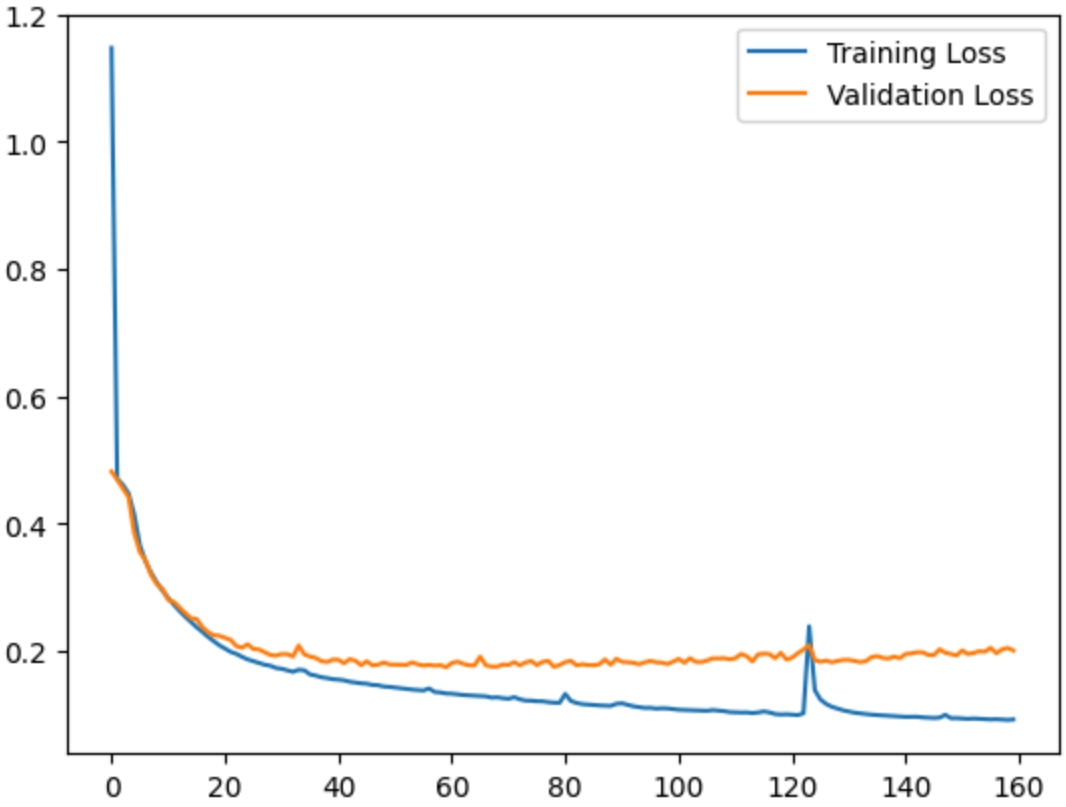
The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



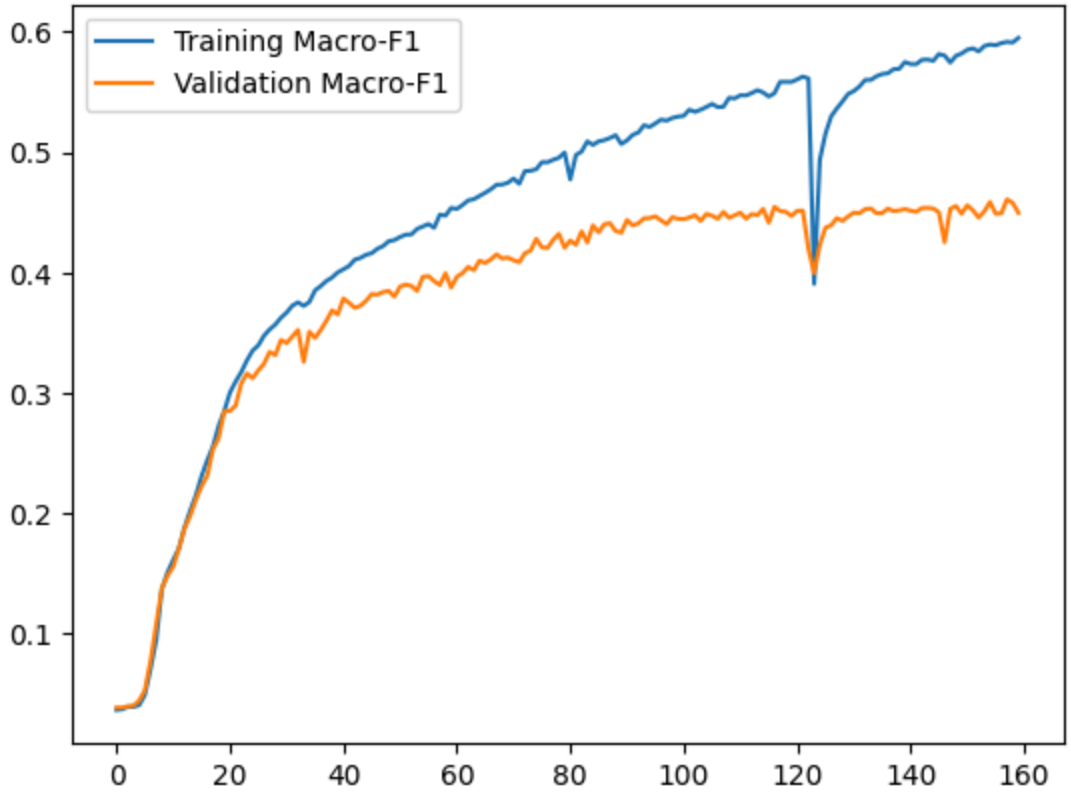
The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Test Loss: 0.3798, Test Macro-F1: 0.4197
```

GRU



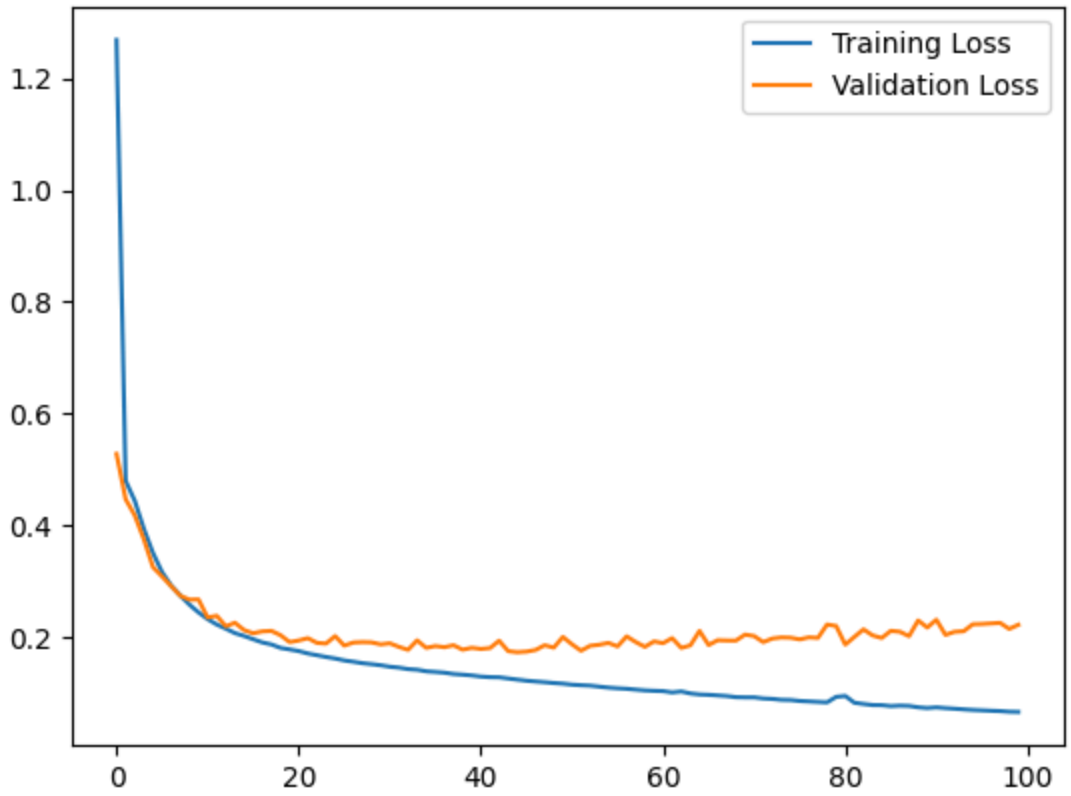
The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



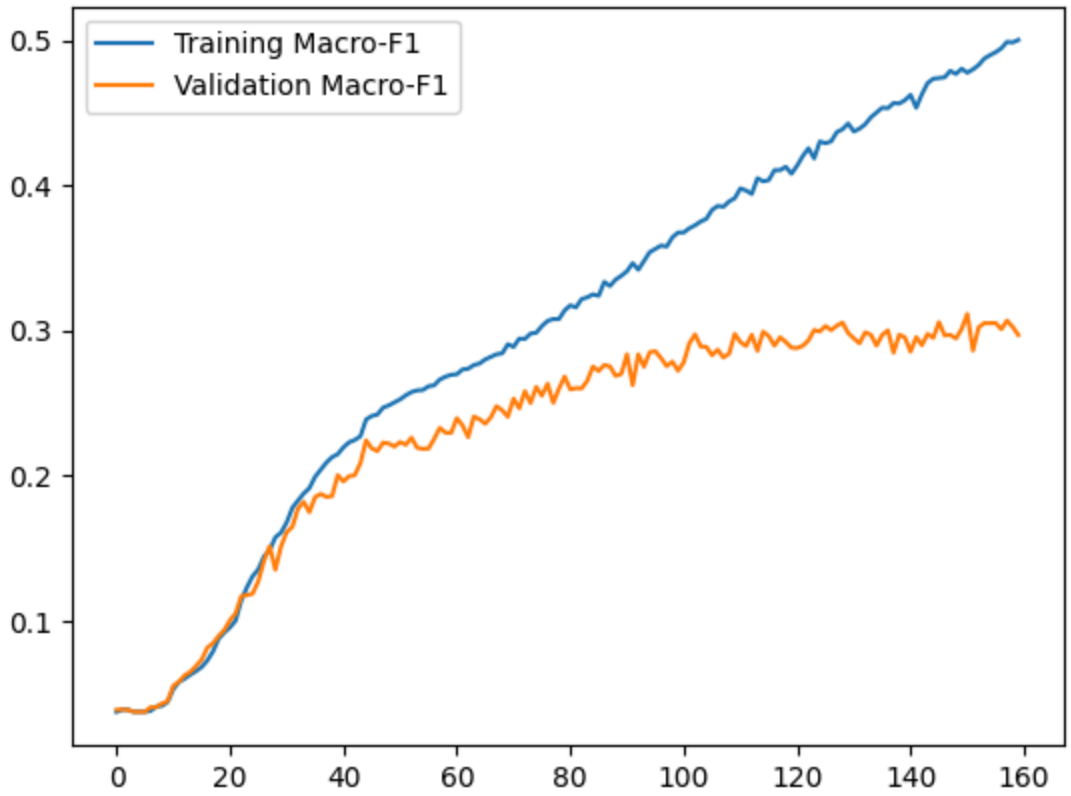
The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Test Loss: 0.2264, Test Macro-F1: 0.4264
```

GloVe
Vanilla RNN



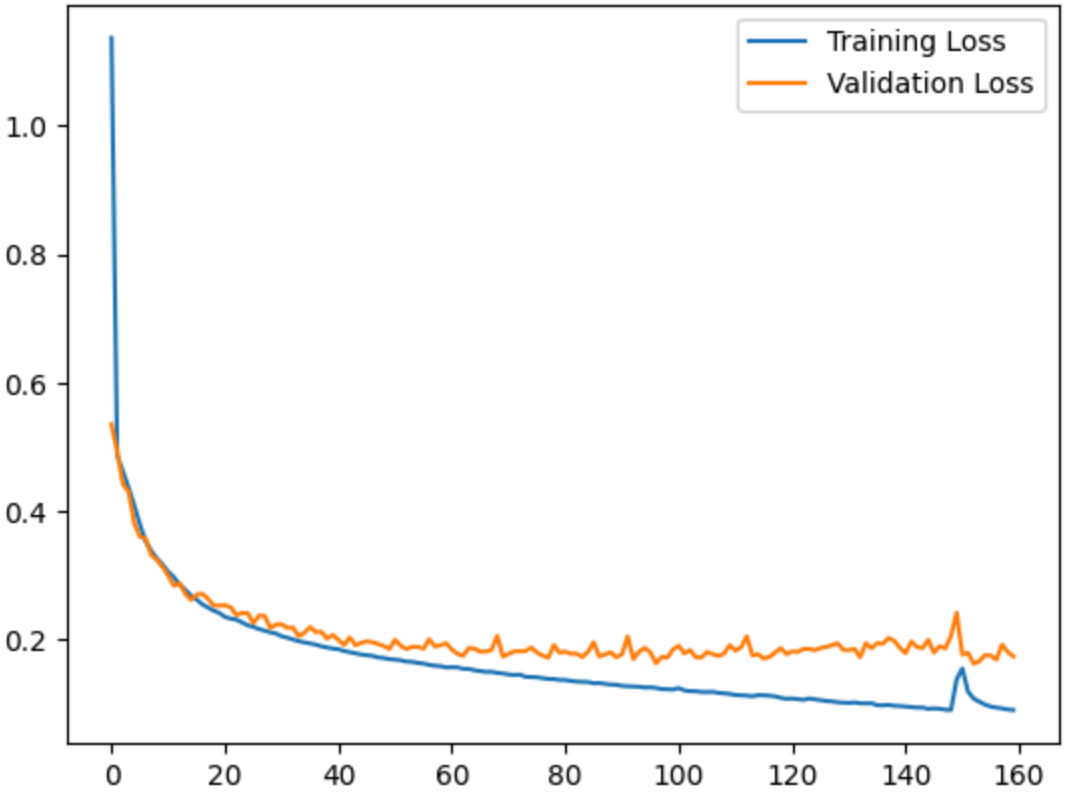
The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



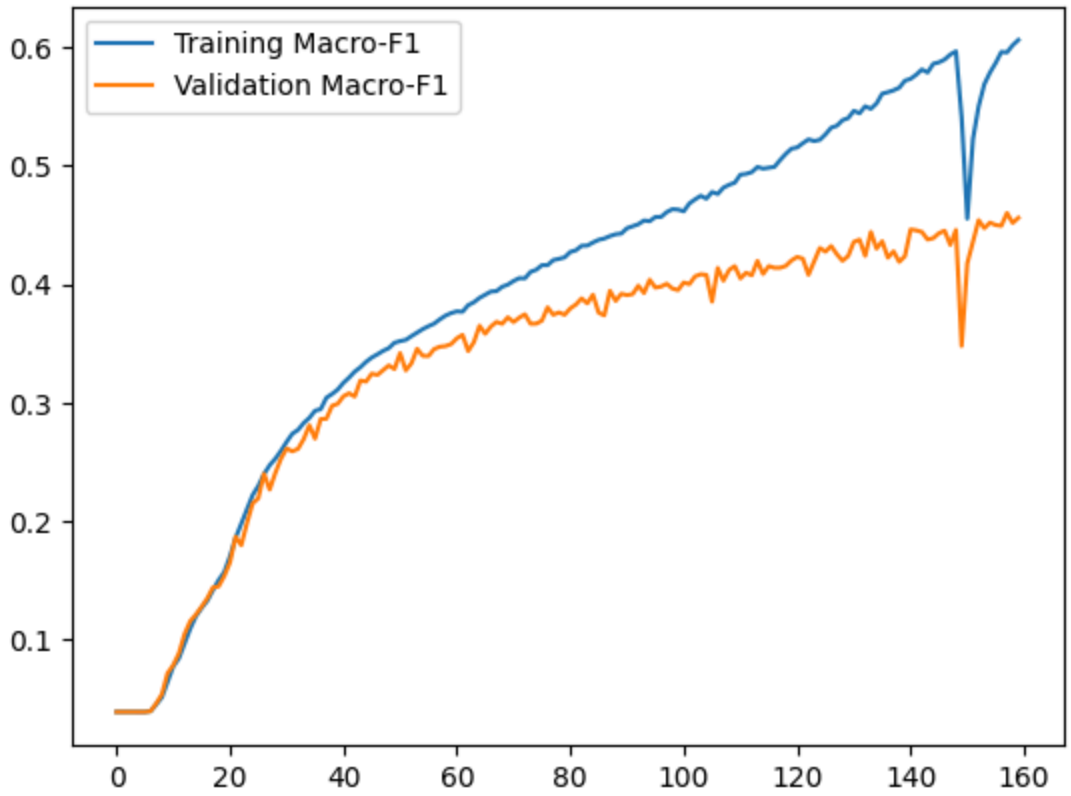
The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Test Loss: 0.3393, Test Macro-F1: 0.2949
```

LSTM



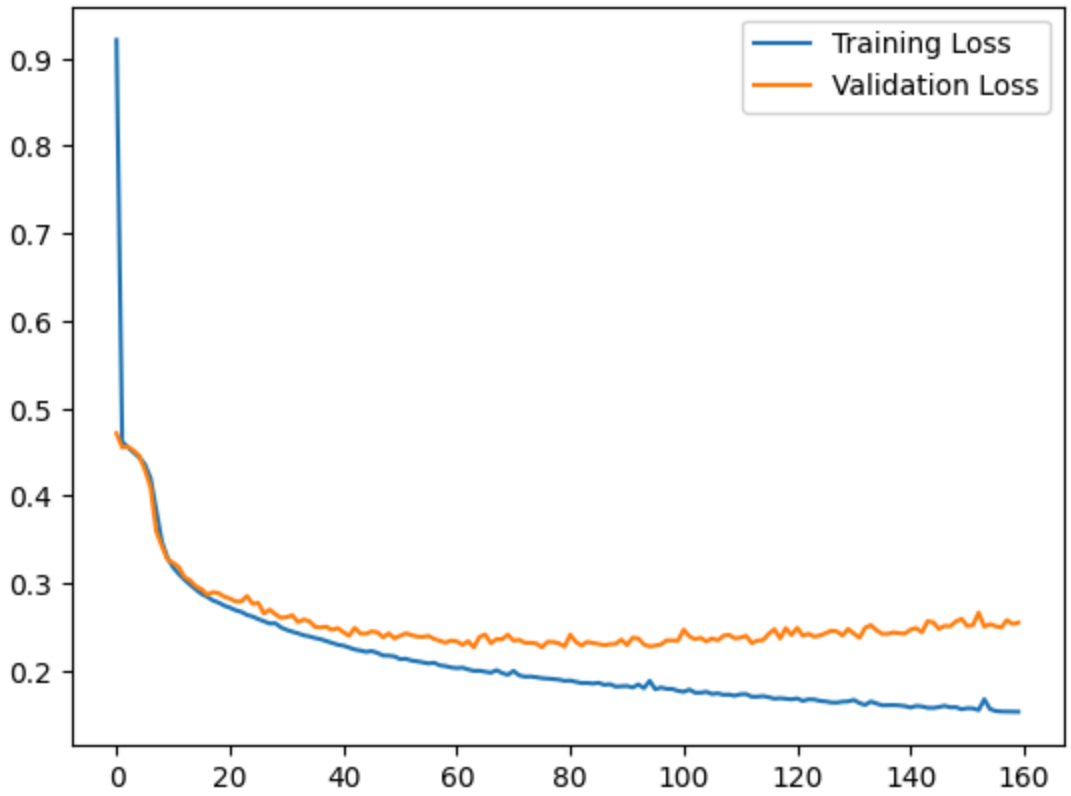
The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



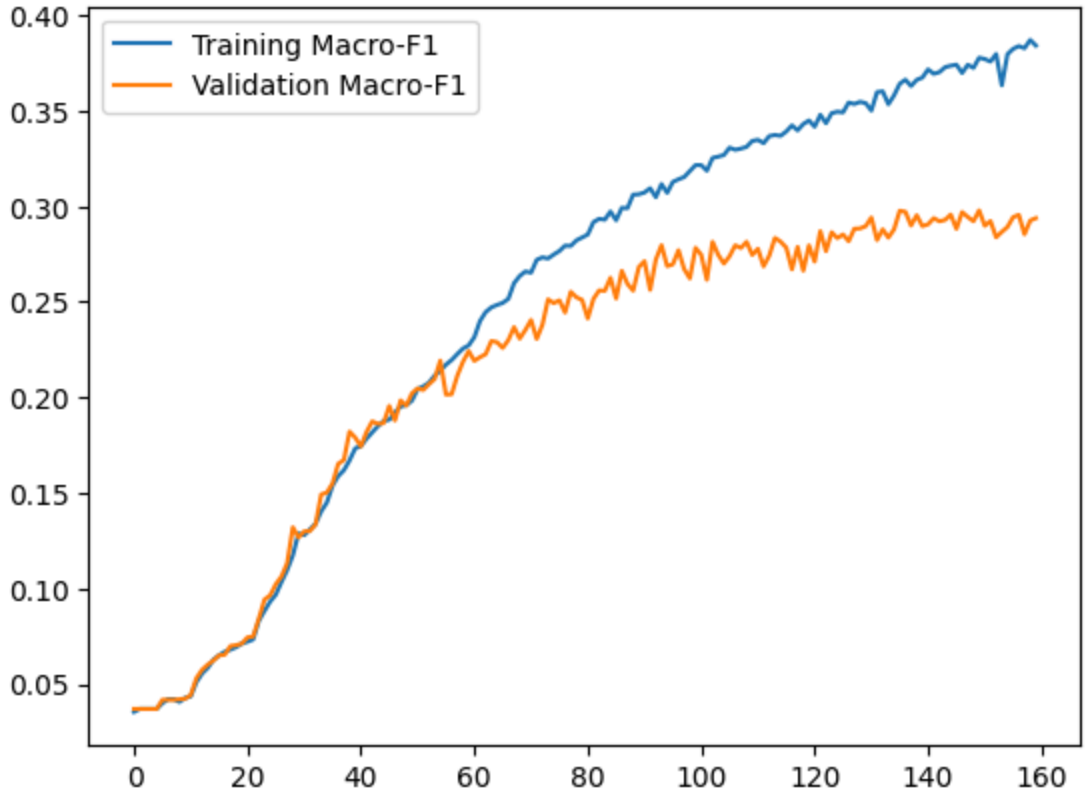
The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Test Loss: 0.1944, Test Macro-F1: 0.4642
```

GRU



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.

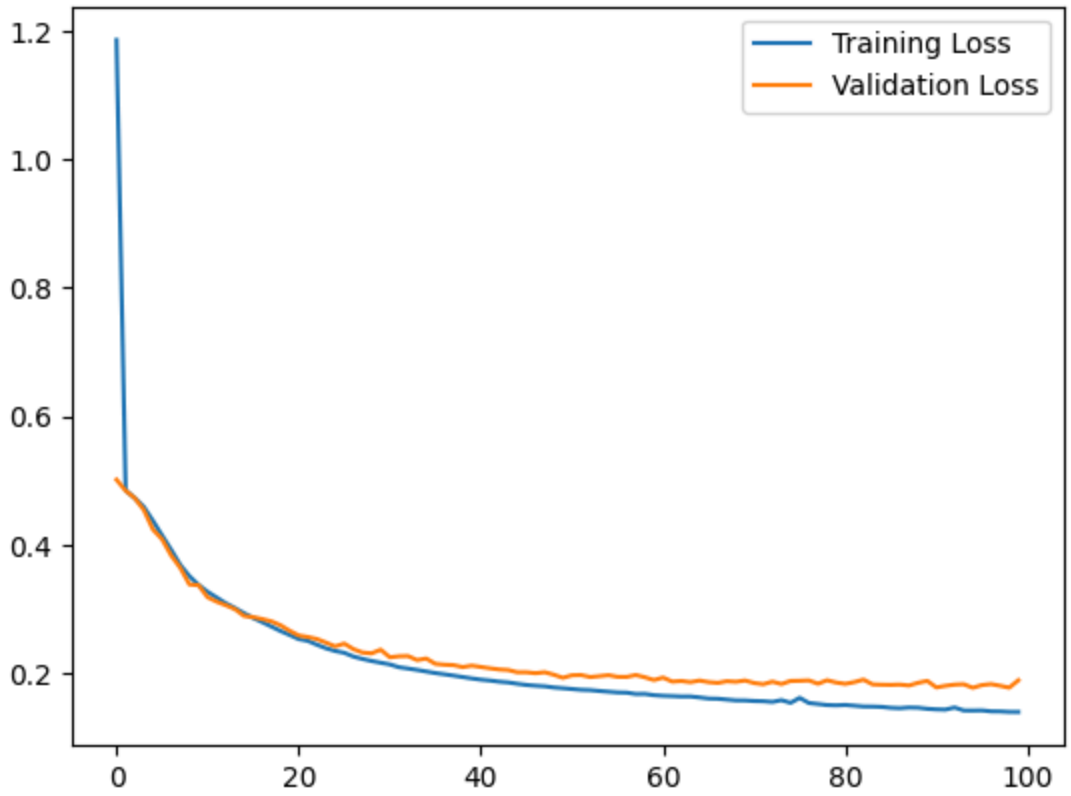


The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

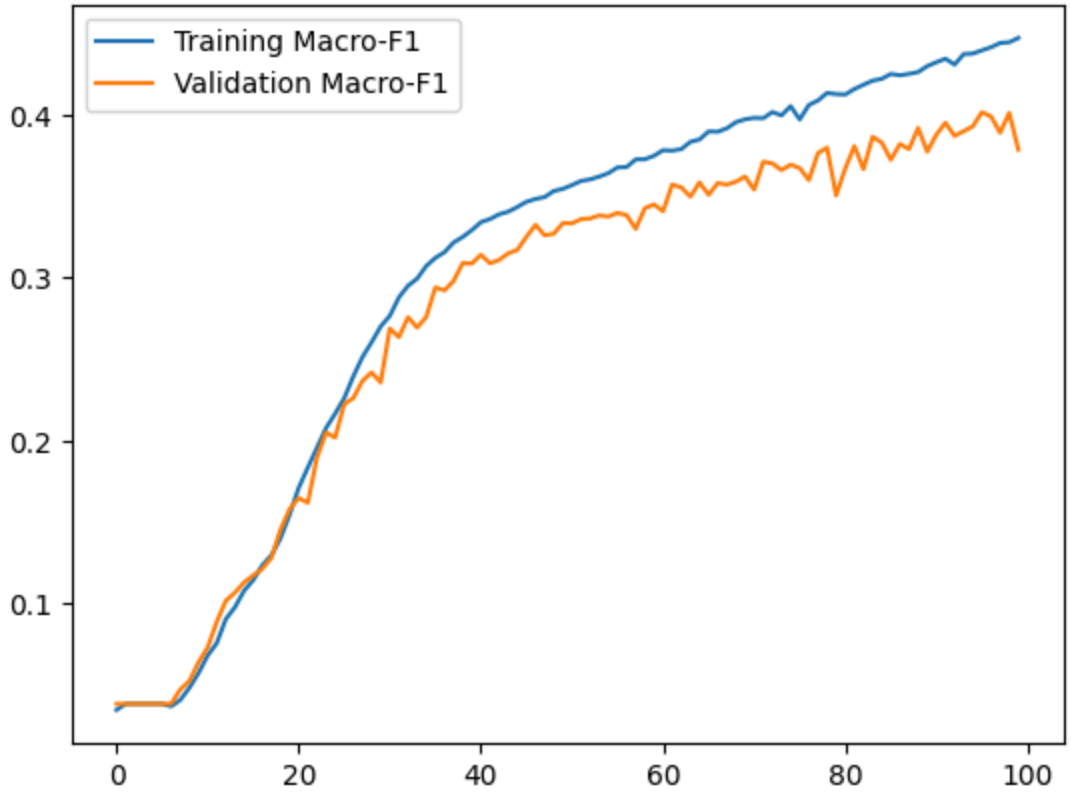
```
Test Loss: 0.2714, Test Macro-F1: 0.2813
```

Fasttext

Vanilla RNN



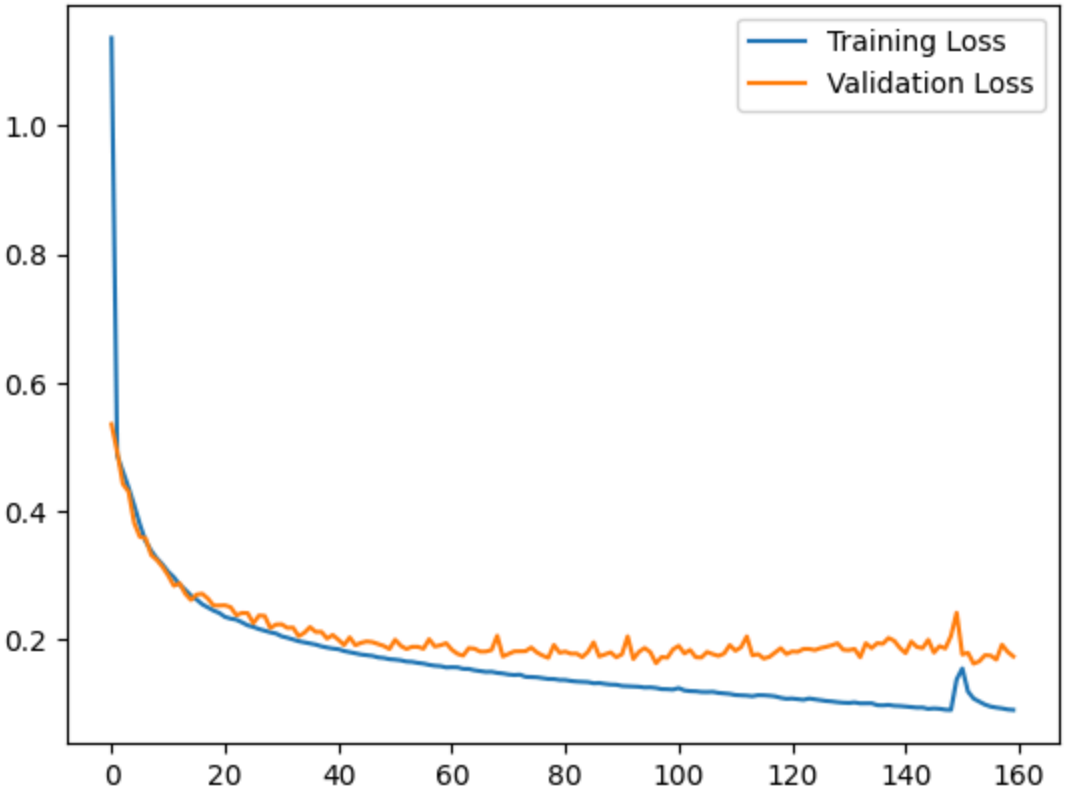
The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



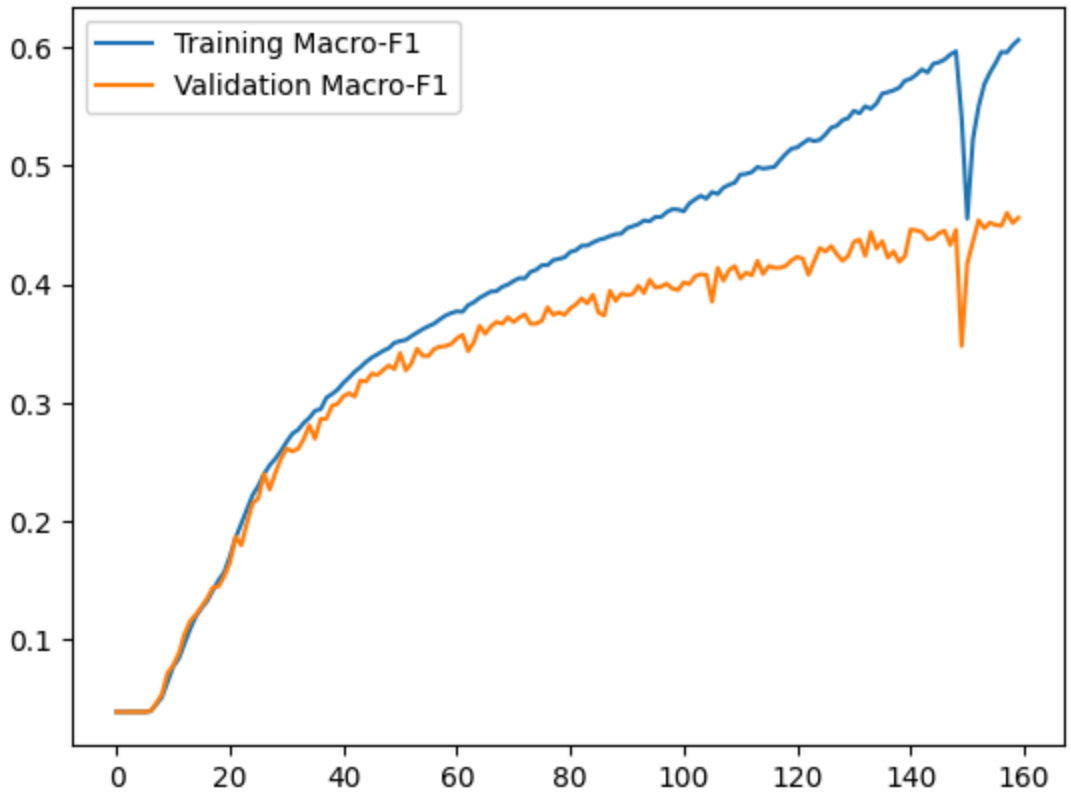
The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Test Loss: 0.2028, Test Macro-F1: 0.3813
```

LSTM



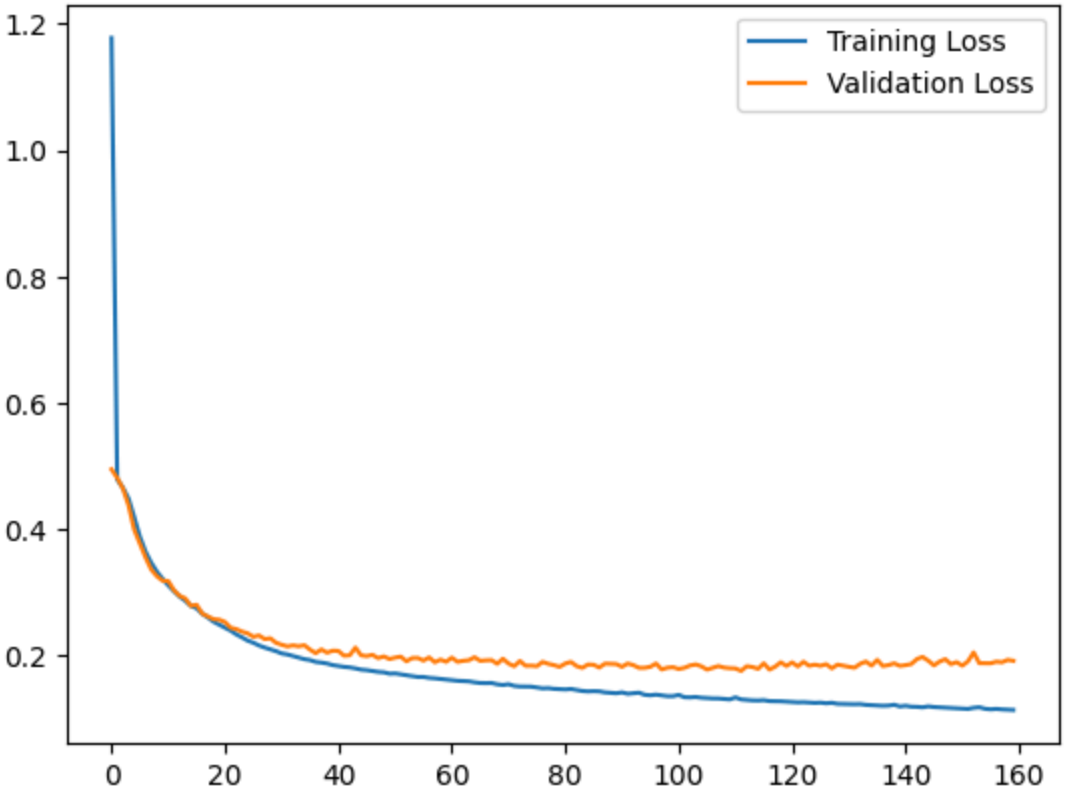
The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



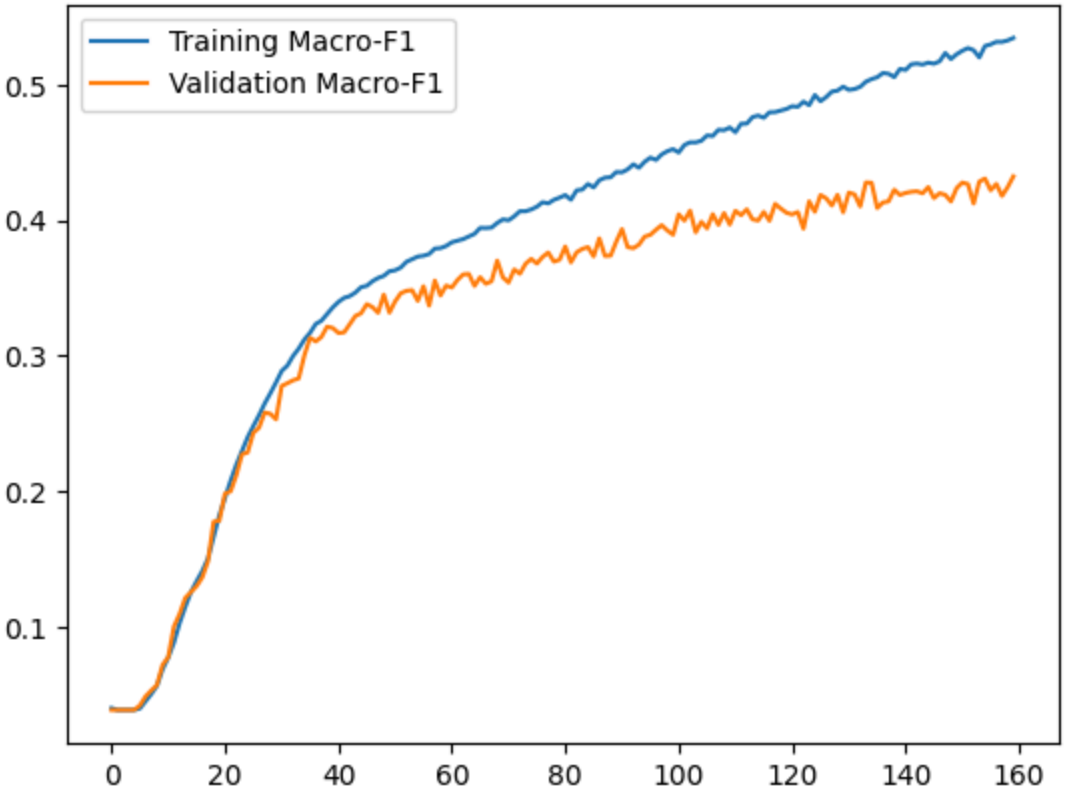
The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Test Loss: 0.1944, Test Macro-F1: 0.4642
```

GRU



The training loss (blue line) significantly decreases as the number of epochs increases, suggesting the LSTM model is learning from the training data. However, the validation loss (orange line) also decreases and come near the training loss for certain number of epochs, indicating the model have been learning good to the training data.



The macro F1 score of the training set (blue line) shows a general increasing trend as the number of epochs increases. Along with it, macro F1 for validation sets is also increasing. This suggests that the model is slowly learning and improving its performance on the training data.

```
Test Loss: 0.2107, Test Macro-F1: 0.4146
```

BiLSTM
NER Dataset
W2V:
The F1 score is 23.95%, the accuracy is 84.51%
FasttextL
F1 score is 10.28% and the accuracy is 85.28%
Glove:
The F1 score is 19.87% and the accuracy is 87.7%

These are all ran on 1epoch because it took too much time for it to run

Results for Task 2 in Part 2

Model	Word Embedding	Loss	Accuracy	Macro F1 Score
LSTM	Word2Vec	0.1194	0.9662	0.4822
RNN	Word2Vec	0.1185	0.9694	0.4973
GRU	Word2Vec	0.0889	0.9778	0.4841
LSTM	GloVe	0.1368	0.9555	0.4649
RNN	GloVe	0.1641	0.9470	0.4732
GRU	GloVe	0.1170	0.9665	0.4741
LSTM	FastText	0.1604	0.9645	0.4973
RNN	FastText	0.1183	0.9650	0.4846
GRU	FastText	0.0900	0.9765	0.4881

Results for Task 1 in Part 2

Model	Word Embedding	Test Loss	Test Macro-F1
RNN	Word2Vec	0.1963	0.4264
LSTM	Word2Vec	0.3798	0.4197
GRU	Word2Vec	0.2264	0.4264
RNN	GloVe	0.2514	0.2603
LSTM	GloVe	0.3393	0.2949
GRU	GloVe	0.2714	0.2813
RNN	FastText	0.2028	0.3813
LSTM	FastText	0.1944	0.4642
GRU	FastText	0.2184	0.4414

Individual Contribution:

1. Saumil Lakra(2021097): Part 1B[text processing], Models of ATE dataset in word2vec, the final compilation code for ATE and NER dataset.
2. Vishal Singh(20215): Embeddings of ATE and NER in Fasttext, GLOVE, graph plots, graph analysation, tables.
3. Sanskar Ranjan(2021096): Part 1A[text processing], BiLSTM CRF model.
4. Jeremiah Rokhum(2021533): BiLSTM CRF and NER dataset and other variations.

It was fairly equal participation among all the members.