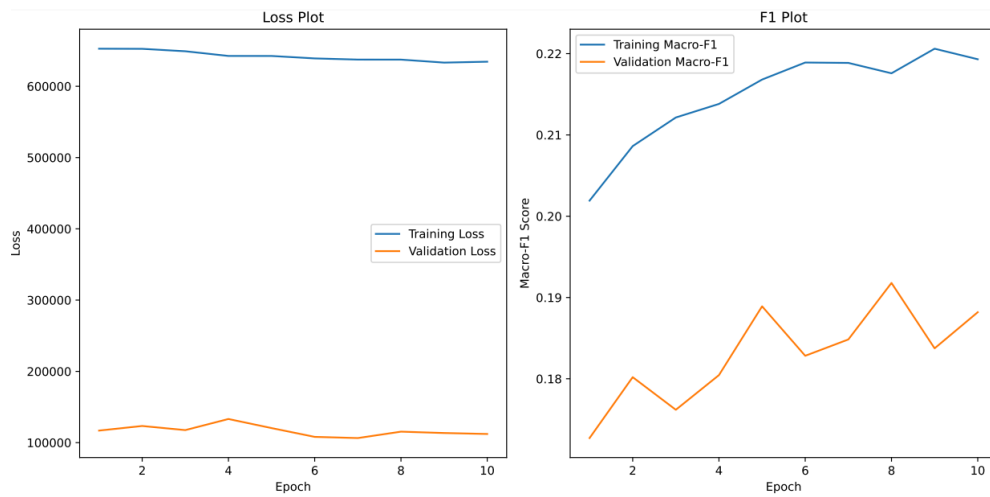


[illegible]

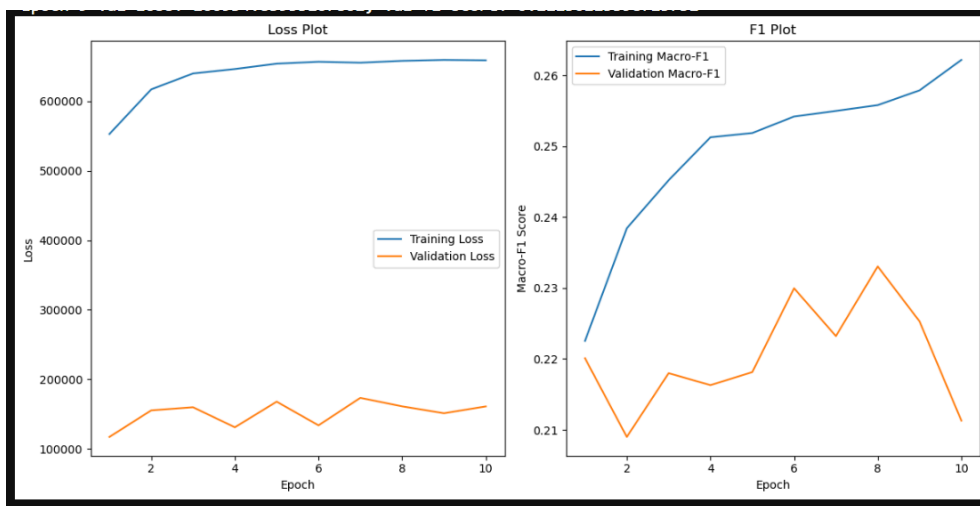
Part - 2

Model 1 - Vanilla RNN Task 1

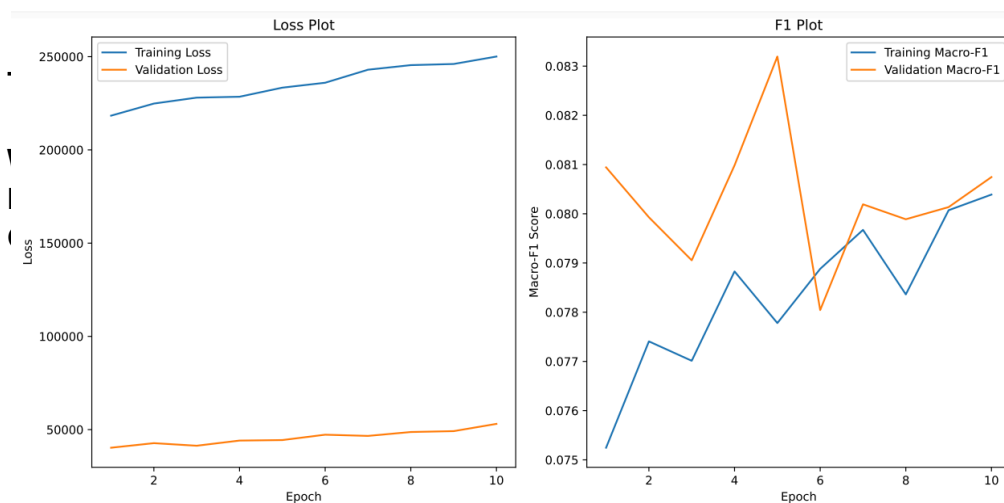
Word2Vec



Fasttext

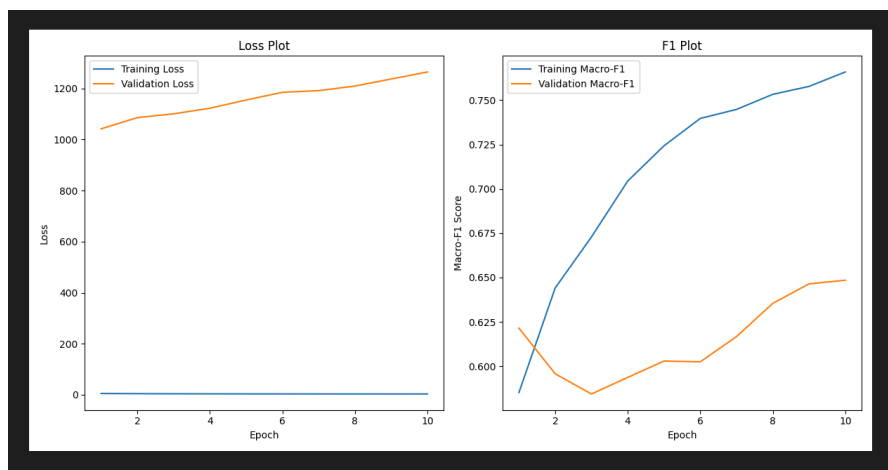


Glove

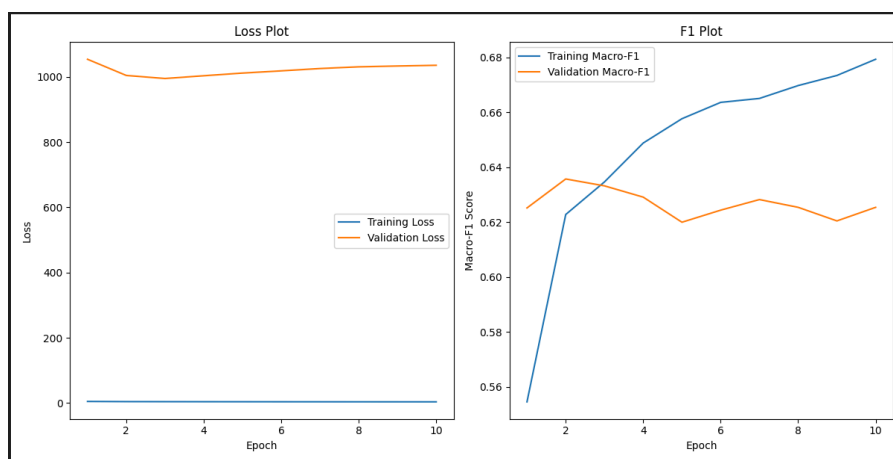


Model 1 - Vanilla RNN Task 1

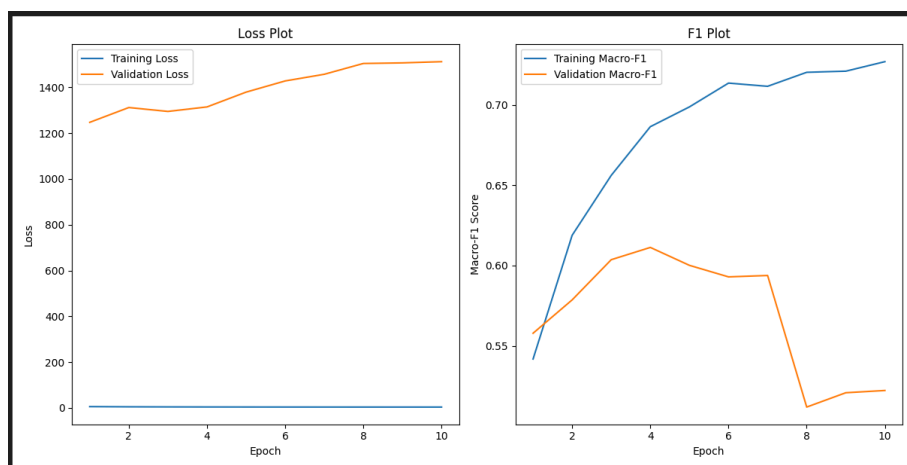
Word2Vec



Fasttext

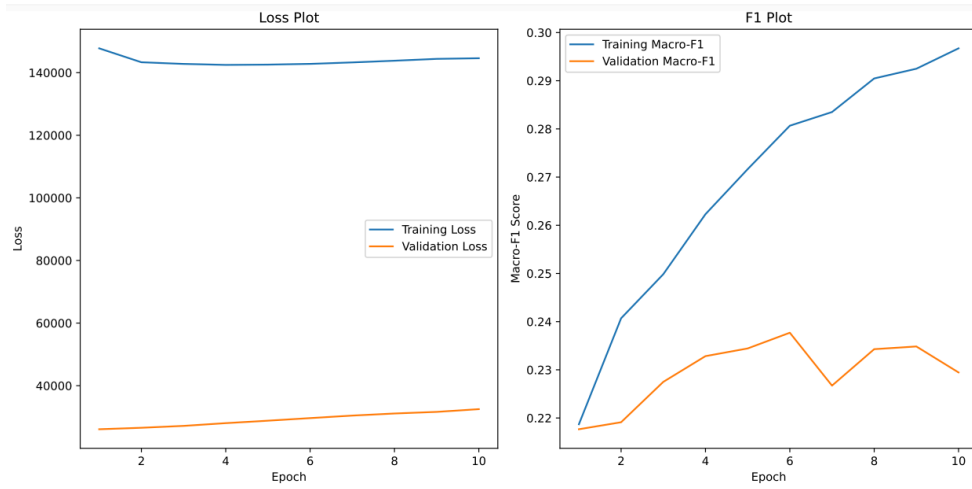


Glove

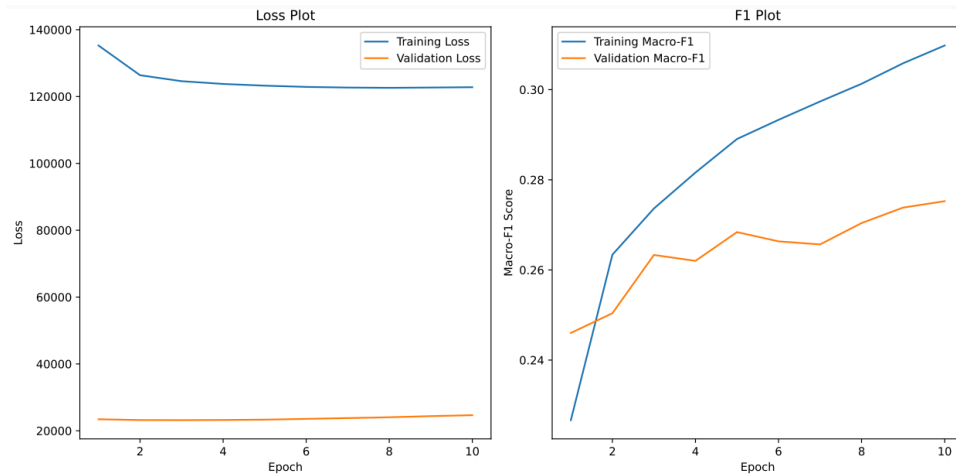


Model 2 - LSTM Task 1

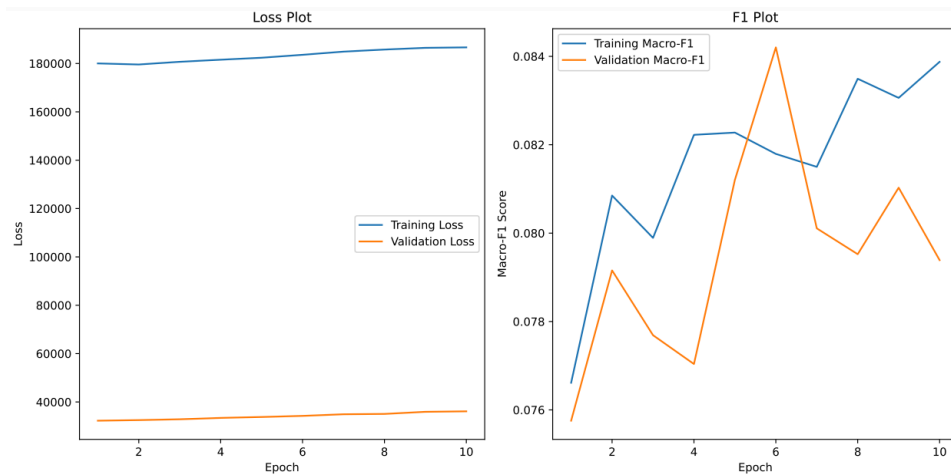
Word2Vec



Fasttext

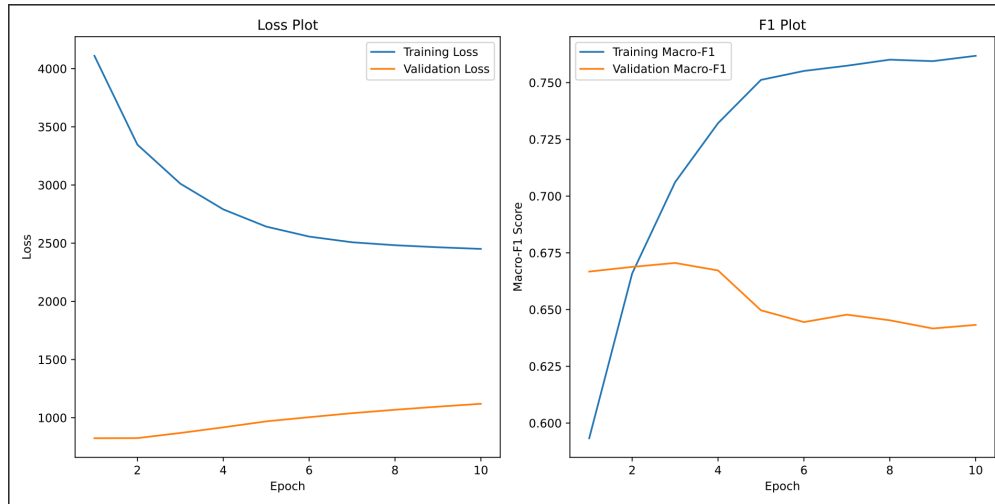


Glove

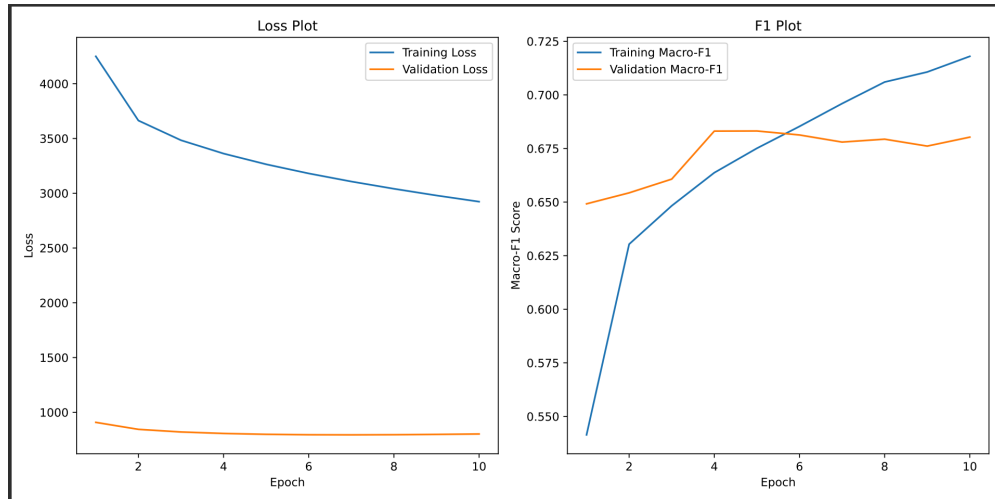


Model 2 Task 2

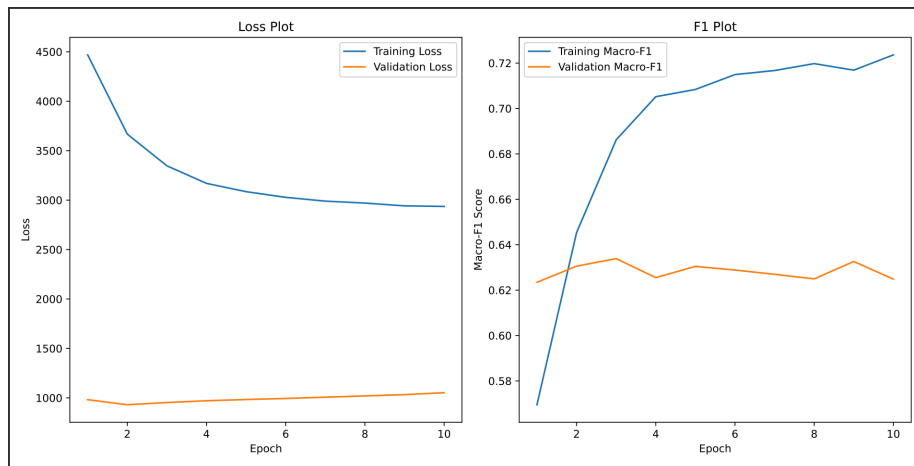
Word2Vec



Fasttext

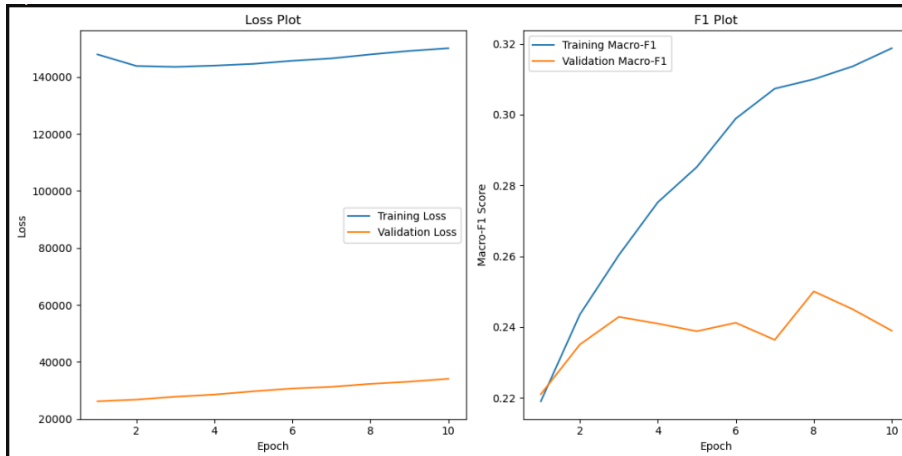


Glove

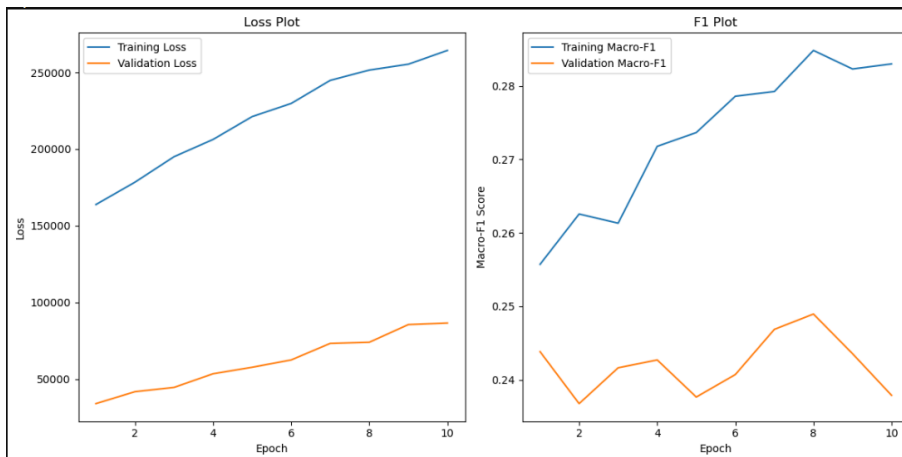


Model 3 - GRU Task 1

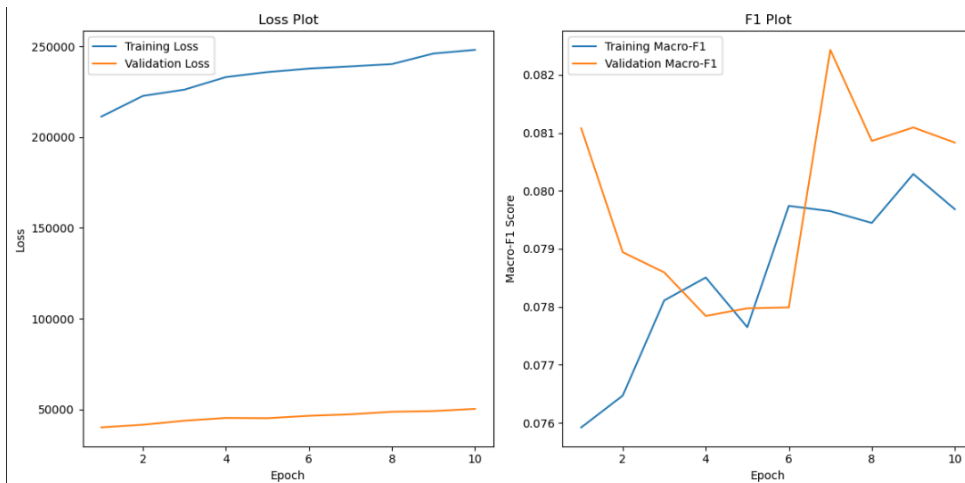
Word2Vec



Fasttext

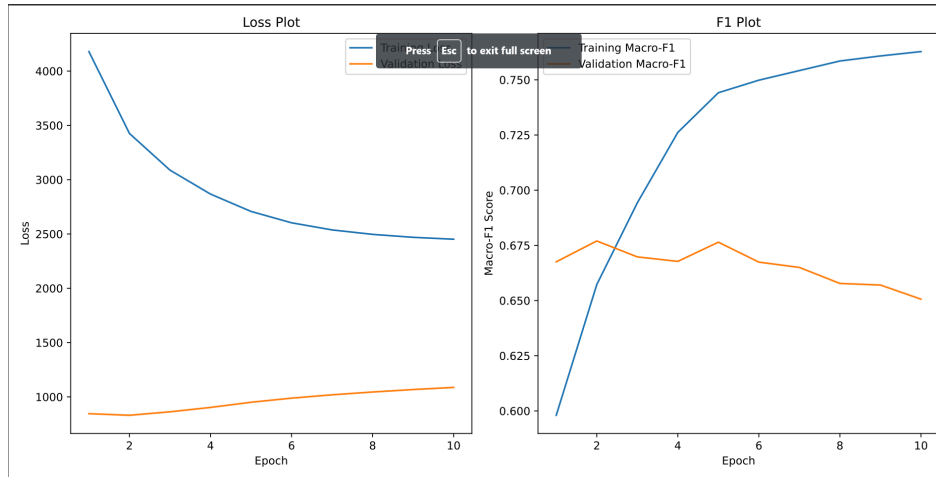


Glove

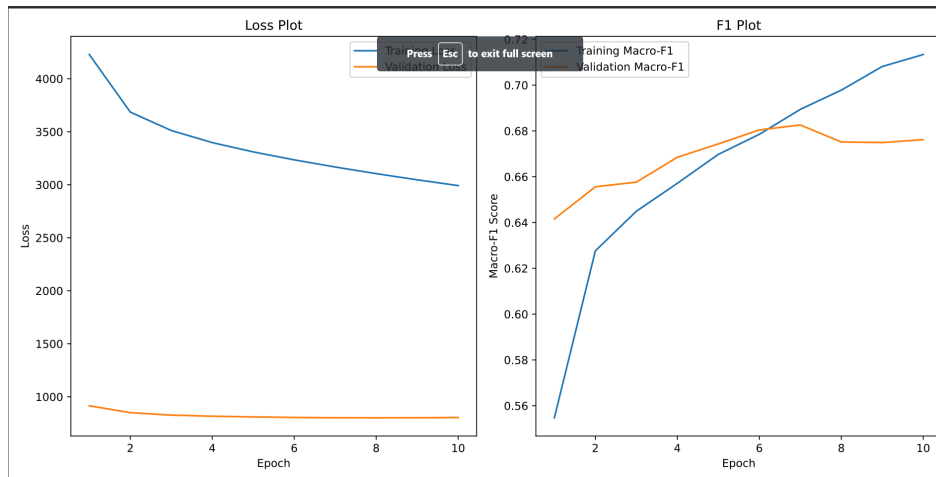


Model 3 Task 2

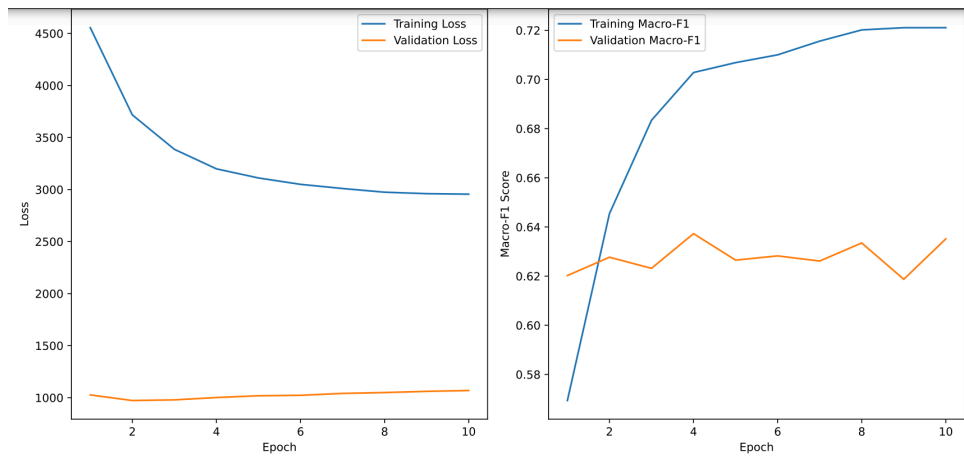
Word2Vec



Fasttext



Glove

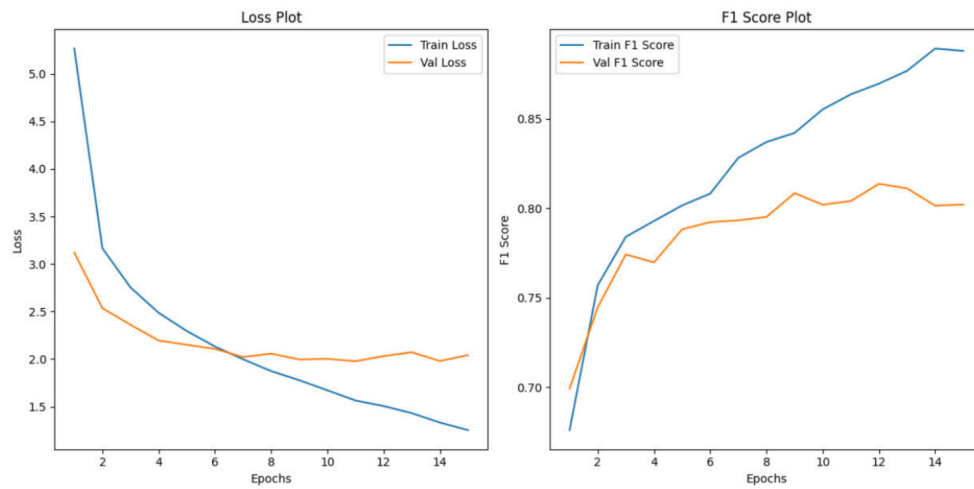


Part - 3
Model 4 Task 1

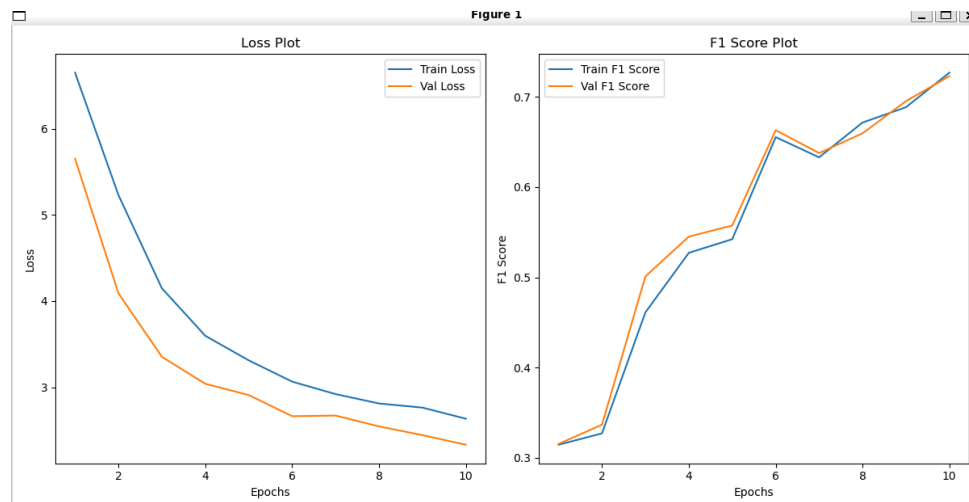
Word2Vec
Fasttext
Glove

Model 4 Task 2

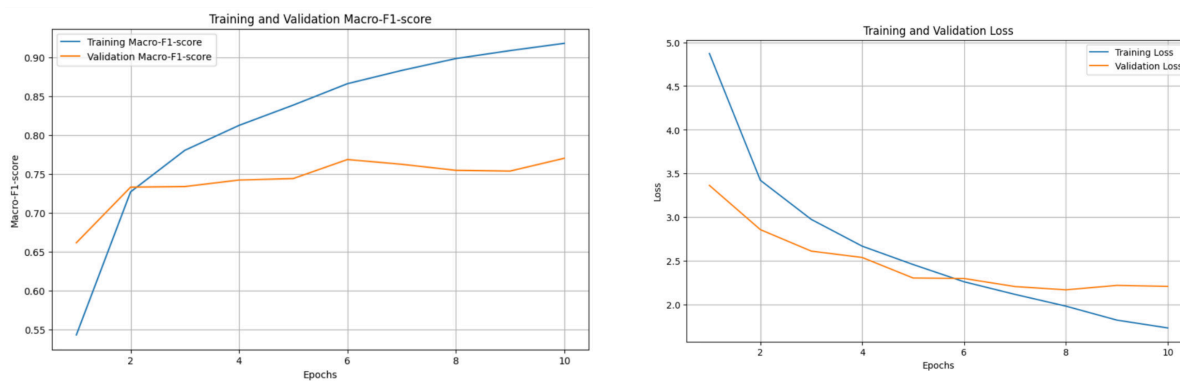
Word2Vec



Fasttext



Glove



Analysis of Plots

1) Loss Plot: Training Loss and Validation Loss V/s Epochs

In all cases training loss decreases with the number of epochs indicating that the model is learning the data on which it is being trained similarly in majority of the cases validation loss also decreases like those in bilstm and gru models for task-2 This

indicates that the model performs well on unseen data, or effective learning.

Conversely, in a few cases like GRU models/RNN models in task 1 the validation loss increases but training loss continues to decrease, this may be due to a case of overfitting of the model.

2) F1 Plot: Training Macro-F1-score and Validation Macro-F1-score V/s Epochs

The training and Validation Macro-F1-scores versus epochs provides insights into the model's performance throughout training. We observe in most cases of task 2 that the macro f1 scores for training and validation increase with the number of epochs.

Indicating a good performance. However, for a few cases of task 1 models the training/validation macro-f1 scores also increase with epochs.

Furthermore, we observe that bilstm models significantly perform better than those used in part - 2 i.e. gru, vanilla RNN & lstm. This can be seen from cases of task 2 where test macro - f1 scores of embeddings: word2vec, GloVe, and fasttext for bilstm increase to 0.78, 0.72 & 0.72 respectively as compared to part - 1 models where the test - macro f1 score lies between 0.60 to 0.65.

Task 1

Model	Embedding	Accuracy	Macro - F1
Vanilla RNN	word2vec	0.837475582268970	0.251589699869998
Vanilla RNN	GloVe	0.827407963936889	0.071787571031267
Vanilla RNN	fasttext	0.775326821938392	0.206682039266887
LSTM	word2vec	0.8524117205108941	0.286512615687765
LSTM	GloVe	0.8311945905334335	0.1211761621747104
LSTM	fasttext	0.860285499624342	0.323062041423525
GRU	word2vec	0.850608564988730	0.278815438936718
GRU	GloVe	0.829511645379414	0.1108778395414184
GRU	fasttext	0.854545454545454	0.2974895492911865
Bi-LSTM CRF	word2vec		
Bi-LSTM CRF	GloVe		
Bi-LSTM CRF	fasttext		

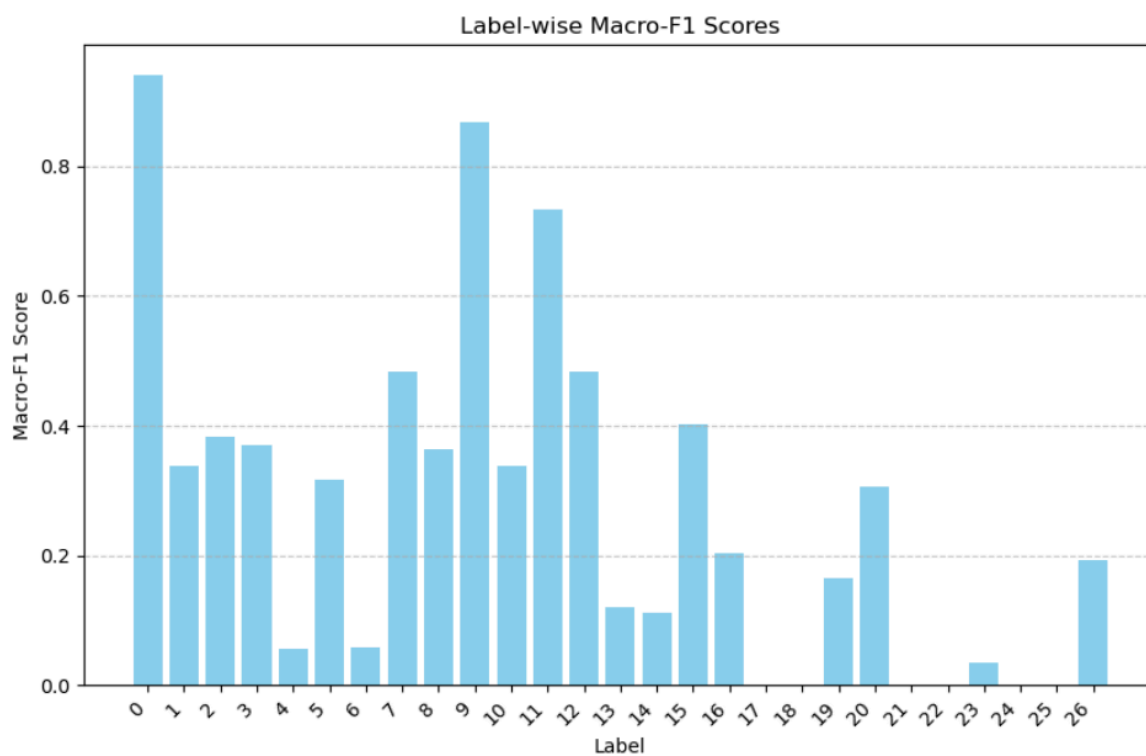
Task 2

Model	Embedding	Accuracy	Macro - F1
Vanilla RNN	word2vec	0.888416167085670	0.605036952725088
Vanilla RNN	GloVe	0.794430477663894	0.516486985362242
Vanilla RNN	fasttext	0.890543415200154	0.630067576450300
LSTM	word2vec	0.904273834848191	0.644236345265209
LSTM	GloVe	0.895764842390253	0.602940538123794
LSTM	fasttext	0.908334944884935	0.663915961866509
GRU	word2vec	0.9075614001160317	0.654639735617689
GRU	GloVe	0.893444208083542	0.605427972017266
GRU	fasttext	0.906401082962676	0.655980855096665
Bi-LSTM CRF	word2vec	0.9444	0.7874
Bi-LSTM CRF	GloVe	0.9056	0.7232

Bi-LSTM CRF	fasttext	0.9252	0.7218
-------------	----------	--------	--------

Label Wise F1 Scores

For Model 2 LSTM - Fasttext embedding



Contributions

Model 1 - Vanilla RNN -> Abhinav

Model 2 - LSTM -> Lakshya

Model 3 - GRU -> Srimant

Model 4 - Bi-LSTM -> Hardik