NLP assignment 2

Group number 11

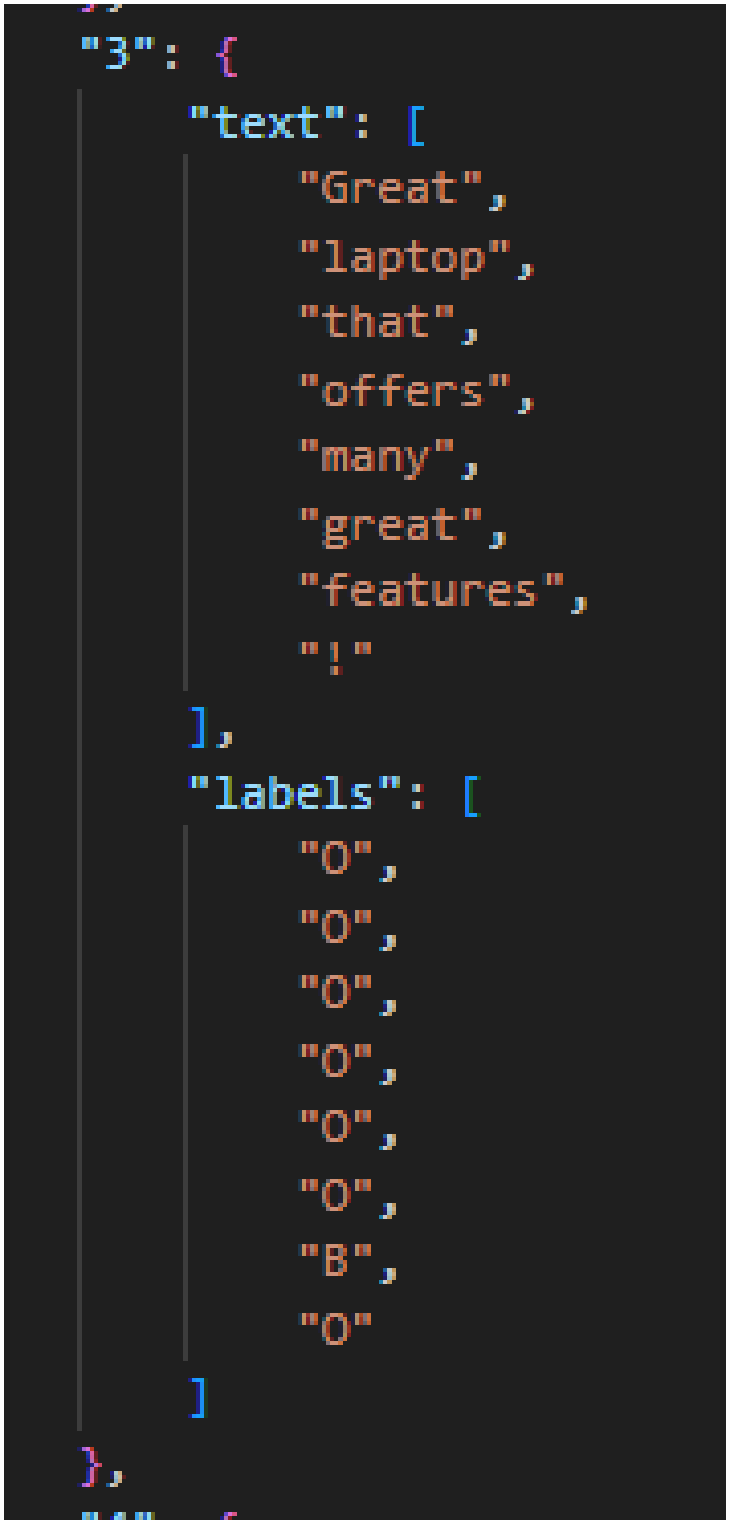
Dataset 1 snippets:


```
"3d966c8225b24464bdfeb759747798fa": {
    "text": [
        "Even",
        "according",
        "to",
        "the",
        "complainant",
        "and",
        "prosecution",
        "witness",
        "."
    ],
    "labels": [
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O"
    ]
},
```

```
"4bbb0629e66146edaf4ac7bde47062fb": {
    "text": [
        "The",
        "factum",
        "of",
        "accident",
        ",",
        "allegation",
        "of",
        "rash",
        "and",
        "negligent",
        "driving",
        "causing",
        "death",
        "of",
        "Sukendra",
        "Pal",
        "Singh",
        "were",
        "denied",
        "."
    ],
    "labels": [
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "B_OTHER_PERSON",
        "I_OTHER_PERSON",
        "I_OTHER_PERSON",
        "O",
        "O",
        "O"
    ]
},
```

Dataset 2 snippets:


```
"2": {
    "text": [
        "Easy",
        "to",
        "start",
        "up",
        "and",
        "does",
        "not",
        "overheat",
        "as",
        "much",
        "as",
        "other",
        "laptops",
        "."
    ],
    "labels": [
        "O",
        "O",
        "B",
        "I",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O",
        "O"
    ]
},
```



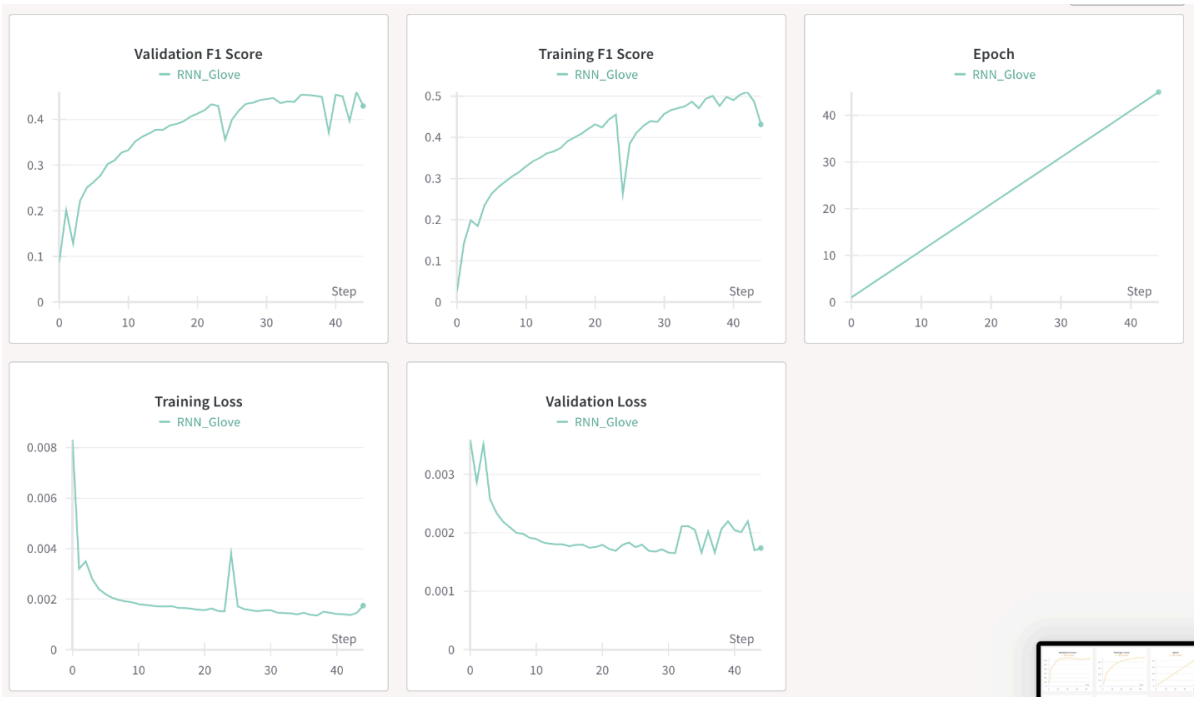
Additional Preprocessing:

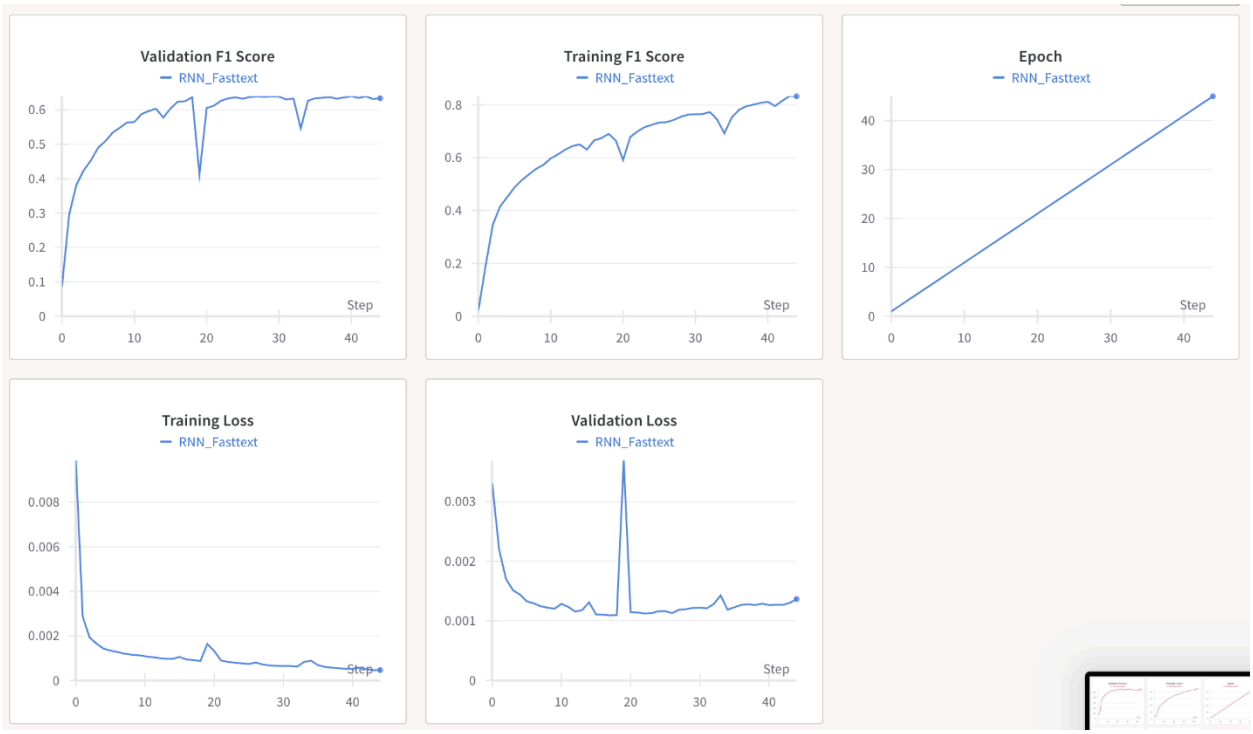
Special emphasis was given on separating tokens that were followed by sentence competition markers like full stop, question mark and other punctuation marks. Rules mentioned in the original json file were applied to generate and label the given texts.

Plots:

Part 2 dataset 1:

a. RNN

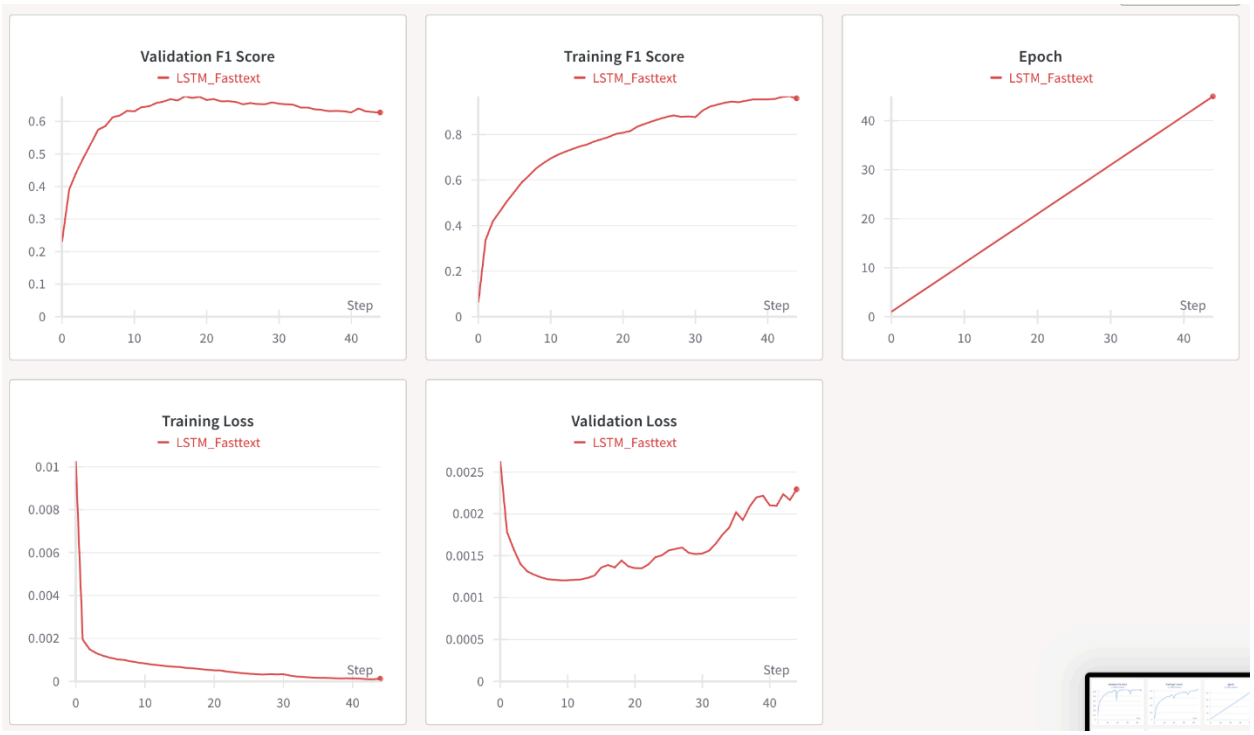



Validation F1 Score
RNN_Fasttext
Step
Training F1 Score
RNN_Fasttext
Step
Epoch
RNN_Fasttext
Step
Training Loss
RNN_Fasttext
Step
Validation Loss
RNN_Fasttext
Step

b. LSTM


Validation F1 Score
LSTM_Word2Vec
Step
Training F1 Score
LSTM_Word2Vec
Step
Epoch
LSTM_Word2Vec
Step
Training Loss
LSTM_Word2Vec
Step
Validation Loss
LSTM_Word2Vec
Step


Validation F1 Score
LSTM_Glove
Training F1 Score
LSTM_Glove
Epoch
LSTM_Glove
Training Loss
LSTM_Glove
Validation Loss
LSTM_Glove
Step

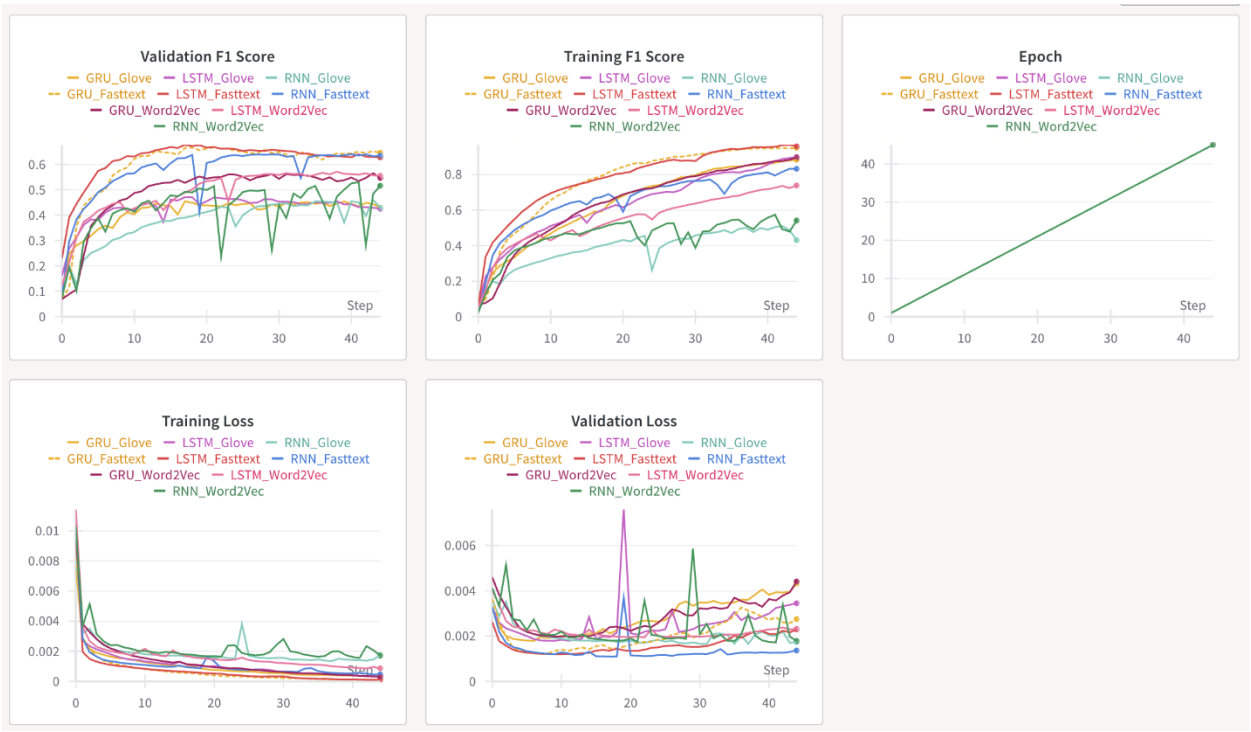

Validation F1 Score
LSTM_Fasttext
Training F1 Score
LSTM_Fasttext
Epoch
LSTM_Fasttext
Training Loss
LSTM_Fasttext
Validation Loss
LSTM_Fasttext
Step

c. GRU


Validation F1 Score
GRU_Glove
Step
Training F1 Score
GRU_Glove
Step
Epoch
GRU_Glove
Step
Training Loss
GRU_Glove
Step
Validation Loss
GRU_Glove
Step


Validation F1 Score
GRU_Fasttext
Step
Training F1 Score
GRU_Fasttext
Step
Epoch
GRU_Fasttext
Step
Training Loss
GRU_Fasttext
Step
Validation Loss
GRU_Fasttext
Step

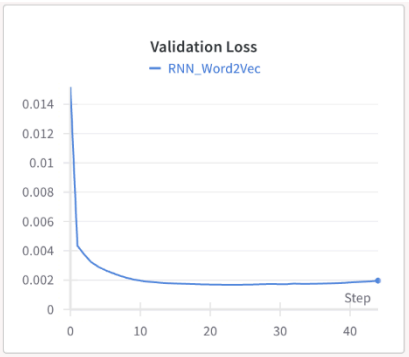
d. Combined

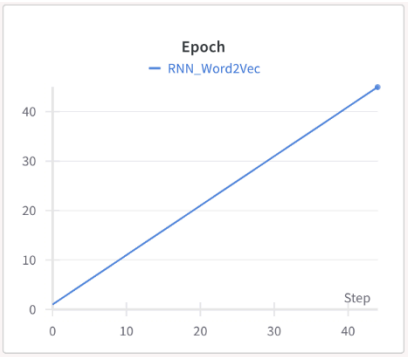


Part 2 dataset 2:

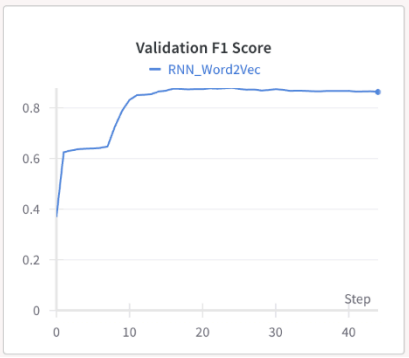
a. RNN


Training F1 Score
RNN_Word2Vec
0.8
0.6
0.4
0.2
0
Step
0
10
20
30
40

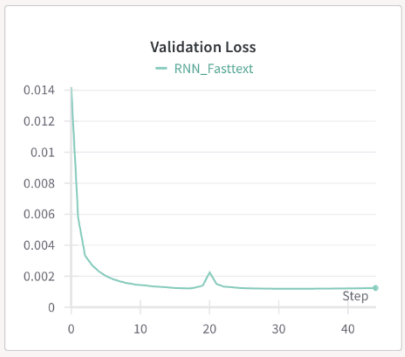

Validation Loss
RNN_Word2Vec
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
Step
0
10
20
30
40

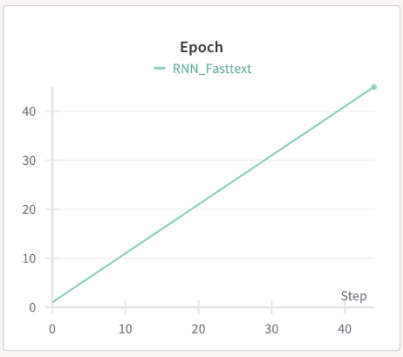

Epoch
RNN_Word2Vec
40
30
20
10
0
Step
0
10
20
30
40


Training Loss
RNN_Word2Vec
0.04
0.03
0.02
0.01
0
Step
0
10
20
30
40

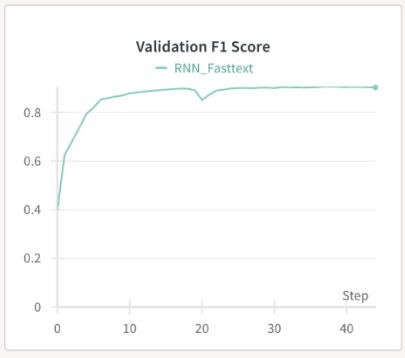

Validation F1 Score
RNN_Word2Vec
0.8
0.6
0.4
0.2
0
Step
0
10
20
30
40

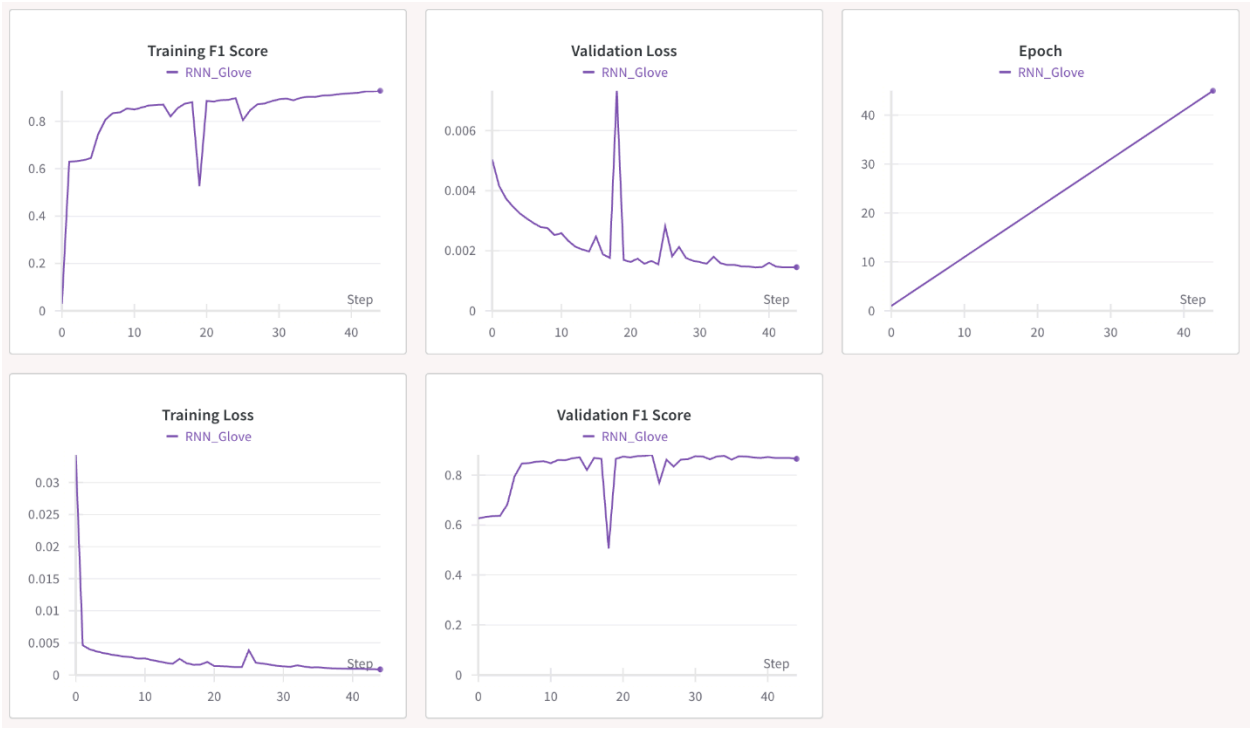

Training F1 Score
RNN_Fasttext
0.8
0.6
0.4
0.2
0
Step
0
10
20
30
40


Validation Loss
RNN_Fasttext
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
Step
0
10
20
30
40

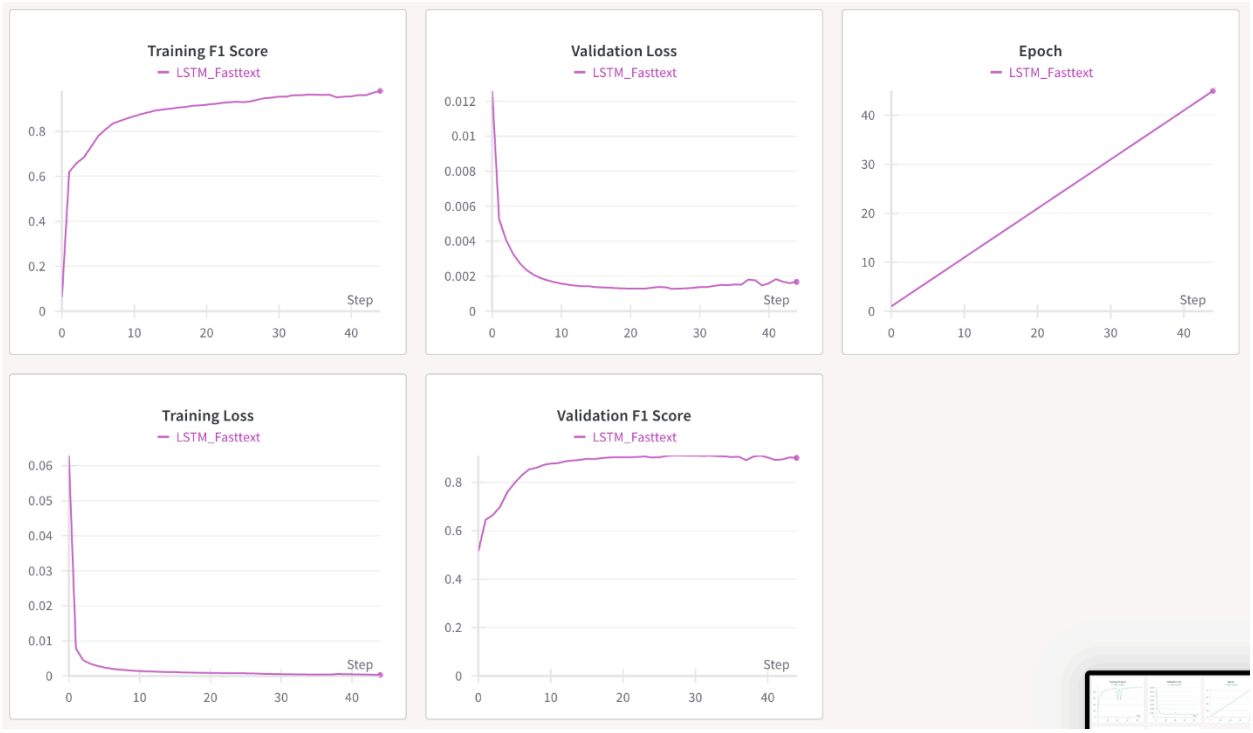

Epoch
RNN_Fasttext
40
30
20
10
0
Step
0
10
20
30
40

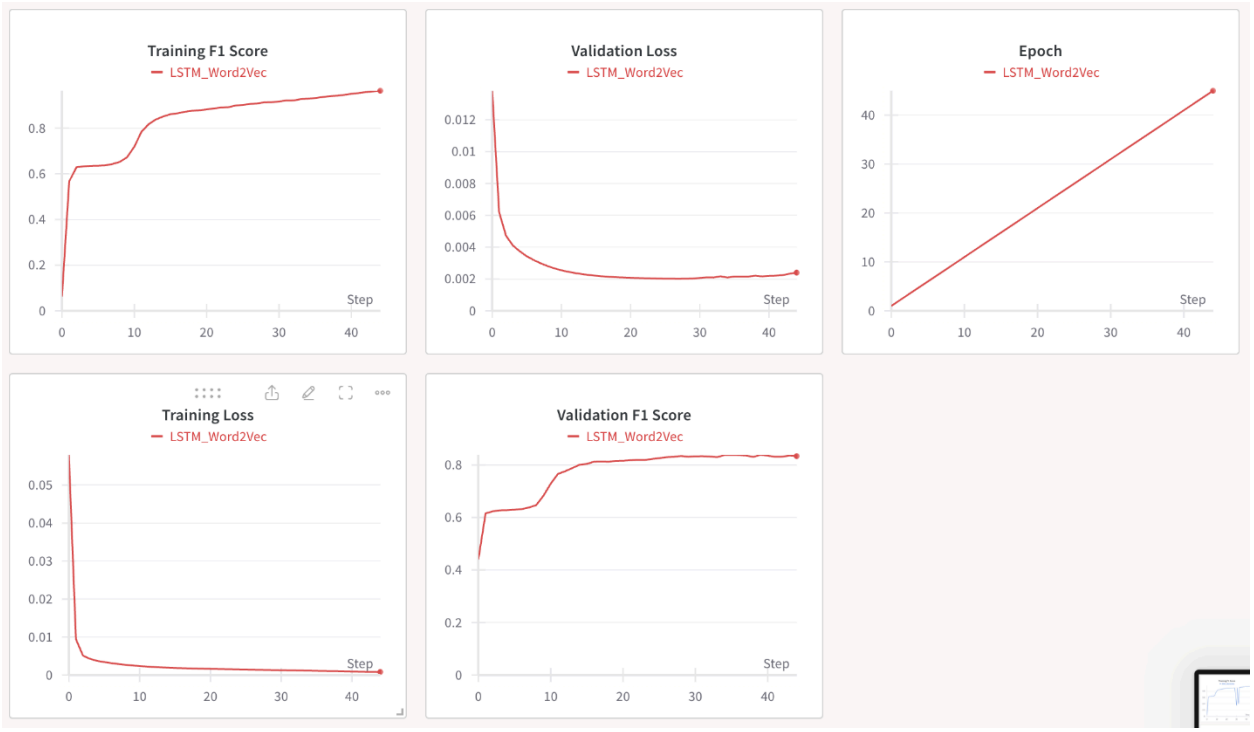

Training Loss
RNN_Fasttext
0.04
0.03
0.02
0.01
0
Step
0
10
20
30
40

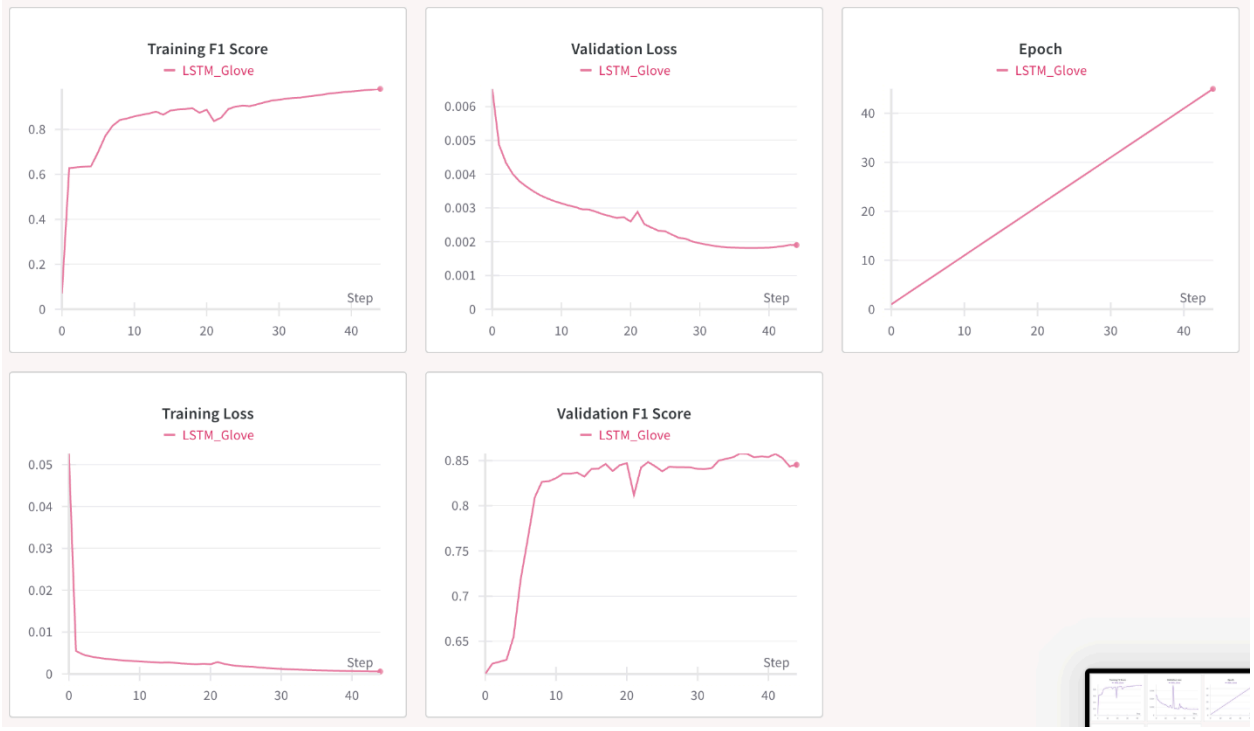

Validation F1 Score
RNN_Fasttext
0.8
0.6
0.4
0.2
0
Step
0
10
20
30
40


Training F1 Score
RNN_Glove
Validation Loss
RNN_Glove
Epoch
RNN_Glove
Training Loss
RNN_Glove
Validation F1 Score
RNN_Glove
Step

b. LSTM

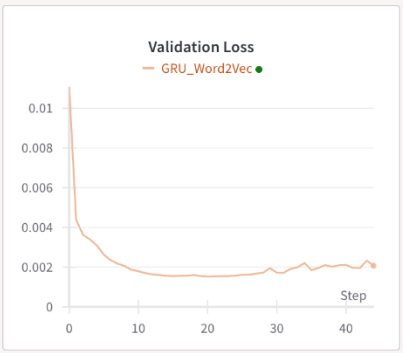

Training F1 Score
LSTM_Fasttext
Validation Loss
LSTM_Fasttext
Epoch
LSTM_Fasttext
Training Loss
LSTM_Fasttext
Validation F1 Score
LSTM_Fasttext
Step

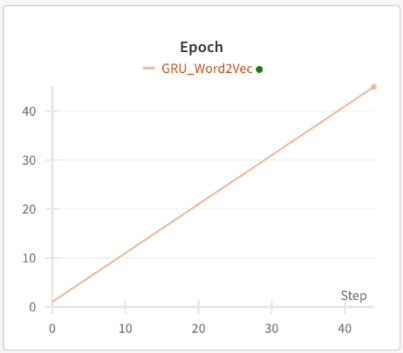

Training F1 Score
LSTM_Word2Vec
Validation Loss
LSTM_Word2Vec
Epoch
LSTM_Word2Vec
Training Loss
LSTM_Word2Vec
Validation F1 Score
LSTM_Word2Vec
Step


Training F1 Score
LSTM_Glove
Validation Loss
LSTM_Glove
Epoch
LSTM_Glove
Training Loss
LSTM_Glove
Validation F1 Score
LSTM_Glove
Step

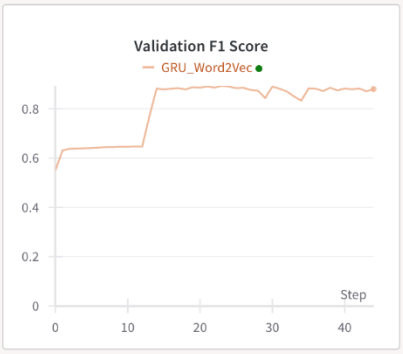
c. GRU

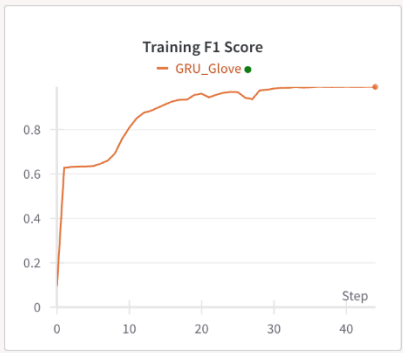

Training F1 Score
GRU_Word2Vec
0.8
0.6
0.4
0.2
0
Step
0
10
20
30
40

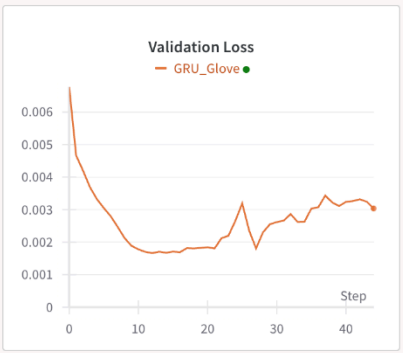

Validation Loss
GRU_Word2Vec
0.01
0.008
0.006
0.004
0.002
0
Step
0
10
20
30
40

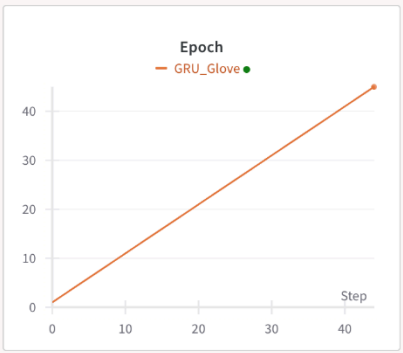

Epoch
GRU_Word2Vec
40
30
20
10
0
Step
0
10
20
30
40


Training Loss
GRU_Word2Vec
0.04
0.03
0.02
0.01
0
Step
0
10
20
30
40

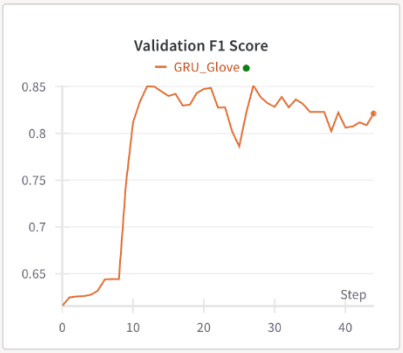

Validation F1 Score
GRU_Word2Vec
0.8
0.6
0.4
0.2
0
Step
0
10
20
30
40


Training F1 Score
GRU_Glove
0.8
0.6
0.4
0.2
0
Step
0
10
20
30
40


Validation Loss
GRU_Glove
0.006
0.005
0.004
0.003
0.002
0.001
0
Step
0
10
20
30
40

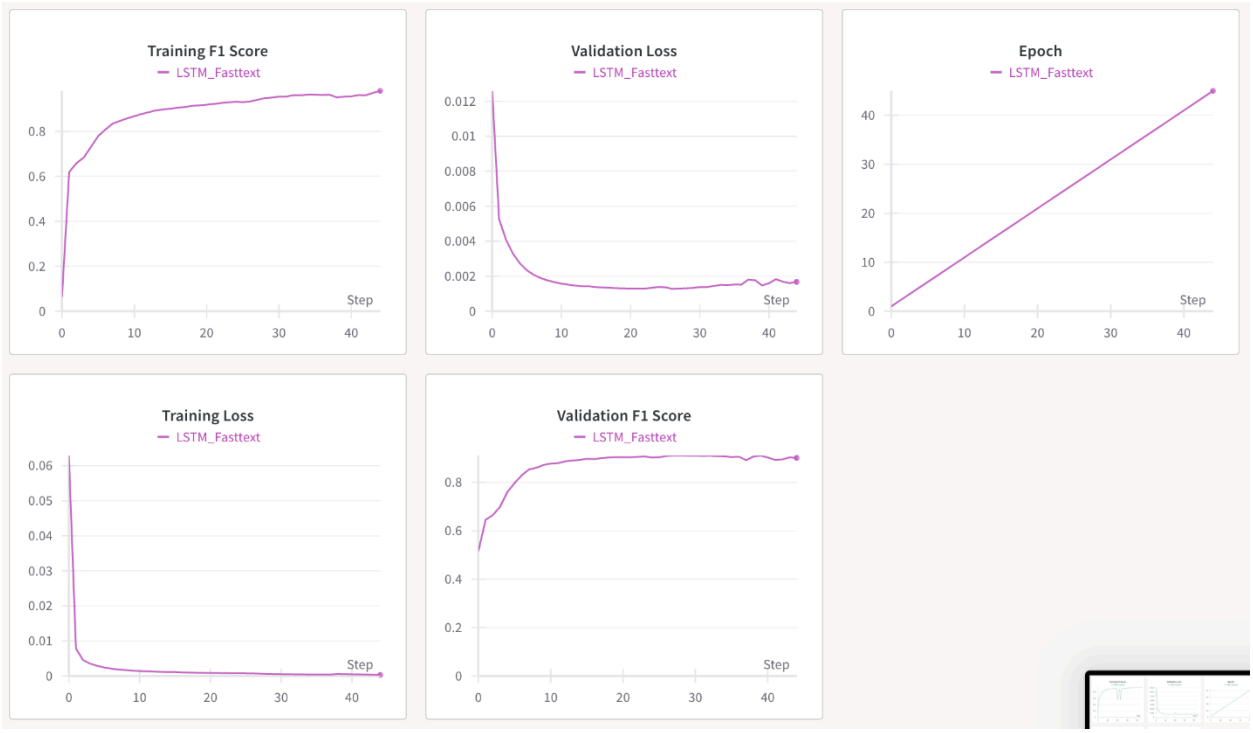

Epoch
GRU_Glove
40
30
20
10
0
Step
0
10
20
30
40

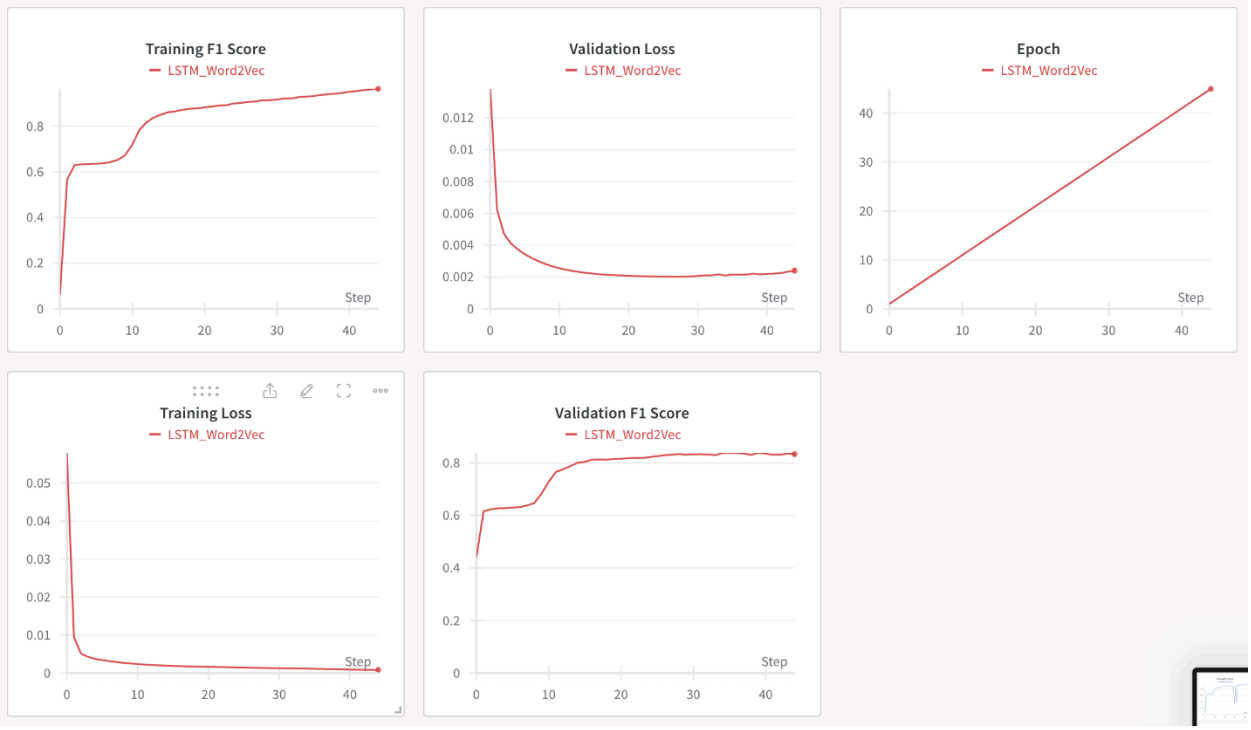

Training Loss
GRU_Glove
0.03
0.02
0.01
0
Step
0
10
20
30
40


Validation F1 Score
GRU_Glove
0.85
0.8
0.75
0.7
0.65
Step
0
10
20
30
40


Training F1 Score
GRU_Fasttext
Validation Loss
GRU_Fasttext
Epoch
GRU_Fasttext
Training Loss
GRU_Fasttext
Validation F1 Score
GRU_Fasttext
Step

d. BiLSTM


Training F1 Score
LSTM_Fasttext
Validation Loss
LSTM_Fasttext
Epoch
LSTM_Fasttext
Training Loss
LSTM_Fasttext
Validation F1 Score
LSTM_Fasttext
Step


Training F1 Score
LSTM_Word2Vec
Step
Validation Loss
LSTM_Word2Vec
Step
Epoch
LSTM_Word2Vec
Step
Training Loss
LSTM_Word2Vec
Step
Validation F1 Score
LSTM_Word2Vec
Step

d. Combined


Training F1 Score
GRU_Word2Vec GRU_Fasttext LSTM_Fasttext
RNN_Fasttext GRU_Glove LSTM_Glove
RNN_Glove LSTM_Word2Vec RNN_Word2Vec
Step
Validation Loss
GRU_Word2Vec GRU_Fasttext LSTM_Fasttext
RNN_Fasttext GRU_Glove LSTM_Glove
RNN_Glove LSTM_Word2Vec RNN_Word2Vec
Step
Epoch
GRU_Word2Vec GRU_Fasttext LSTM_Fasttext
RNN_Fasttext GRU_Glove LSTM_Glove
RNN_Glove LSTM_Word2Vec RNN_Word2Vec
Step
Training Loss
GRU_Word2Vec GRU_Fasttext LSTM_Fasttext
RNN_Fasttext GRU_Glove LSTM_Glove
RNN_Glove LSTM_Word2Vec RNN_Word2Vec
Step
Validation F1 Score
GRU_Word2Vec GRU_Fasttext LSTM_Fasttext
RNN_Fasttext GRU_Glove LSTM_Glove
RNN_Glove LSTM_Word2Vec RNN_Word2Vec
Step

General Analysis and Comparison of word embeddings

1. **Word2Vec:**
    a. *Pros:*
        i. Effective at capturing semantic relationships between words.
        ii. Can handle out-of-vocabulary words to some extent.
    b. *Cons:*
        i. May not capture subword information.
        ii. Fixed-size embeddings.
2. **GloVe (Global Vectors for Word Representation):**
    a. *Pros:*
        i. Utilizes global statistics, capturing word co-occurrence information effectively.
        ii. Generally performs well on semantic tasks.
    b. *Cons:*
        i. May struggle with rare or out-of-vocabulary words.
        ii. Fixed-size embeddings.
3. **FastText:**
    a. *Pros:*
        i. Considers subword information, effective for handling rare words.
        ii. Can handle out-of-vocabulary words by breaking them down into subword components.
    b. *Cons:*
        i. May generate larger embeddings.
        ii. May not capture long-range semantic relationships as well.

Analysis of the plots:

1. We observe that the f1 scores of all the embeddings follow the order: fasttext > word2vec > glove in the dataset 1.
2. We observe that the f1 scores of all the embeddings follow the order: fasttext > glove > word2vec in the dataset 2.
3. The better performance of fasttext in both datasets can be attributed to the fact that it works very well in smaller context and does not consider longer dependencies and

semantics. Since our problem was not focused on deriving or considering long term semantics, addition of the same would have hindered the performance.

4. We also observed that word2vec outperformed glove embedding in dataset 1 whereas opposite was true for dataset 2.
5. This can be attributed to the fact that in first problem of NER, many word labels were out of vocabulary which is a weak point of glove embedding, hence word2vec outperformed it.
6. In the second dataset of ATE, glove performed better than word2vec because it captures co-occurrence information quite well that can be quite useful in associating terms with a single entity like emotion or object.

Table comparing the performance of all the models:

Dataset 1

Model number	Embedding Used	Accuracy	Macro f1
RNN word2vec	word2vec	94.634	0.538
RNN glove	glove	93.678	0.441
RNN fasttext	fasttext	95.961	0.640
LSTM word2vec	word2vec	94.674	0.571
LSTM glove	glove	93.978	0.470
LSTM fasttext	fasttext	95.261	0.676
GRU word2vec	word2vec	94.634	0.539
GRU glove	glove	93.678	0.451
GRU fasttext	fasttext	94.961	0.647
BiLSTM word2vec	word2vec	95.122	0.653
BiLSTM glove	glove	94.672	0.537
BiLSTM fasttext	fasttext	96.012	0.701

Dataset 2

Model number	Embedding Used	Accuracy/100	Macro f1
RNN word2vec	word2vec	0.977	0.879
RNN glove	glove	0.983	0.744
RNN fasttext	fasttext	0.986	0.900
LSTM word2vec	word2vec	0.981	0.846
LSTM glove	glove	0.984	0.751
LSTM fasttext	fasttext	0.986	0.904
GRU word2vec	word2vec	0.982	0.866
GRU glove	glove	0.982	0.695
GRU fasttext	fasttext	0.987	0.908
BiLSTM word2vec	word2vec	0.985	0.854
BiLSTM glove	glove	0.987	0.821
BiLSTM fasttext	fasttext	0.988	0.924

Contribution of members:

Each member contributed equally to the project and most of the things were done in sync overlapping with each other.