CS563: Natural Language Processing

Indian Institute of Technology Patna

Assignment 3

Feed Forward NN and RNNs

Group Members:

- Muhammed Sinan C K (1901CS38)
- Vaishakh Sreekanth Menon (1901CS68)
- Varsha Tumburu (1901CS69)

Aim

We target to implement Feed-Forward NN and RNN for Binary and multi-class sentiment analysis on the ACL IMDB dataset and the SemEval dataset.

Datasets

IMDB Dataset

This is a binary class dataset which consists of reviews of movie from multiple users. The reviews are stored in a txt file and are usually a paragraph long. The aim of the dataset is to provide reviews which tag positive/negative sentiment and accordingly predict that.

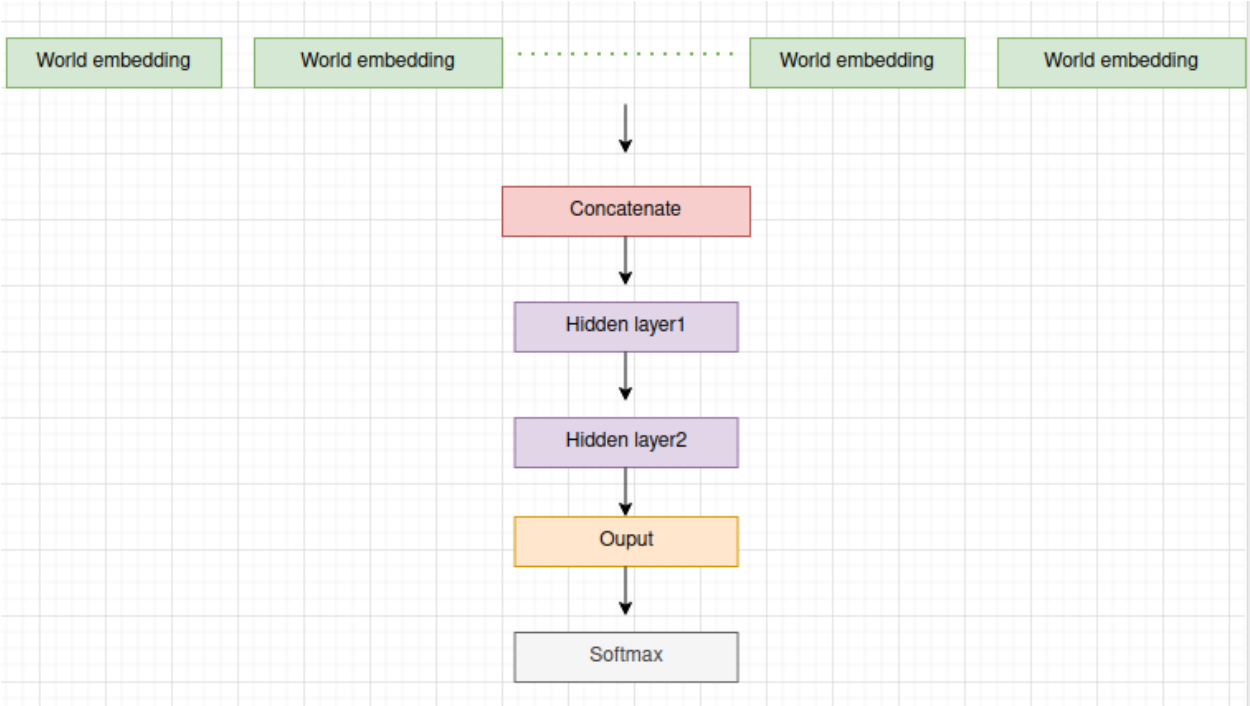
SemEval Dataset

This is a multi-class dataset which consists of tweets from various Twitter users. In a similar manner, the tweets are classified positive/negative/neutral sentiment. We must train models to pick up words to classify unclassified tweets in the same way.

Feed Forward Neural Network

Explain and draw the architecture of Feed-Forward NN that you are proposing with justification. Describe the features of Feed-Forward NN.

Information flows only in one direction in a feed forward neural network from input nodes through hidden nodes to output nodes. FFNNs consist of an input/output layer, multiple hidden layers, activation functions, weights and bias. Backpropagation is the major factor that helps in training allowing the model to be fit better utilizing its loss and gradient.
FFNNs are commonly used for classification and regression tasks, and can be used in combination with other techniques such as dropout, batch normalization, and early stopping to improve their performance.



Dimension of Feedforward architecture

- Activation function - Relu
- Concatenate of word embedding - [batch size =32,seq length, embedding dim=100]
- Hidden layer1 - [batch size, 256]
- Hidden layer2- [batch size, 128]
- Output - [batch size, number of classes]
- Softmax - [batch size, number of classes]

Recurrent Neural Network

Clearly discriminate between binary class and multi class loss functions.

In the case of binary class classification, there are only two classes into which we can classify into whereas for multiple class classification there are more than two classes. We use Binary cross entropy loss for the binary class problem and multiple cross entropy function for the multi class entropy problem.

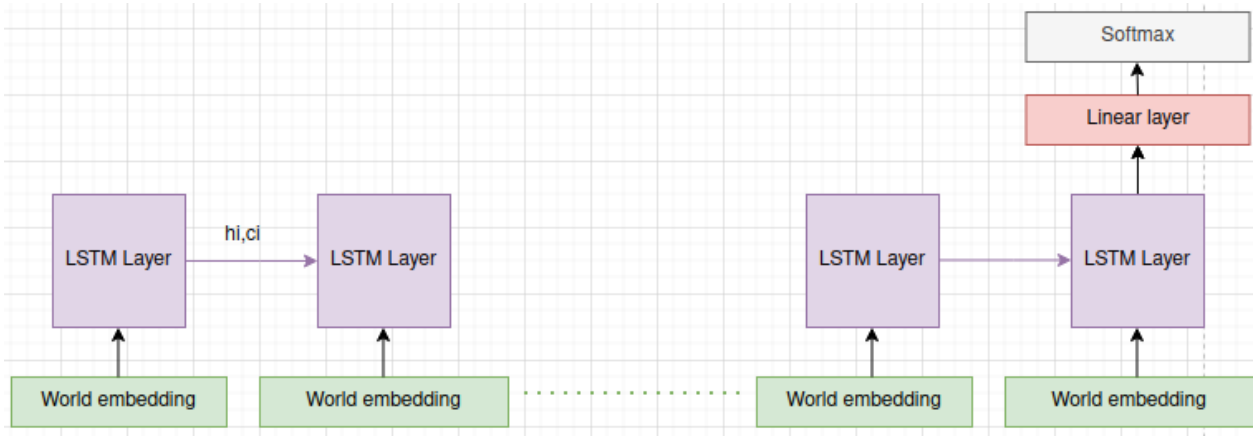
Binary cross entropy function :

Loss = -(y * log(p) + (1 - y) * log(1 - p))

Multi class cross entropy function :

Loss = -∑(y * log(p))

Recurrent neural network architecture



Dimensions

- Activation function - Relu
- LSTM layer input- [seq_length,batch size,embedding_dim]
- LSTM layer output - [batch size, hidden dimension=256]
- Linear Layer - [batch size, number of classes]
- Softmax output - [batch size, number of classes]

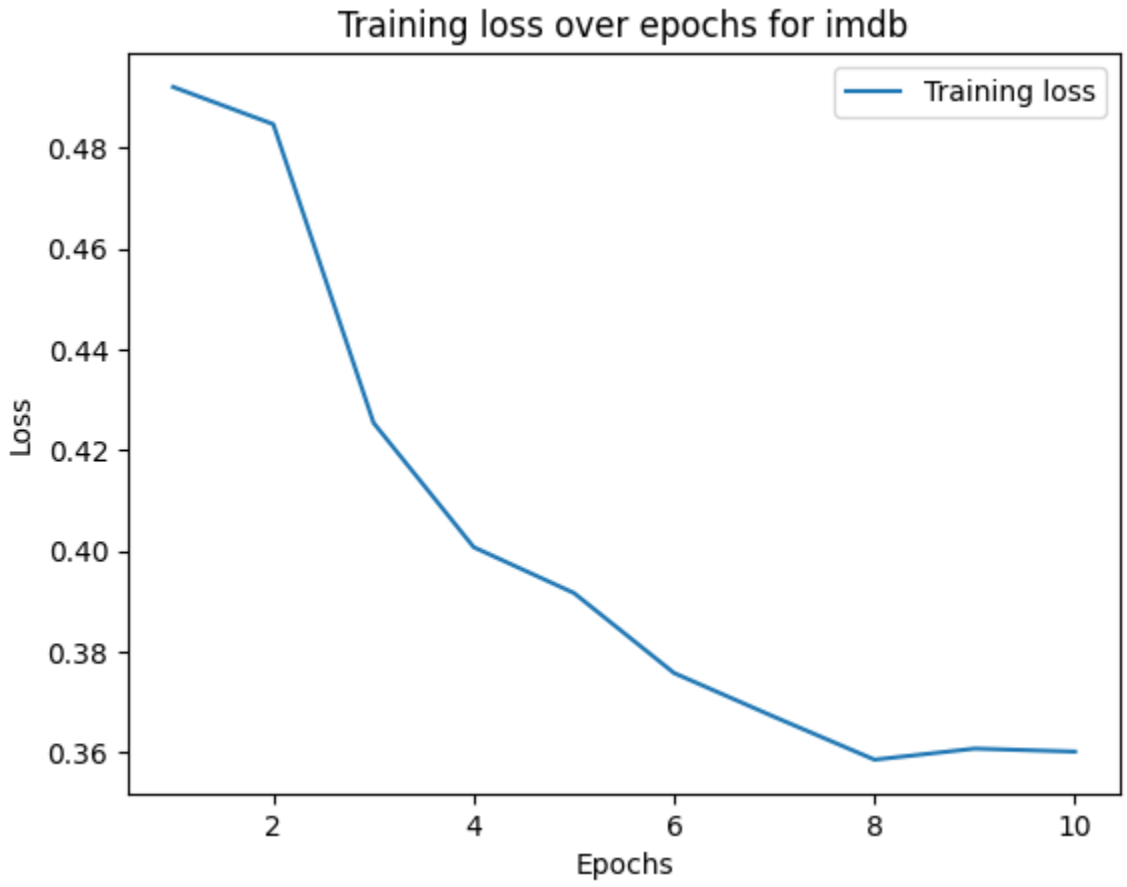
Results

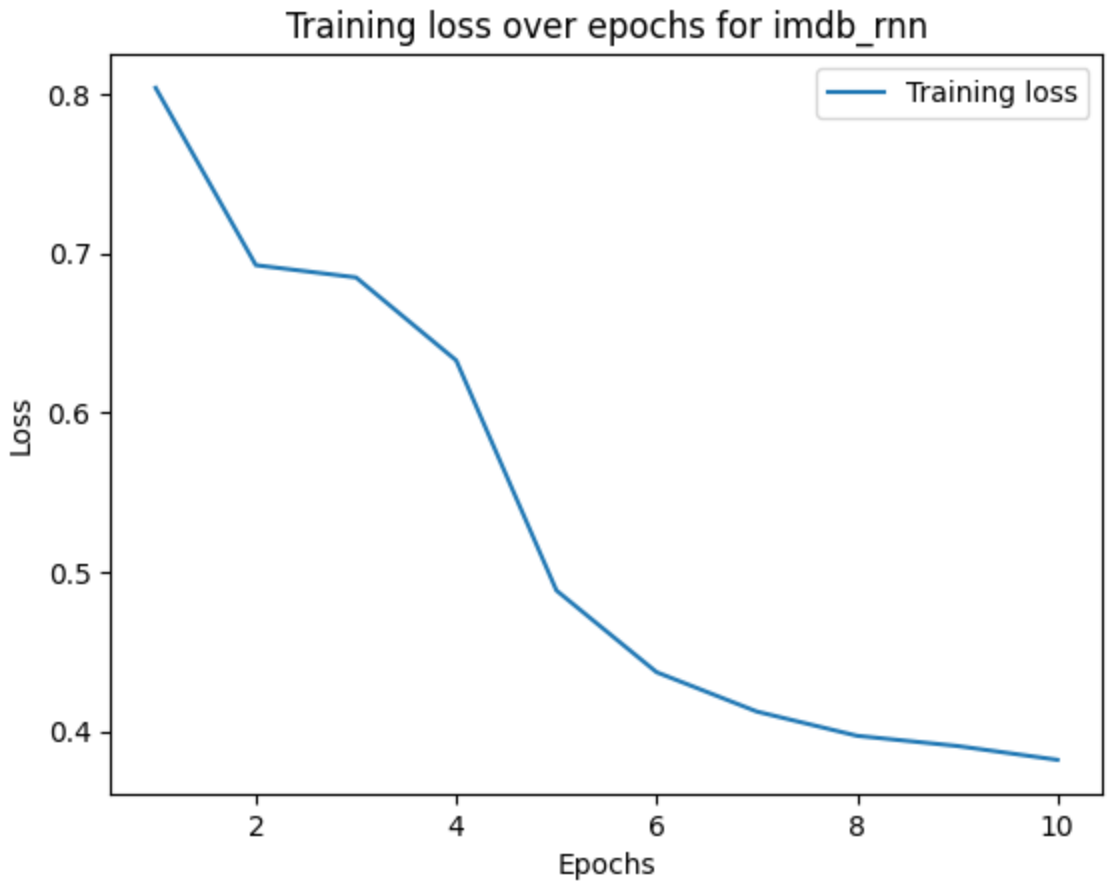
	Imdb	**imdb_rnn**	**semeval**	**semeval_rnn**
Accuracy	0.7554	**0.8011**	0.5008	**0.5654**
Precision	0.7555	**0.8025**	0.4131	**0.5533**
Recall	0.7554	**0.8011**	0.4104	**0.4605**
F1-score	0.7554	**0.8009**	0.3984	**0.4439**

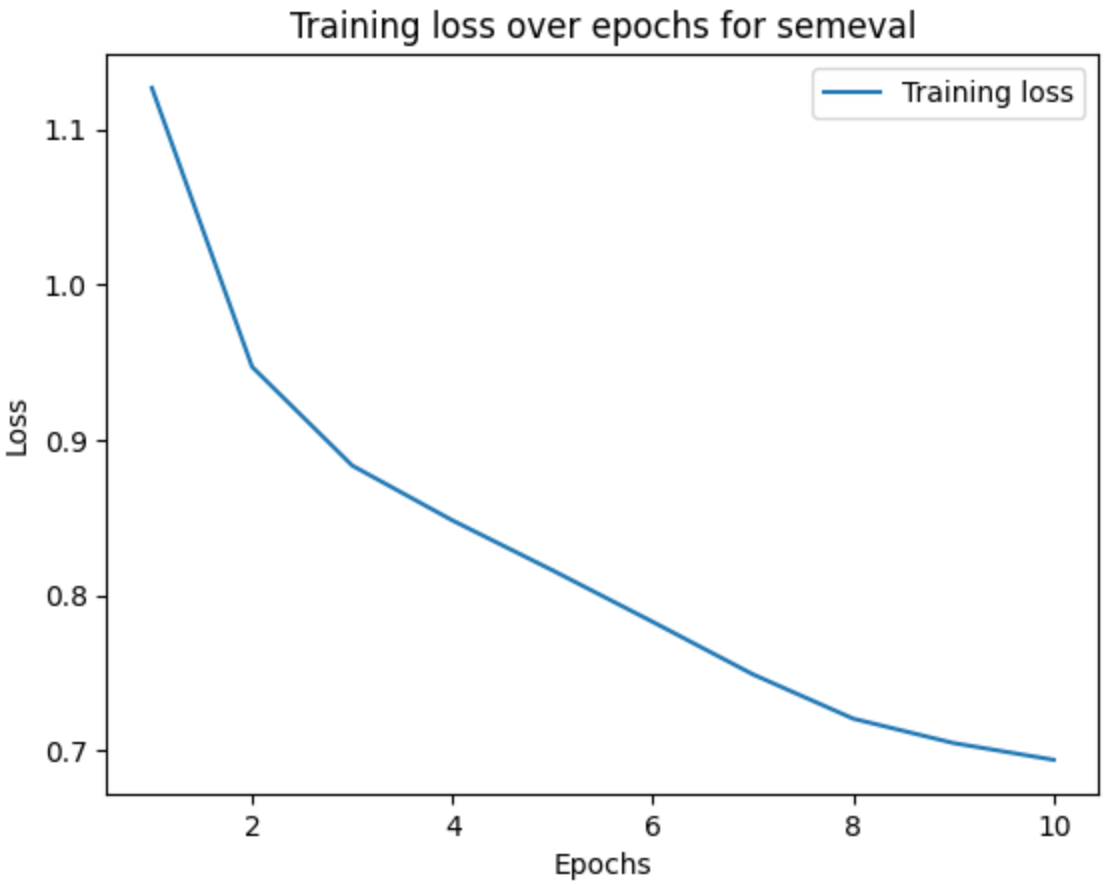
Findings: RNN architecture performed better in both datasets due to recurrency.

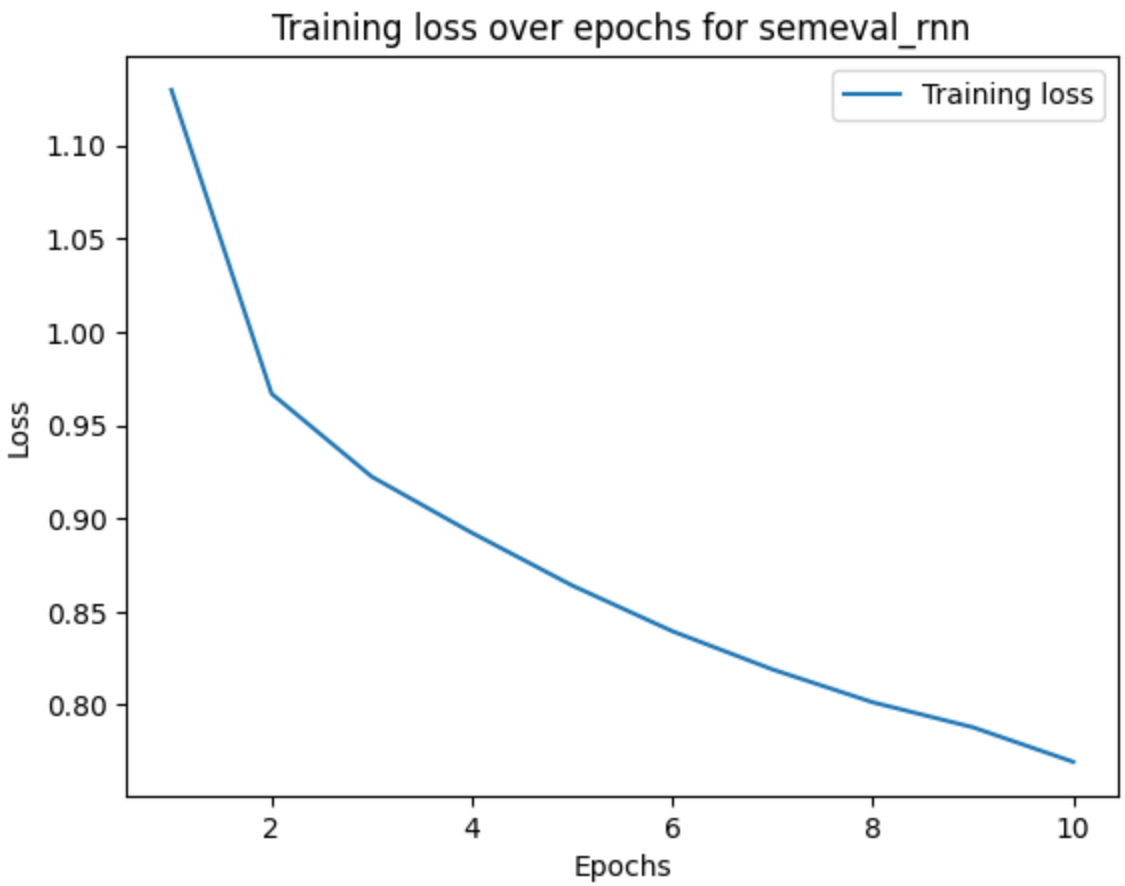
Model Logs

Training loss for four cases

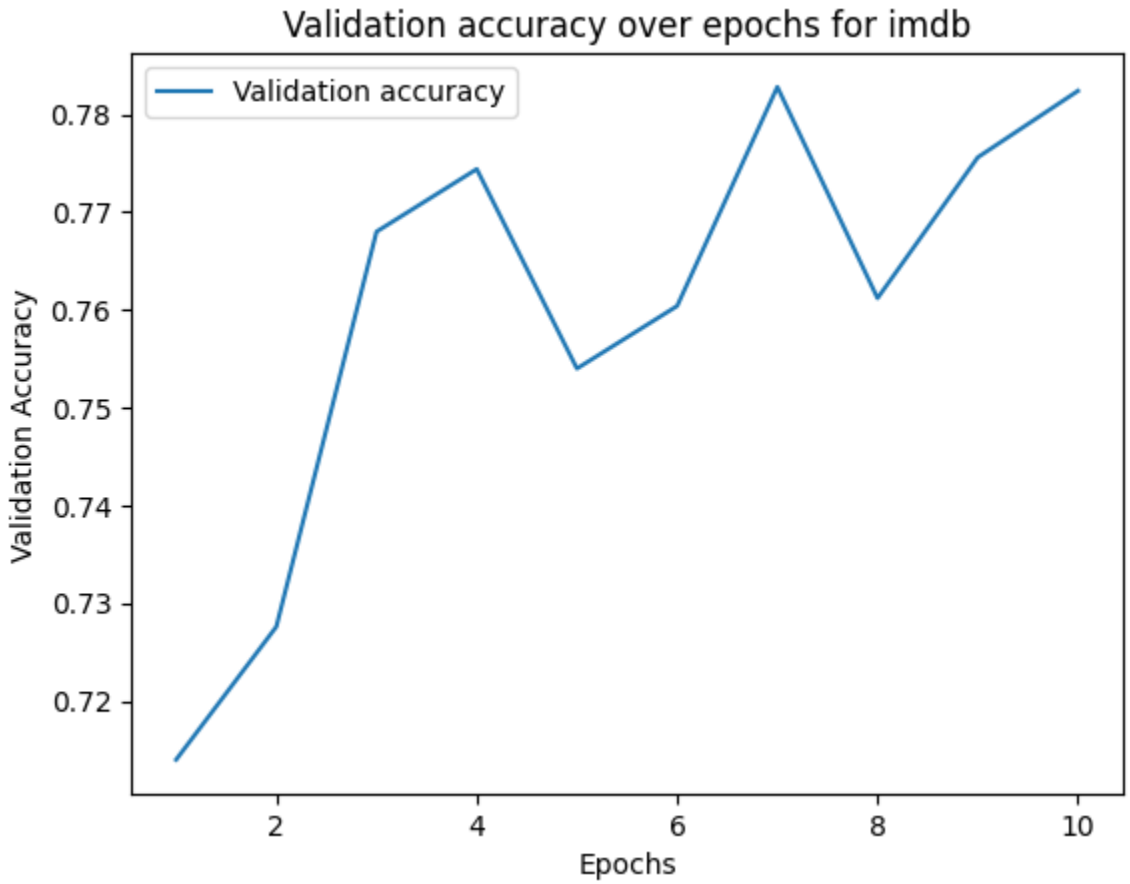



Training loss over epochs for imdb_rnn
Training loss
Loss
Epochs
0.8
0.7
0.6
0.5
0.4
2
4
6
8
10

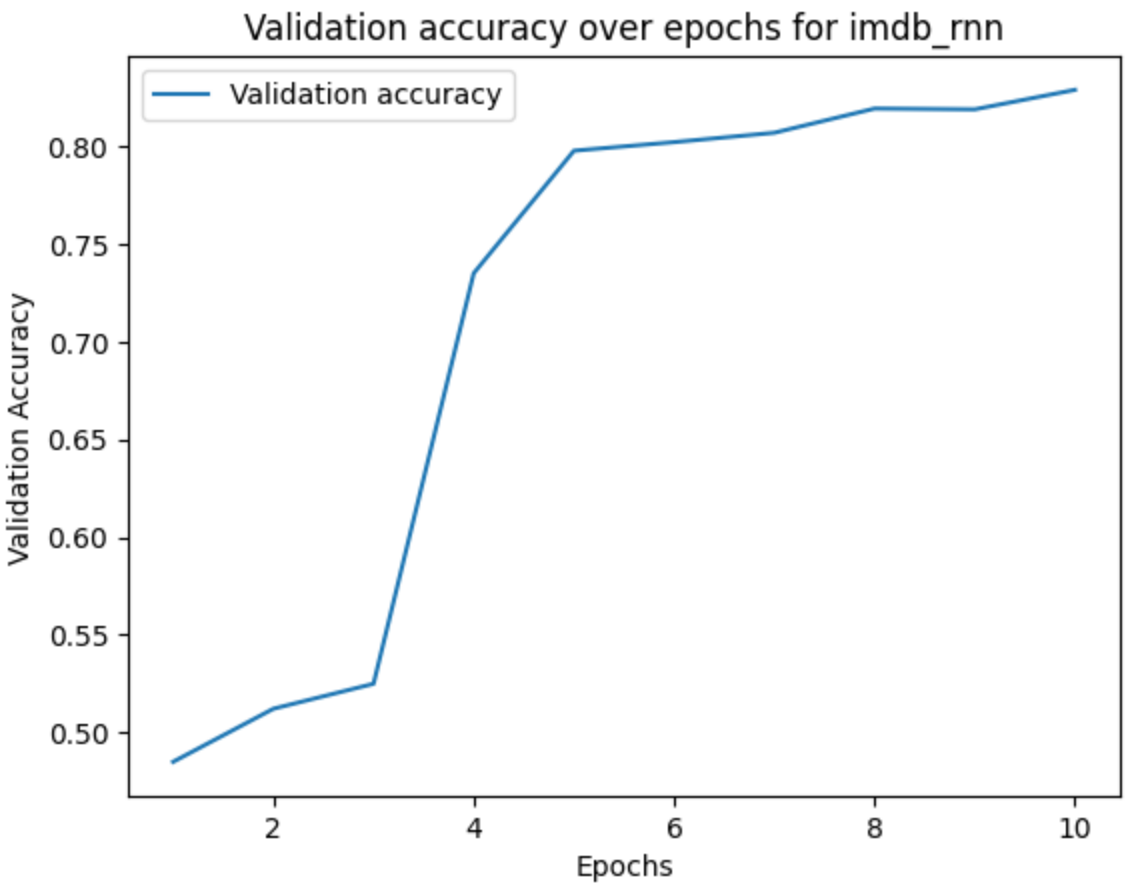

Training loss over epochs for semeval
Training loss
1.1
1.0
0.9
0.8
0.7
Loss
2
4
6
8
10
Epochs

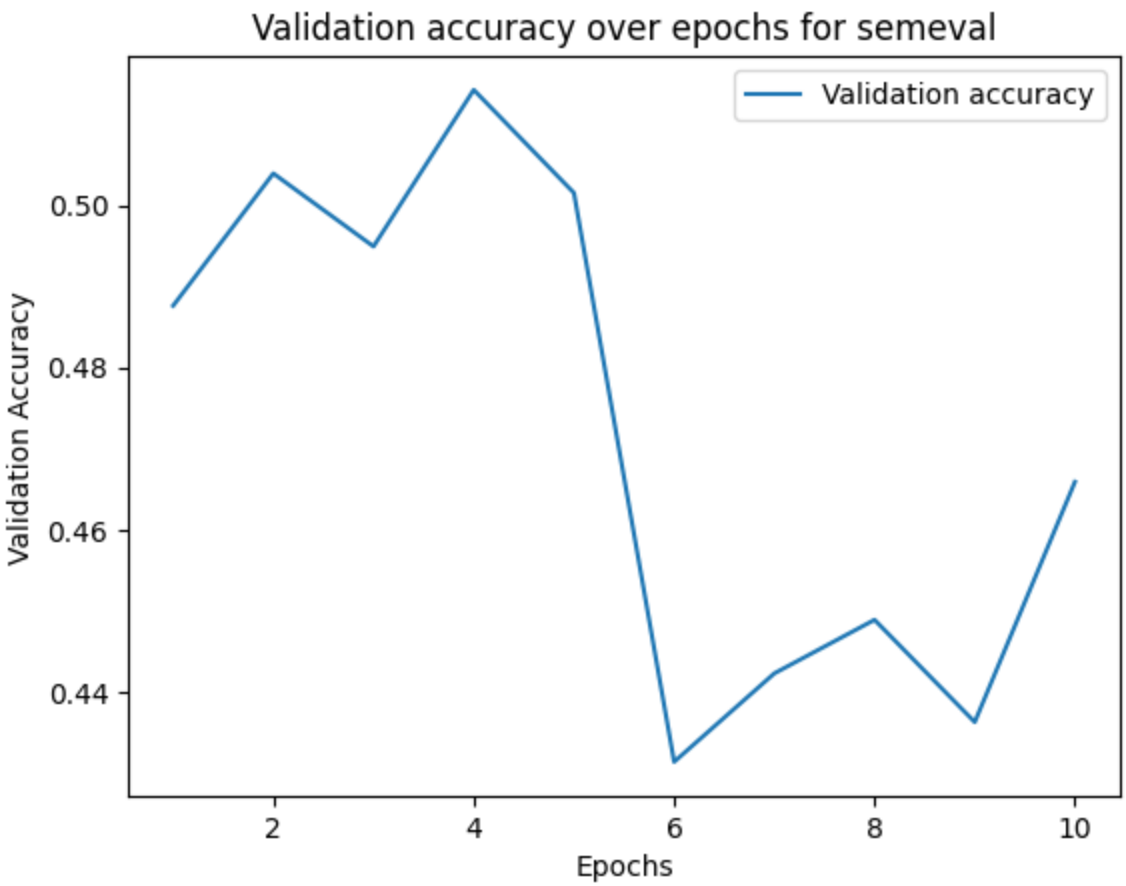

Training loss over epochs for semeval_rnn
Training loss
1.10
1.05
1.00
0.95
0.90
0.85
0.80
Loss
2
4
6
8
10
Epochs

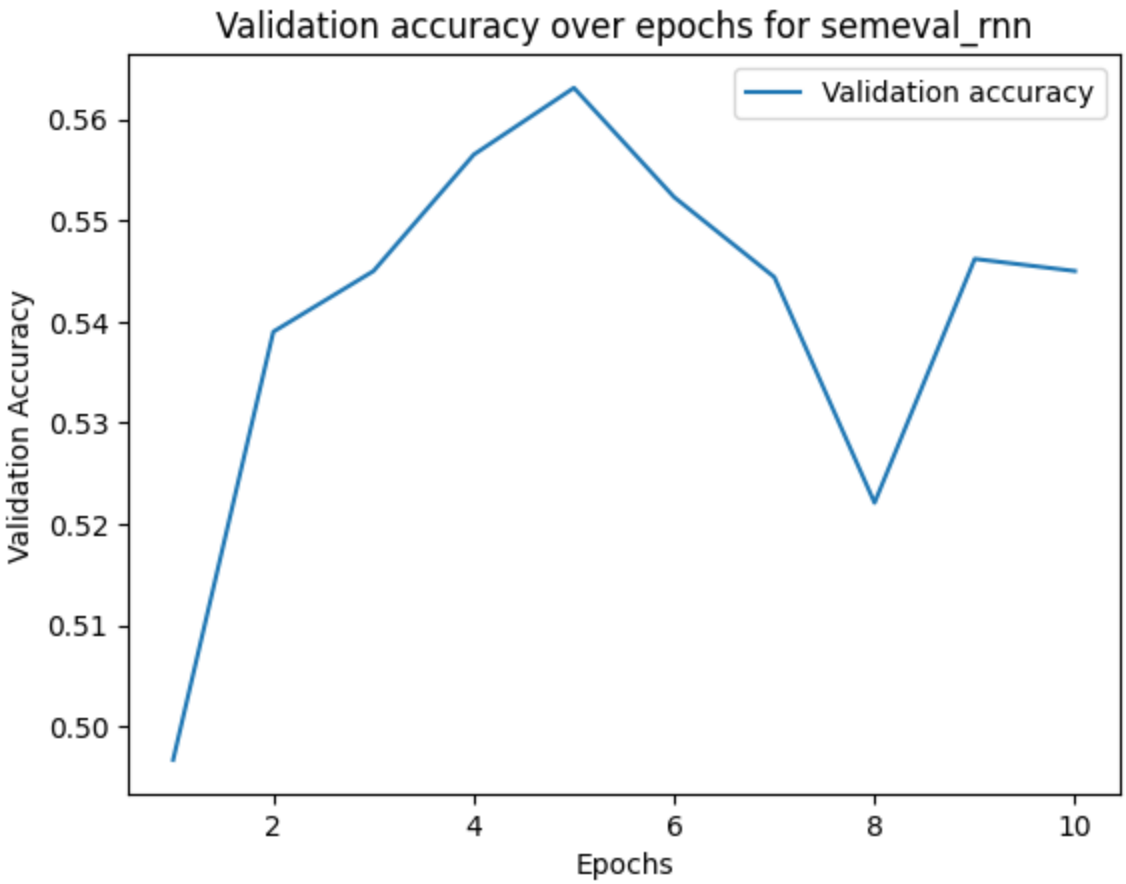
Validation accuracy on four model



s


Validation accuracy over epochs for imdb_rnn
Validation accuracy
Validation Accuracy
0.80
0.75
0.70
0.65
0.60
0.55
0.50
2
4
6
8
10
Epochs


Validation accuracy over epochs for semeval
Validation accuracy
Validation Accuracy
0.50
0.48
0.46
0.44
2
4
6
8
10
Epochs


Validation accuracy over epochs for semeval_rnn
Validation accuracy
Validation Accuracy
0.56
0.55
0.54
0.53
0.52
0.51
0.50
2
4
6
8
10
Epochs