



**AHSANULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY**  
*Department of Computer Science and Engineering*

SOFT COMPUTING LAB  
CSE4238

ASSIGNMENT - 3

---

## Recurrent Neural Network(RNN)

---

Submitted to

**Mr. Nibir Chandra Mandal**  
Lecturer, CSE, AUST

**Sanzana Karim Lora**  
Lecturer, CSE, AUST

Submitted by

**Arafat Ibne Yousuf**      **17.01.04.140**

September 25, 2021

## Contents

<b>1</b>	<b>Data Visualization</b>	<b>1</b>
<b>2</b>	<b>Data pre-processing</b>	<b>1</b>
<b>3</b>	<b>Results</b>	<b>3</b>
<b>4</b>	<b>Prediction</b>	<b>4</b>
<b>5</b>	<b>Git Link</b>	<b>4</b>

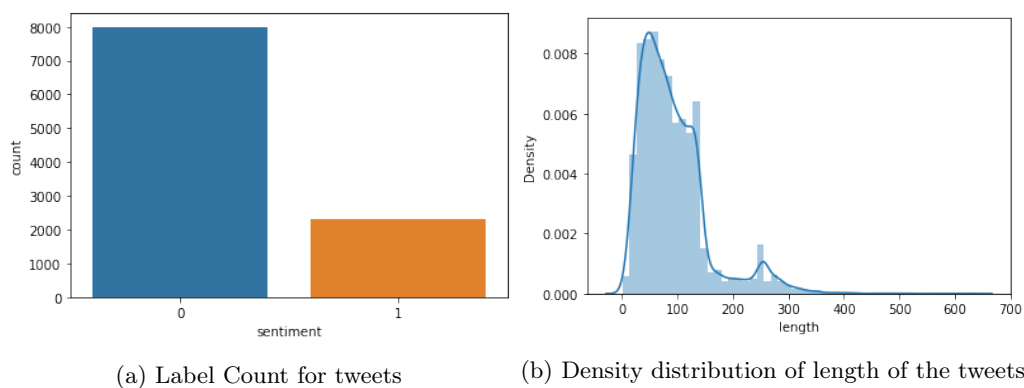
## 1 Data Visualization

The dataset contains two columns and 10315 rows. The first column is the tweets and the second column is sentiment which represents which class the tweet belongs to(0/1).

	tweets	sentiment
0	just had a real good moment. i missssssssss hi...	0
1	is reading manga <a href="http://plurk.com/p/mzp1e">http://plurk.com/p/mzp1e</a>	0
2	@comeagainjen <a href="http://twitpic.com/2y2lx">http://twitpic.com/2y2lx</a> - <a href="http://...">http://...</a>	0
3	@lapcat Need to send 'em to my accountant tomo...	0
4	ADD ME ON MYSPACE!!! <a href="http://myspace.com/LookThunder">myspace.com/LookThunder</a>	0

Figure 1: Data Visualization

From label count of the sentiment we can see that most of the sentiments are labeled as 0 (positive).



(a) Label Count for tweets

(b) Density distribution of length of the tweets

## 2 Data pre-processing

For pre-processing, some steps have been taken such as removing most frequent words, removing of punctuation, removing address id.

```
0    just had a real good moment. i missssssssss hi...
1                                is reading manga
2                                -
3    Need to send 'em to my accountant tomorrow. Od...
4                                ADD ME ON MYSPACE!!!
Name: tweets, dtype: object
```

Figure 3: Address ID Removal

```

0   just had a real good moment i missssssssss him...
1                                   is reading manga
2
3   Need to send em to my accountant tomorrow Oddl...
4                                   ADD ME ON MYSPACE
Name: tweets, dtype: object

```

Figure 4: Punctuation Mark Removal

```

to           3758
the          3448
a            2949
I            2839
and          2574
you          2143
depression   1807
of           1778
my           1705
is           1659
dtype: int64

```

(a) Most Frequent Words

```

Keough       1
snakebite    1
phenomenal   1
Recovering   1
refreshment  1
Roman         1
deffo        1
waves        1
Iver         1
Aluminum     1
dtype: int64

```

(b) Most Rare Words

```

0   just had real good moment i missssssssss him s...
1                                   reading manga
2
3   Need send em accountant tomorrow Oddly wasnt e...
4                                   ADD ME ON MYSPACE
Name: tweets, dtype: object

```

Figure 6: Frequent Words Removal

### 3 Results

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 50, 32)	160000
simple_rnn (SimpleRNN)	(None, 50, 128)	20608
simple_rnn_1 (SimpleRNN)	(None, 50, 64)	12352
simple_rnn_2 (SimpleRNN)	(None, 64)	8256
dense (Dense)	(None, 1)	65

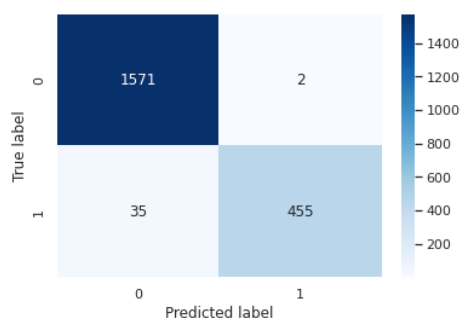
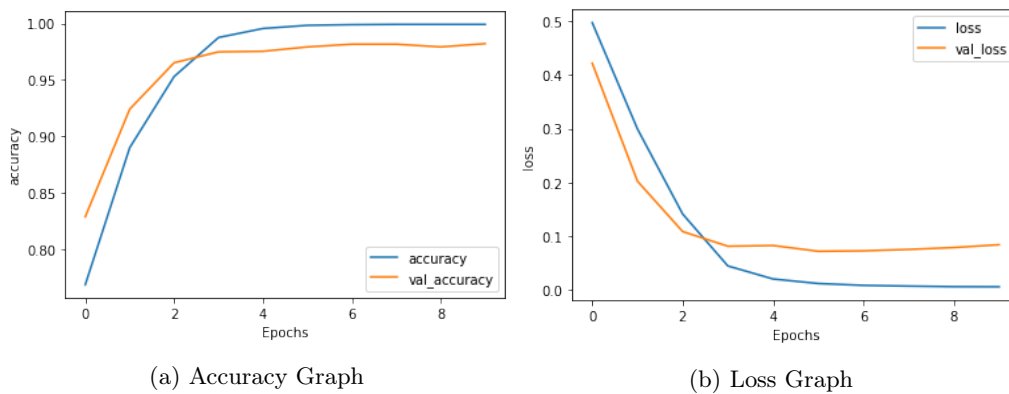
```

Total params: 201,281
Trainable params: 201,281
Non-trainable params: 0

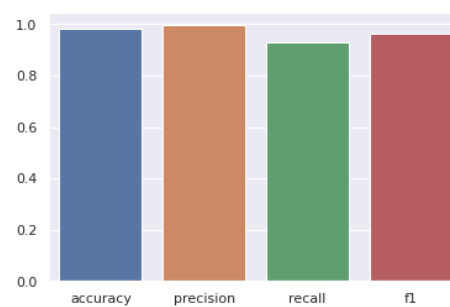
```

Figure 7: Model

The comparison between (accuracy and validation accuracy) and (loss and validation loss) can be visualized by plotting graph which is shown below:



(a) Confusion Matrix



(b) Accuracy, Precision, Recall, F1 Score

The exact value of our performance measurement is given below:

Accuracy	98.21%
Precision	99.56%
Recall	92.86%
F1 score	96.09%

## 4 Prediction

```
1 test_review_neg = "so excited 2day, looking 4ward 2 my long weekend "  
2 predict_sentiment(test_review_neg)  
  
0.000864: Positive sentiment  
1
```

Figure 10: Positive Sentiment Prediction

```
1 test_review_neg = "My mom died last week. I don't have any relatives. I'm alone and it is so hard. Today is the first day without anxiety and depression."  
2 predict_sentiment(test_review_neg)  
  
1.0: Negative sentiment  
0
```

Figure 11: Negative Sentiment Prediction

## 5 Git Link

### Assignment 3