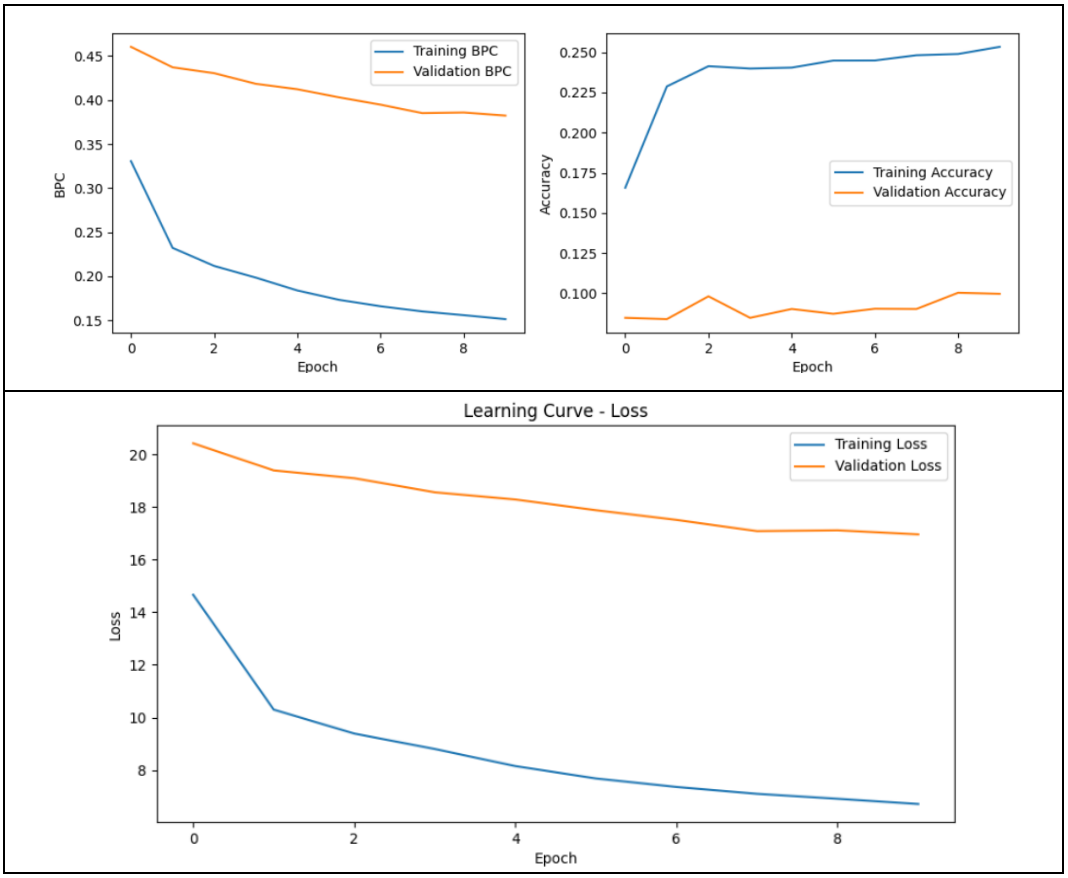
RNN

1.

(1) Network Architecture

```
SimpleRNN(
  (embedding): Embedding(100, 128)
  (rnn): RNN(128, 128, batch_first=True)
  (fc): Linear(in_features=128, out_features=100, bias=True)
)
```

(2) Learning curve



(3) Training error rate

0.2534

(4) Validation error rate

0.0997

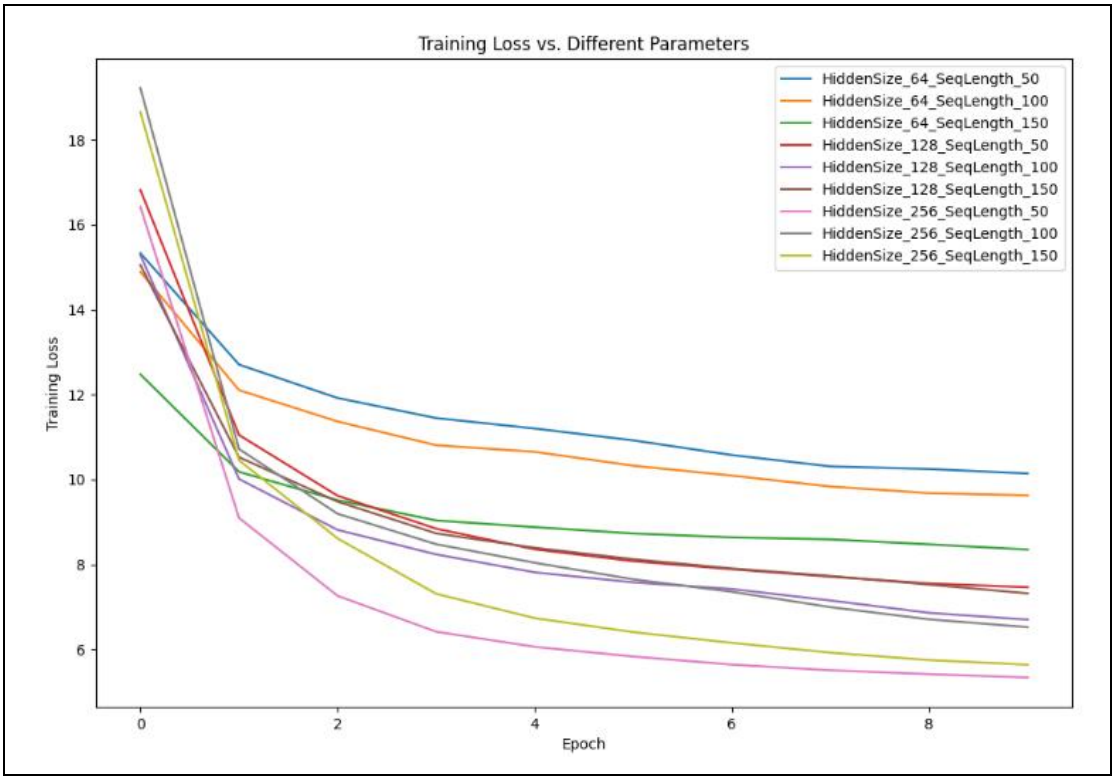
2.

(1)

```
Epoch [1/10], Train Loss: 14.6674, Train Accuracy: 0.1657, Train BPC: 0.3306, Valid Loss: 20.4269, Valid Accuracy: 0
.0848, Valid BPC: 0.4605
Generated Text:
There's no remedy; me me me me me me me me me me me me me me me me me me me me me me me me me me me me me me m
e
Epoch 2/10: 100%|                                   | 29685/29685 [00:27<00:00, 1071.81batch/s, Loss=10.3, Accuracy=0.229]
Epoch [2/10], Train Loss: 10.2989, Train Accuracy: 0.2288, Valid Loss: 19.3970, Valid Accuracy: 0.0840
Epoch 3/10: 100%|                                   | 29685/29685 [00:27<00:00, 1063.38batch/s, Loss=9.39, Accuracy=0.241]
Epoch [3/10], Train Loss: 9.3876, Train Accuracy: 0.2414, Valid Loss: 19.0997, Valid Accuracy: 0.0982
Epoch [3/10], Train Loss: 9.3876, Train Accuracy: 0.2414, Train BPC: 0.2116, Valid Loss: 19.0997, Valid Accuracy: 0.
0982, Valid BPC: 0.4305
Generated Text:
There's no remedy; the the the the the the the the the the the the the the the the the the the the the the the the t
he
Epoch 4/10: 100%|                                   | 29685/29685 [00:27<00:00, 1069.31batch/s, Loss=8.8, Accuracy=0.24]
Epoch [4/10], Train Loss: 8.8001, Train Accuracy: 0.2399, Valid Loss: 18.5649, Valid Accuracy: 0.0848
Epoch 5/10: 100%|                                   | 29685/29685 [00:27<00:00, 1065.31batch/s, Loss=8.15, Accuracy=0.241]
Epoch [5/10], Train Loss: 8.1503, Train Accuracy: 0.2405, Valid Loss: 18.2901, Valid Accuracy: 0.0903
Epoch [5/10], Train Loss: 8.1503, Train Accuracy: 0.2405, Train BPC: 0.1837, Valid Loss: 18.2901, Valid Accuracy: 0.
0903, Valid BPC: 0.4123
Generated Text:
There's no remedy; me me me me me me me me me me me me me me me me me me me me me me me me me me me me me me m
e
Epoch 6/10: 100%|                                   | 29685/29685 [00:27<00:00, 1068.97batch/s, Loss=7.68, Accuracy=0.245]
Epoch [6/10], Train Loss: 7.6801, Train Accuracy: 0.2449, Valid Loss: 17.8834, Valid Accuracy: 0.0873
```

```
Epoch 7/10: 100%|                                   | 29685/29685 [00:27<00:00, 1066.40batch/s, Loss=7.36, Accuracy=0.245]
Epoch [7/10], Train Loss: 7.3559, Train Accuracy: 0.2449, Valid Loss: 17.5139, Valid Accuracy: 0.0904
Epoch [7/10], Train Loss: 7.3559, Train Accuracy: 0.2449, Train BPC: 0.1658, Valid Loss: 17.5139, Valid Accuracy: 0.
0904, Valid BPC: 0.3948
Generated Text:
There's no remedy; an an an an an an an an an an an an an an an an an an an an an an an an an an an an an an an a
n
Epoch 8/10: 100%|                                   | 29685/29685 [00:27<00:00, 1075.48batch/s, Loss=7.1, Accuracy=0.248]
Epoch [8/10], Train Loss: 7.0973, Train Accuracy: 0.2482, Valid Loss: 17.0875, Valid Accuracy: 0.0903
Epoch 9/10: 100%|                                   | 29685/29685 [00:27<00:00, 1073.47batch/s, Loss=6.91, Accuracy=0.249]
Epoch [9/10], Train Loss: 6.9077, Train Accuracy: 0.2490, Valid Loss: 17.1181, Valid Accuracy: 0.1004
Epoch 10/10: 100%|                                  | 29685/29685 [00:27<00:00, 1068.33batch/s, Loss=6.71, Accuracy=0.253]
Epoch [10/10], Train Loss: 6.7079, Train Accuracy: 0.2534, Valid Loss: 16.9634, Valid Accuracy: 0.0997
Epoch [10/10], Train Loss: 6.7079, Train Accuracy: 0.2534, Train BPC: 0.1512, Valid Loss: 16.9634, Valid Accuracy: 0
.0997, Valid BPC: 0.3824
Generated Text:
There's no remedy; the the the the the the the the the the the the the the the the the the the the the the the the t
he
```

3.



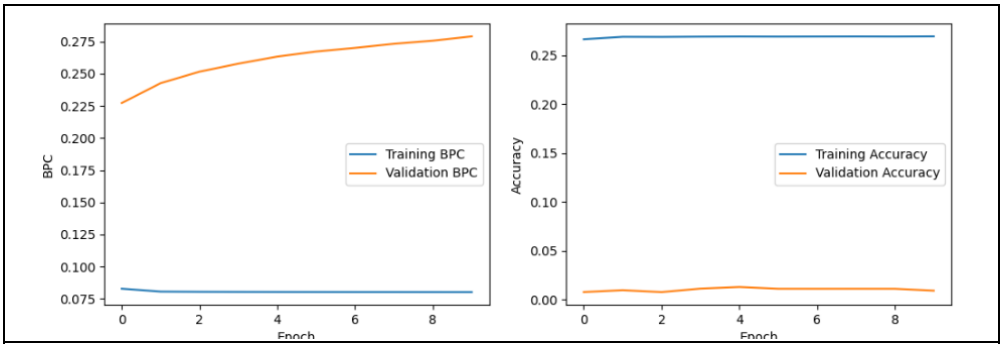
4.

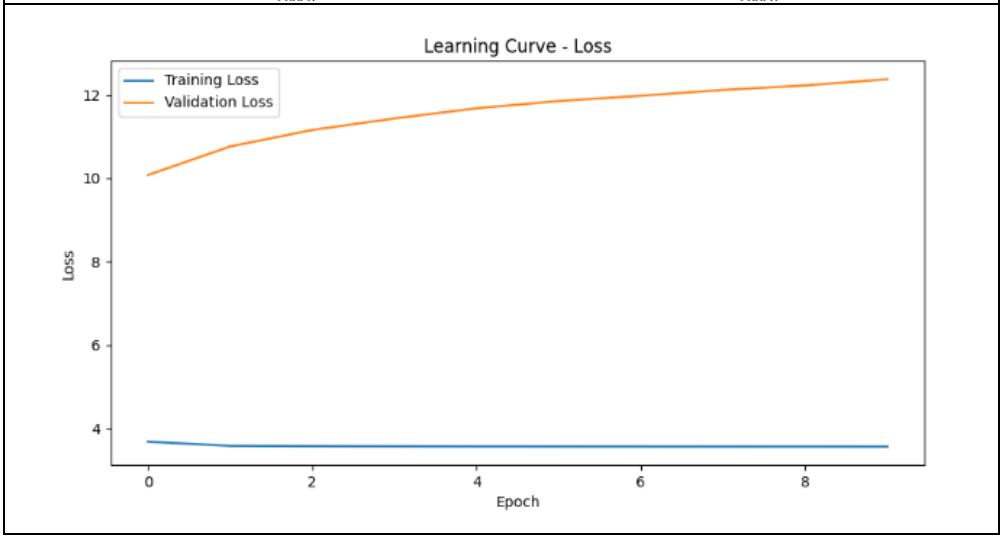
LSTM

4.1

(1) Network Architecture

(2) Learning curve





(3) Training error rate

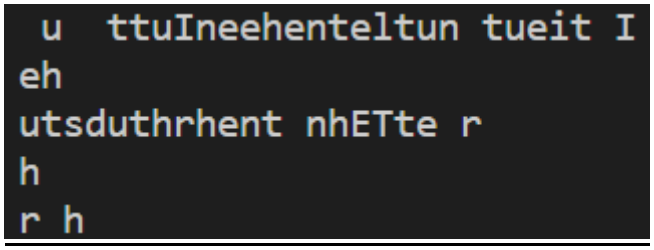
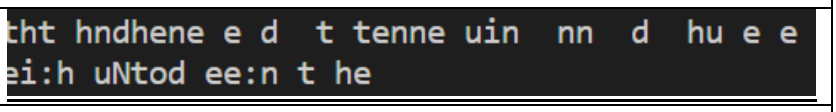
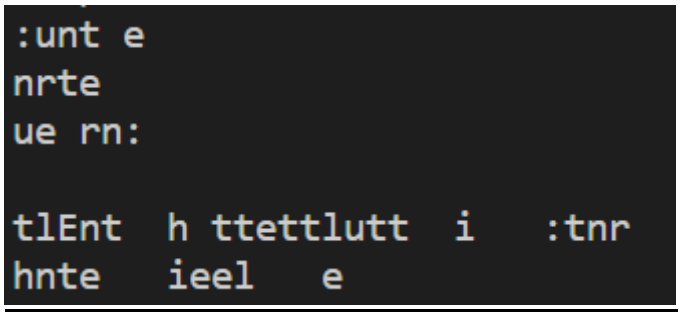
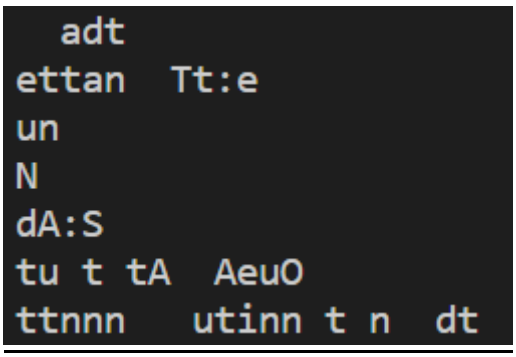
0.2694

(4) Validation error rate

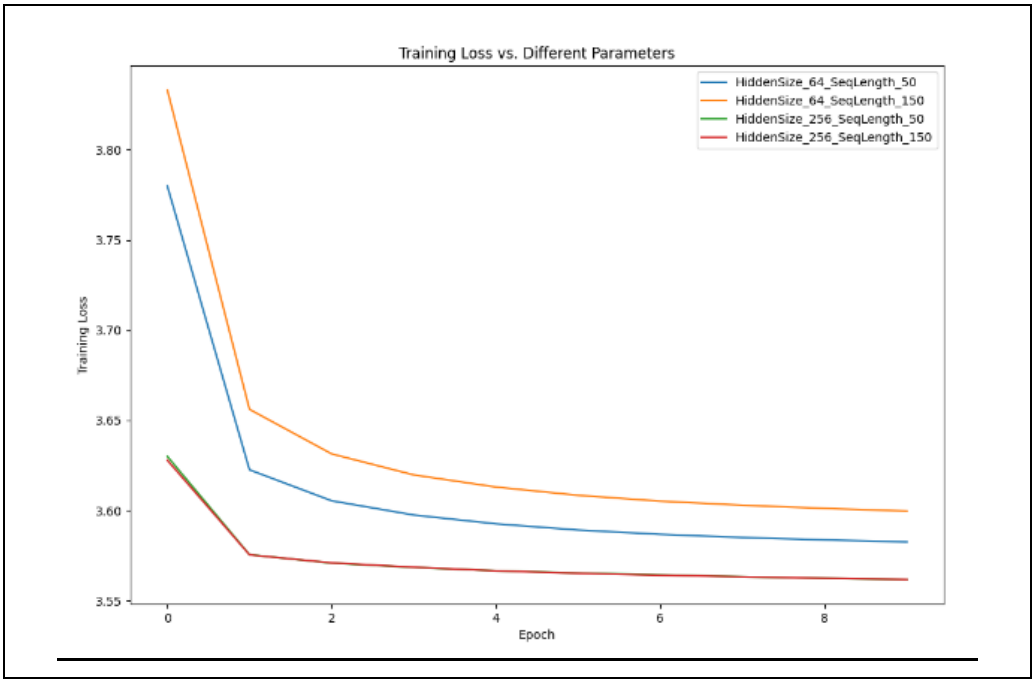
0.0092

4.2

1	hn:ttth te u nnti ntt huo:ht nt e eton tE tt nretu

3	 u ttuIneehenteltun tueit I eh utsduthrhent nhETte r h r h
5	 tht hndhene e d t tenne uin nn d hu e e ei:h uNtod ee:n t he
7	 :unt e nrte ue rn: tlEnt h ttettlutt i :tnr hnte ieel e
10	 adt ettan Tt:e un N dA:S tu t tA AeuO ttnnn utinn t n dt

4.3



Difference

- The performance of LSTM is a little bit higher than RNN. However, because of insufficient training epoch, performances of both of them are not good enough.
- The convergence of LSTM is more faster than RNN.
- The training time of LSTM is more longer than RNN.

5.

Model: RNN

Prime: “JULIET”

```
Generated Text with Priming "JULIET":
JULIEThgrmUd t wistheXETIOml, we a[yoZARB!
AUS:
O---JUSe&cbll!
HENd, me !
hat l A, mThe
Wat B!


She hil AUCAnSem&car ave Bbor, atgre
r a[y OLEar CEk a[yoMI
```