

# Assignment 3: Recurrent Neural Networks for Language Modeling

Device: NVIDIA GeForce RTX 4090 D

## 3. Evaluation and generation

RNN needs 10G memory and LSTM needs 14G memory.

	4	N/A	N/A	3021331	C	...iang/miniconda3/envs/nlp/bin/python	10842MiB	
	5	N/A	N/A	3021575	C	...iang/miniconda3/envs/nlp/bin/python	14008MiB	

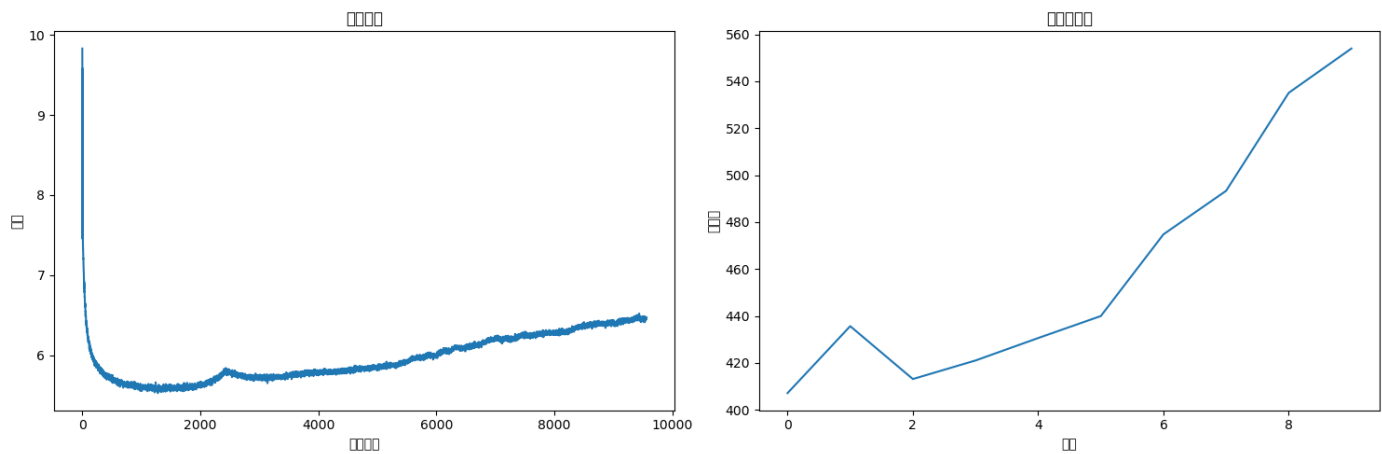
From the training results, LSTM seems to have better performance (loss) in training under same conditions:

```
EMBEDDING_DIM = 300
HIDDEN_DIM = 512
NUM_LAYERS = 2
BATCH_SIZE = 1024
NUM_EPOCHS = 10
LEARNING_RATE = 0.005
SEQUENCE_LENGTH = 35
```

However, Ive noticed that both RNN and LSTM method gained higher Perplexity during training, which against our training purposes.

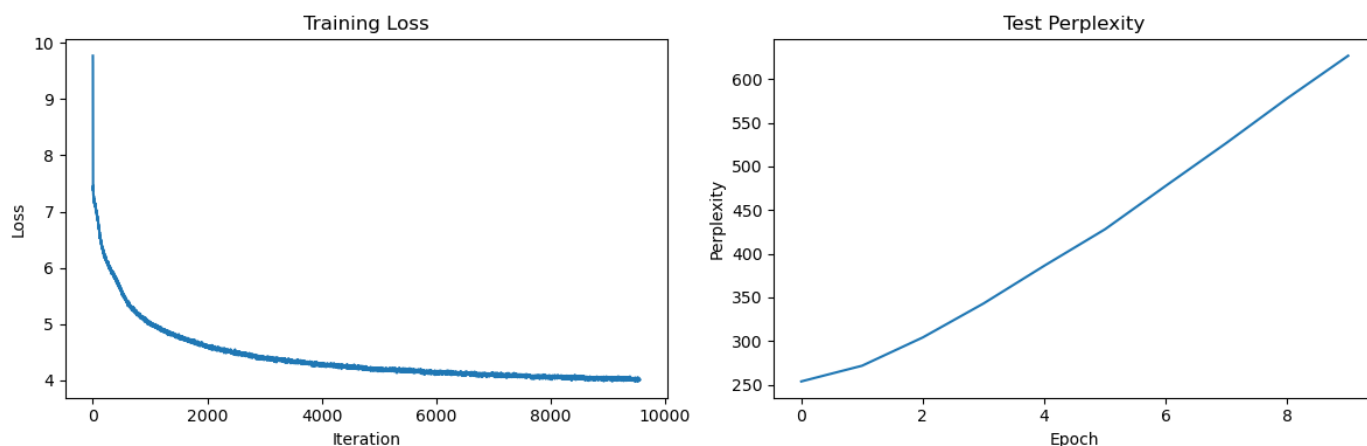
### RNN Loss and Perplexity

Epoch [10/10], Step [900/956], Loss: 6.4591  
Epoch [10/10], Test Perplexity: 553.9734



### LSTM Loss and Perplexity

Epoch: 10/10, Batch: 900/956, Loss: 4.0324  
Epoch: 10/10, Test Perplexity: 626.7841



## RNN Generated Text Samples

### Prefix: Harry looked at

Generated: harry looked at a goblet were horrors checked tonight .he had to out it wall stood walk lunch malfoy .what dont know we could to be invited ?positive is of for are her knowing that hermione happened first any later breathing caught them that the life in the purple looked it right does

### Prefix: The castle was

Generated: the castle was popular but gulped well up okay disappearing and do with a job .it asked sure give great plot held faithful face nor hermione .they looked turning up here with hagrid .what shouldnt this too has conjured a smile on face was sure because harry and voldemort .but black once at

### Prefix: Hermione said

Generated: hermione said a desk on having gone and made for full of in the window of no parents i minerva had reminded him was hermione .something good .look you dont he was the is bizarrely liechtenstein loved betterlooking that learn being enormous way from his head still people with moodys coincidence

### Prefix: Ron couldn't

Generated: ron couldn't lies himself want over to shortly .harry morning them over them be her by his down on harry i stop you you ?said halloween stepping from the call of magic it had certain lawn might her way inside again told .an kind little chirruping he dilemma only subject of term

### Prefix: Dumbledore smiled

Generated: dumbledore smiled classes at unwanted off asking your ?harry face and he fingers everywhere the meetings investigations dwindling and his usual where .the glasses in harry .he check noticed he would come said harry who ill .and it hard loony good to room .as all albus professor mean i cant help

## LSTM Generated Text Samples

### 前綴: Harry looked at

生成: harry looked at him .your father works at least a couple of toadstools .harry there was no difference he dipped on his trainers as he shrank to the front of his empty cereal bowl on the wall and then became low he saw his eyes dart in the golden light ive changed to

前缀: **The castle was**

生成: the castle was called .the doorbell rang there was a cold man whose back had a permanently bloody face .and the dormitory door opened and hagrid looked down to see that she was very angry and the firebolt had vanished from the end of a and a pair of old headmasters

前缀: **Hermione said**

生成: hermione said by the last thing in the subject .hermione had definitely packed off out in her dormitory .harry was halfway up the corridor on which there was no snow from fellow the room harry had ever seen .he reached the entrance hall after breakfast at the end of the lesson .mr

前缀: **Ron couldn't**

生成: ron couldn't with the luggage of muggle four .ah !said mrs weasley tartly .oh no she said .its a i said youd come back ter moaning myrtles .and if i suppose we got monday afternoons work still ron how did they get the right thing ?did i lick the train

前缀: **Dumbledore smiled**

生成: dumbledore smiled .not enough .he ive done it to my face with me .james was the two there was the last time a death eater was there he was slowly starting the same .hermione had been taken to each of them from professor trelawney in another corner mr dursley pointed at harry

It seems that lower Perplexity reflects to a better generation.

## Why Perplexity Increases During Training

We found that the initial perplexity is the lowest and while training loss decreased, test perplexity increased instead.

### 1. Overfitting

- The model becomes too specialized to the training data
- It loses its ability to generalize to unseen data
- This leads to higher perplexity on the test set

### 2. Learning Rate Issues

- Learning rate might be too high
- Model overshoots the optimal parameters
- Results in unstable training and increasing perplexity

### 3. Data Distribution Mismatch

- Training and test sets might have different distributions
- Model performs well on training data but poorly on test data
- This discrepancy manifests as increasing perplexity

## 4. Compared randomly initialized and pretrained model

正在下载GloVe预训练词向量...

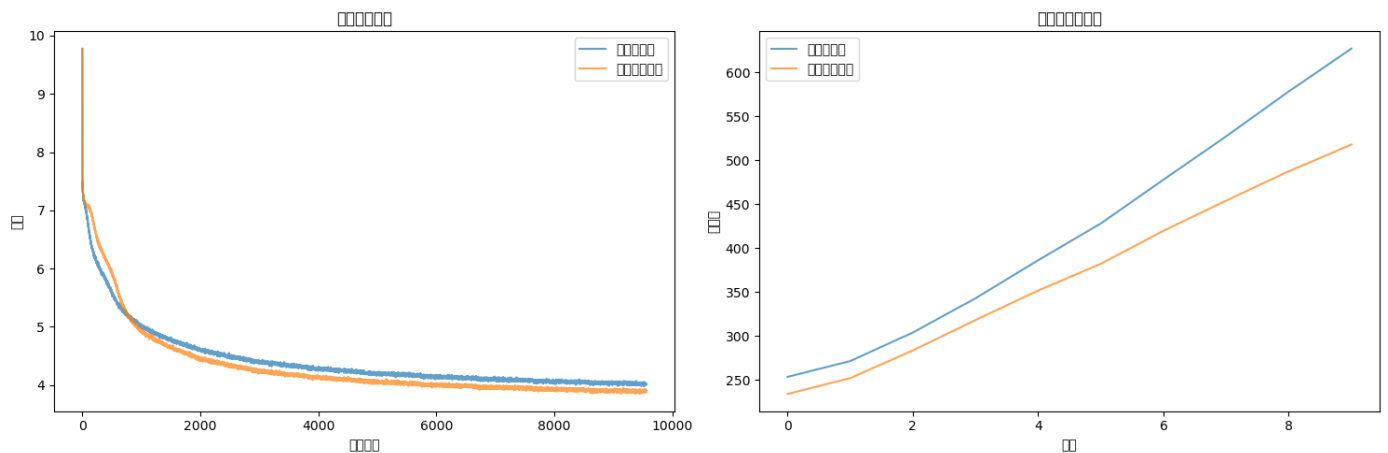
在预训练词向量中找到的词数: 13727/17544

Epoch: 10/10, Batch: 800/956, Loss: 3.9077

Epoch: 10/10, Batch: 900/956, Loss: 3.9196

Epoch: 10/10, Test Perplexity: 517.7275

### Pretrained model Loss and Perplexity (blue random yellow retrained)



### 最终测试集困惑度对比

随机初始化: 626.78 预训练词向量: 517.73

预训练词向量模型表现更好, 保存为best\_model\_pretrained.pth