Report
RNN-Based Text Generation

Lyle He

Contents

1 Part1: RNN-Based Text Generation 2
1.1 RNN Model Implementation 2
1.2 Dataset Preparation 2
1.3 Training (10 Points) 3
1.4 Text Generation 4
1.5 Analysis 5

1 Part1: RNN-Based Text Generation

1.1 RNN Model Implementation

According to the assignment requirement, I have implemented a basic RNN model and a LSTM model with different other components, such as dropout, fully connected layers, and activation function. The model architecture is shown in the below.

1. Basic RNN Model

```
BasicRNN(
(embedding): Embedding(7014, 128)
(rnn): RNN(128, 128, batch_first=True, dropout=0.3)
(fc1): Linear(in_features=128, out_features=7014, bias=True)
(fc2): Linear(in_features=7014, out_features=7014, bias=True)
(dropout): Dropout(p=0.3, inplace=False)
)
```

2. LSTM Model

```
LSTMModel(
(embedding): Embedding(7014, 128)
(lstm): LSTM(128, 128, batch_first=True)
(fc1): Linear(in_features=128, out_features=7014, bias=True)
(fc2): Linear(in_features=7014, out_features=7014, bias=True)
(dropout): Dropout(p=0.3, inplace=False)
)
```

1.2 Dataset Preparation

I choosed a book named **Frankenstein** by Mary Shelley as the dataset for this task. The book is in the public domain and can be downloaded from the Project Gutenberg website[4]. The book is a novel and contains 7014 words.

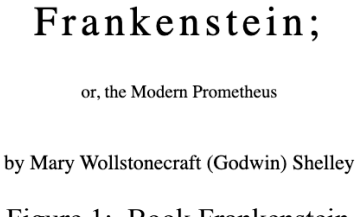

Figure 1: Book Frankenstein

Preprocessing:

1. I tokenized the text into a long string
2. I used a slide window with a size of 100 to slide through the text and get the sequences.
3. converted them into numerical format suitable for training the RNN model.

The result is shown in the below.

index: 0 [‘you’, ‘will’, ‘rejoice’, ‘to’, ‘hear’, ‘that’, ‘no’, ‘disaster’, ‘has’, ‘accompanied’] …
index: 1 [‘will’, ‘rejoice’, ‘to’, ‘hear’, ‘that’, ‘no’, ‘disaster’, ‘has’, ‘accompanied’, ‘the’] …
index: 2 [‘rejoice’, ‘to’, ‘hear’, ‘that’, ‘no’, ‘disaster’, ‘has’, ‘accompanied’, ‘the’] …
index: 3 [‘to’, ‘hear’, ‘that’, ‘no’, ‘disaster’, ‘has’, ‘accompanied’, ‘the’, ‘of’] …
index: 4 [‘hear’, ‘that’, ‘no’, ‘disaster’, ‘has’, ‘accompanied’, ‘the’, ‘commencement’, ‘of’, ‘an’] …
index: 5 [‘that’, ‘no’, ‘disaster’, ‘has’, ‘accompanied’, ‘the’, ‘commencement’, ‘of’, ‘an’, ‘enterprise’] …

```
print("word vocabulary lenght", len(word_to_index))
step = 0
for i in range(len(tokens) - window_size):
    if step > 5:
        break
    print(f"index: {step}", tokens[i:i+window_size-1][:10], "...")
    step += 1
```

```
word vocabulary lenght 7014
index: 0 ['you', 'will', 'rejoice', 'to', 'hear', 'that', 'no', 'disaster', 'has', 'accompanied'] ...
index: 1 ['will', 'rejoice', 'to', 'hear', 'that', 'no', 'disaster', 'has', 'accompanied', 'the'] ...
index: 2 ['rejoice', 'to', 'hear', 'that', 'no', 'disaster', 'has', 'accompanied', 'the', 'commencement'] ...
index: 3 ['to', 'hear', 'that', 'no', 'disaster', 'has', 'accompanied', 'the', 'commencement', 'of'] ...
index: 4 ['hear', 'that', 'no', 'disaster', 'has', 'accompanied', 'the', 'commencement', 'of', 'an'] ...
index: 5 ['that', 'no', 'disaster', 'has', 'accompanied', 'the', 'commencement', 'of', 'an', 'enterprise'] ...
```

Figure 2: Dataset Preparation

1.3 Training (10 Points)

According to the assignment requirement, I have trained the basic RNN model and the LSTM model with different hyperparameters and the best hyperparameters are shown in the below table.

Hyperparameters	Value
Model	Basic RNN
RNN hidden size	128
Learning rate	0.001
full connection layer units	128
Epochs	10
Optimizer	Adam
Loss function	cross entropy loss

Hyperparameters	Value
Model	LSTM
LSM hidden size	128
Learning rate	0.001
full connection layer units	128
Epochs	10
Optimizer	Adam
Loss function	cross entropy loss

The training process is shown below.

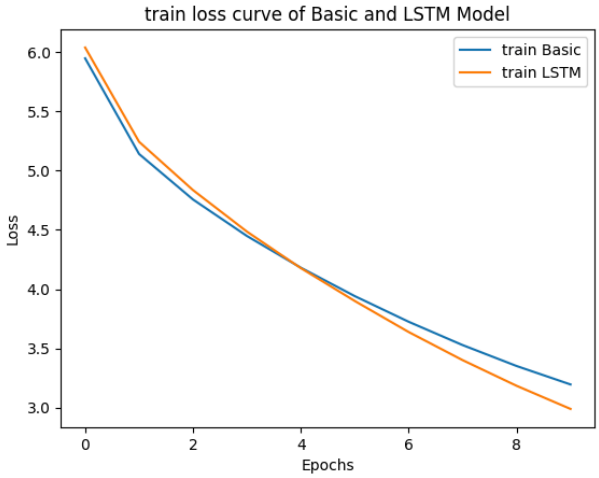

Figure 3: Dataset Preparation

1.4 Text Generation

As required in the assignment, I have implemented a Basic RNN model and a LSTM model with different other components, such as dropout, fully connected layers, and activation function shown in the previous chapter. In addition, I have also implemented a text generation function to generate text according to the assignment requirement. In the below table, I show the prompt text and the generated 100 words text from the RNN model and the LSTM model.

model	prompt text	generated text
Best Basic RNN model	To be or not to be	to be or not to be distinguished from land ; i should not intercept the spirits that preside over thee , sank at sunset to have been shown early and death . yet i can not inspire me with prey , and i must create another like the happiness of ice , and conceived . i have longed to come in the traveller of the leaves . i am interrupted by the causes of the half made me , i have no land ; i have a momentary upon me , and i am yet dizzy with the accumulation . i had cast off myself in
Best LSTM RNN model	To be or not to be	to be or not to be the most alarming symptoms that it was conceived . when i arrived i look saw the spot where i reflected , i was unable ; i am solitary and shone ; but the only renders a care was northeast and my portion of his companions , what do you to seek . what had now was he ! generous life , when i shall be in what truth , then , the exertion , which had fallen in hope . yet do i say this ; he pressed on me ; his victim is forcible and touching ; nor can
Best Basic RNN model	You are so	you are so eminently ? i have wandered one beneficial ; but i concealed my return . years will pass but that arrived at the whole deep and contemplated with ice , yet i am not so miserable , and i shall perish in these wild and enthusiastic

		man ! but i have given the images of the country ate , but i am no coward an obstacle ; i disencumbered the whale fishers on several expeditions ; and when i momentarily expect the assassin of my history , and i hastily wiped before me in the north pacific , and conceived the
Best LSTM RNN model	You are so	you are so ardently desire to you could . heaven forbid ! how would reason an end to him , but he eluded my eyes sparkled , and the wind arose ; a feverish joy is not that you give me if i have undergone now not feel more than a great and sudden world ; but the ice cracked round me and by his bed , and in spite of the walls seized upon my soul , and his eyes closed fancy , and , with all the feelings which arrived to me . yet i am going to unexplored regions ,
Best Basic RNN model	How do I find this book	how do i find this book are self but being struck by the feeling that he should pursue me to the shame of the seas rather look in frankenstein , seem to ashes up in the foe . i had cast my agitation ; wrap long and difficult voyage , the emergencies of which had before gave him to stay . urged thus far and failed ? you know you that hurried you always replenished of exquisite weeks while regarding me that i had invoked to london , september 2d . the sledge was still visible , margaret , what comment can i find rest but
Best LSTM RNN model	How do I find this book	how do i find this book in these rambles for whom i have often engaged my despair . unable to destroy you so long confine to my journey from the mainland where i was concealed , and on her lips , sleeping that night ; this one has not not so rise within me . be happy , dear victor , how can i answer this letter , but nothing before my guiding spirit still been in mine . but he showed close me ; he asked to him and be swallowed up ; i am when so strange that i hoped that i should ask

1.5 Analysis

As we can see from the above table and the loss curve, the LSTM model has a better performance than the basic RNN model. The generated text from the LSTM model is more coherent and meaningful than the basic RNN model. The LSTM model can capture the long-term dependencies and the gradient vanishing problem of the basic RNN model. The generated text from the basic RNN model is not coherent and meaningful. The LSTM model can generate text that is more similar to the prompt text and the generated text from the basic RNN model is not similar to the prompt text.