

APRIL 9, 2020

Text Generation with RNNs

Homework Assignment by prof.Huaque



SANGJEE DONDRUB
sangjeedondrub@live.com
QINGHAI NORMAL UNIVERSITY

Contents

1	Github	1
2	Environment	1
3	Data	1
4	Models	2
5	Train and evaluate	2

1 Github

<https://github.com/sangjeedondrub/text-gen-homework>

2 Environment

```
1 || pip install --upgrade tensorflow
```

```
1 || import tensorflow as tf
2 ||
3 || print(tf.__version__)
```

2.1.0

3 Data

```
1 || head -n 10 lu-drub-gong-gyan.txt
```

དབུ་མའི་ཟབ་གནད་སྤྲོད་པའི་ལེགས་བཤད་ལྟ་སྐྱབ་དགོངས་རྒྱན་ཞེས་བྱ་བ་བཞུགས་སོ།
དགེ་འདུན་ཆོས་འཕེལ།
བདུད་ཀྱི་མཆོན་ཆ་རྩལ་མོ་དག་ལ་འཇམ་མཉེན་མེ་ཏོག་ལན་དུ་སྦྱིན།
ལྷ་མ་བྱིན་ལོ་ལྟོགས་དོ་འཕངས་ཆེ་མི་སྤྲིའི་བརྒྱལ་ཞུགས་དང་དུ་བྲངས།
རང་གི་དག་ལ་སྤང་མིག་ཙམ་ཡང་བལྟ་བར་མི་ལྷན་ཤུགས་སྤྲོད་སྤྲོད།
འཇིགས་རུང་འཁོར་བའི་དག་ཆེན་སྦྱོབས་པའི་གྲོགས་སུ་ཤེས་ལྷན་སུ་ཡིས་བཀྱར།
དགེ་ལེགས་ཞི་བསེལ་བདུད་ཅིའི་འབྱུང་གནས་ཐར་པའི་ལམ།
འདུན་པས་ཕྱ་ཞིབ་མ་ཞོར་མཆོན་པ་སྤྲོད་པའི་མིག།
ཆོས་རྒྱ་སྤྲོད་པའི་སངས་མཆོག་ནི་འདི་འོ་ཞེས།
འཕེལ་འགྲིབ་མེད་པར་རིགས་སྤྲོད་ཆོགས་ཀྱིས་རྟག་དུ་གྱས།
གང་གི་མཆོད་པའི་ཉི་མའི་དཀྱིལ་འཁོར་ཀྱན་དུ་བཟང་པོའི་མཁའ་དབྱིངས་ནས།

Total number of lines

```
,##+begin_src sh :exports both :results output :dir ../ wc -l lu-drub-gong-gyan.txt ##+end_src
```

918 lu-drub-gong-gyan.txt

Table 1. As Character

Total number of characters	81674
Total vocab	65
Total Patterns	81664

Table 2. As Syllable

Total number of syllables	21755
Total vocab	1189
Totoal pattern	21752

4 Models

Character model

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(64, None, 256)	16640
gru_1 (GRU)	(64, None, 1024)	3938304
dense_1 (Dense)	(64, None, 65)	66625
Total params: 4,021,569		
Trainable params: 4,021,569		
Non-trainable params: 0		

Syllable model

Layer (type)	Output Shape	Param #
embedding_4 (Embedding)	(1, None, 256)	304384
gru_4 (GRU)	(1, None, 1024)	3938304
dense_4 (Dense)	(1, None, 1189)	1218725
Total params: 5,461,413		
Trainable params: 5,461,413		
Non-trainable params: 0		

5 Train and evaluate

```
1 | git clone https://github.com/sangjeedondrub/text-gen-homework --depth=1
2 | cd text-gen-homework
```

Train character-level model

```
1 | python text-gen.py
```

Train and evaluate syllable-level model

```
1 | python text-gen.py --use_syllable
```

The evaluation results will be saved to `sample.syllable.txt` and `sample.char.txt` files