Recipe Generation using Recurrent Neural Network (RNN)

Overview

In this experiment we will use character level Recurrent Neural Network (RNN) to generate cooking recipes. We will use Tensorflow 2.0 and Keras to implement the RNN model. Dataset used in this experiment is a collection of recipes from different cuisines in JSON format.

We will use Keras RNN utils for creating mini-batches for training. Keras RNN utilities are very convenient to work with. We will also need to learn longer sequences and pad short in order to make all inputs of same length. Longer sequences will help RNN to learn longer dependencies and also to avoid unrolling RNN for too many steps.

For this experiment we will use Tensorflow 2.0 and Keras 2.3.1.

Note: The recipes in this notebook are generated just for fun and for learning purposes. The recipes are not safe to eat or cook!

Importing dependencies

```
%tensorflow_version 2.x
import os
import numpy as np
import pandas as pd
import tensorflow as tf
import tqdm
import numpy
import sys
import collections
import time
import functools
import pickle

# to save models after training
from tensorflow.keras.callbacks import ModelCheckpoint
print("TensorFlow Version:", tf.__version__)
print("Keras Version:", tf.keras.__version__)
!pip install -q h5py pyyaml
```

```
TensorFlow Version: 2.0.0
Keras Version: 2.3.1
Collecting h5py
  Downloaded https://files.pythonhosted.org/packages/1d/03/2a64983a36655c4ed49840a6962d8382bf9f2c0bb4e7cd64d0f14ab8a6f0/h5py-2.10.0-cp36-cp36m-manylinux1_x86_64.whl
Collecting pyyaml
  Downloaded https://files.pythonhosted.org/packages/e3/e8/b3212641ee2718d556df0f23f78de8303f068fe29cdd505831b67491888f/PyYAML-5.3.1.tar.gz
Building wheels for pyyaml: started
  Building wheels for pyyaml: finished with status 'done'
  Stored in directory: /root/.cache/pip/wheels/ad/dc/cd/f7fe6f693b14b4e9e8f8b8b2ff8d566f380f7f6f5bb2e010f1
Successfully built pyyaml
Installing collected packages: h5py, pyyaml
Successfully installed h5py-2.10.0 pyyaml-5.3.1
```

Exploring dataset

Let's go through several available dataset and explore them and work out. One of the requirement we have this dataset is need it in form of recipes but we only a lot of ingredients but also a cooking method to make a recipe. We will need to see if we can find such dataset.

- Recipe Ingredients Dataset dataset have ingredients properties
- Recipe1M+ dataset registration is required
- Food.com (Kaggle) dataset seems to be a good candidate dataset. It would be nice to have ratings
- Recipe Box (100,000 recipes with ingredients properties, good

Loading the dataset

Let's download the dataset using Kaggle cmdline tool. Using `!pip install kaggle` utility is convenient because it prompts you to upload your `kaggle.json` file so that it can download dataset from Kaggle. Following commands will download the dataset to the current working directory and unpack it. We will use `unzip` and `rm` commands to deal with the compressed files.

```
!pip install -q kaggle
!mkdir ~/.kaggle
!cp kaggle.json ~/.kaggle/
!ls ~/.kaggle
!ls -lha kaggle.json
!kaggle datasets download -d shuyangli94/food-com-recipes-and-user-interactions
!unzip food-com-recipes-and-user-interactions.zip
!rm interactions_test.csv interactions_validation.csv RAW_interactions.csv
!ls -lrtha *.csv
```

```
ls: cannot access '~/.kaggle': No such file or directory
-rw-r--r-- 1 root root  2.3K Apr 18 19:37 kaggle.json
Downloading food-com-recipes-and-user-interactions.zip, 39 MB in 1s
100%|#######################################|Loading dataset...
100%|#######################################|Loading dataset...
100%|#######################################|Loading dataset...
```

```
!git clone https://github.com/karpathy/char-rnn
```

```
Cloning into 'char-rnn'...
remote: Enumerate objects: 53, done.
remote: Total 53 (delta 0), reused 0 (delta 0), pack-reused 53
Receiving objects: 100% (53/53), 6.82 MiB | 6.82 MiB/s, done.
Resolving deltas: 100% (18/18), done.
```

```
!cd char-rnn/data/tinyshakespeare/ && ls -l
```

```
total 1084
-rw-r--r-- 1 root root 1084310 Mar 13 00:57 input.txt
```

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt
```

```
First Citizen:
Before we proceed any further, hear me speak.

All:
Speak, speak.

First Citizen:
You are all resolved rather to die than to famish?

All:
Resolved. resolved.

First Citizen:
First, you know Caius Marcius is chief enemy to the people.

All:
We know't, we know't.

First Citizen:
Let us kill him, and we'll have corn at our own price.
Is't a verdict?

All:
No more talking on't; let it be done: away, away!

Second Citizen:
One word, good citizens.

First Citizen:
We are accounted poor citizens, the patricians good.
What authority surfeits on would relieve us: if they
would yield us but the superfluity, while it were
wanting, the
```

Create TensorFlow dataset

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | wc
```

```
  137  1707 10000
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | wc
```

```
  137  1707 10000
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 1
```

```
   1000 e
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 2
```

```
   1000 e
    600 o
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 3
```

```
   1000 e
    600 o
    571 t
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 4
```

```
   1000 e
    600 o
    571 t
    515 h
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 5
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 6
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 7
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 8
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 9
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 10
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 11
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 12
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 13
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 14
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 15
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 16
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 17
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 18
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 19
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 20
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 21
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 22
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 23
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 24
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 25
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 26
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 27
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 28
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 29
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 30
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 31
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 32
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 33
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 34
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 35
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 36
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 37
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 38
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 39
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 40
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 41
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 42
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 43
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 44
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 45
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 46
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 47
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 48
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 49
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 50
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 51
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 52
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 53
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 54
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
     20 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 55
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
     20 Z
     19 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 56
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
     20 Z
     19 Z
     18 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 57
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
     20 Z
     19 Z
     18 Z
     17 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 58
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
     20 Z
     19 Z
     18 Z
     17 Z
     16 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 59
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
     20 Z
     19 Z
     18 Z
     17 Z
     16 Z
     15 Z
```

```
!cat char-rnn/data/tinyshakespeare/input.txt | head -c 10000 | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 60
```

```
   1000 e
    600 o
    571 t
    515 h
    465 s
    377 a
    335 i
    324 n
    319 r
    268 d
    244 l
    205 u
    172 m
    167 c
    158 w
    137 g
    117 y
    110 p
    108 b
    101 A
     97 I
     91 k
     86 v
     79 F
     76 N
     74 M
     70 O
     69 T
     66 E
     64 R
     61 S
     58 C
     55 W
     53 L
     51 B
     49 U
     47 D
     46 G
     44 P
     42 H
     41 Y
     39 K
     38 V
     34 J
     31 X
     30 Q
     27 Z
     26 Z
     25 Z
     24 Z
     23 Z
     22 Z
     21 Z
     20 Z
     19 Z
     18 Z
     17 Z
     16 Z
     15 Z
     14 Z
```

Split sequences into input and output texts

Input sequences, separate and shift the first letter from the rest and target text. For example, my sequence length is 4 and our text is "Beta". The input sequence would be "Bet", and the target sequence would be "eta".

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt
```

```
First Citizen:
Before we proceed any further, hear me speak.

All:
Speak, speak.

First Citizen:
You are all resolved rather to die than to famish?

All:
Resolved. resolved.

First Citizen:
First, you know Caius Marcius is chief enemy to the people.

All:
We know't, we know't.

First Citizen:
Let us kill him, and we'll have corn at our own price.
Is't a verdict?

All:
No more talking on't; let it be done: away, away!

Second Citizen:
One word, good citizens.

First Citizen:
We are accounted poor citizens, the patricians good.
What authority surfeits on would relieve us: if they
would yield us but the superfluity, while it were
wanting, the
```

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt | wc
```

```
  137  1707 1000
```

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt | tr -c "a-zA-Z\n" " " | wc
```

```
  137  1707 1000
```

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 1
```

```
    600 o
```

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 2
```

```
    600 o
    571 t
```

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort -rn | head -n 3
```

```
    600 o
    571 t
    515 h
```

```
!head -c 1000 char-rnn/data/tinyshakespeare/input.txt | tr -c "a-zA-Z\n" " " | grep -o . | sort | uniq -c | sort
```

Image source: Tree generated with an RNN-textbook

```
def generate_tree(model, start_pixelseq, temperature=1.0):
    """
    Generate a tree pixel sequence from the trained model.
    Uses temperature sampling for diversity.
    Returns: (seq_length, 3) array of [x, y, pen_state]
    """
    seq = [start_pixelseq]
    for step in range(MAX_SEQ_LEN - 1):
        inp = np.array(seq[-SEQ_LEN:])
        if len(inp) < SEQ_LEN:
            pad = np.zeros((SEQ_LEN - len(inp), 3))
            inp = np.concatenate([pad, inp])
        inp = inp.reshape(1, SEQ_LEN, 3)
        pred = model.predict(inp, verbose=0)[0]
        # Temperature sampling
        pen_state = np.argmax(pred[2:])
        if pen_state == 1:
            break
        new = pred[:2] + np.random.normal(0, 0.01, 2)
        new = np.clip(new, 0.0, 1.0)
        new = np.round(new, 3)
        seq.append(np.concatenate([new, [0]]))
    return np.array(seq)
```

```
start = np.array([0.500, 0.900, 0])
tree = generate_tree(model, start, temperature=0.8)
print(f"Generated tree: {len(tree)} points")
```

```
Generated tree: 14 points
```

```
def plot_tree(points, title=None):
    fig, ax = plt.subplots(1, 1, figure=(4, 4))
    ax.plot(points[:, 0], points[:, 1], 'o-', linewidth=2, markersize=4)
    ax.set_aspect('equal')
    ax.grid(True, alpha=0.3)
    if title:
        ax.set_title(title)
    plt.show()
plot_tree(tree, "Generated tree")
```



```
def plot_multi(trees, cols=4):
    n = len(trees)
    rows = (n + cols - 1) // cols
    fig, axes = plt.subplots(rows, cols, figure=(3*cols, 3*rows))
    for i, tree in trees:
        ax = axes[i // cols, i % cols]
        ax.plot(tree[:, 0], tree[:, 1], 'o-', linewidth=1.5, markersize=3)
        ax.set_aspect('equal')
        ax.grid(True, alpha=0.3)
        ax.set_title(f"#{i}")
    plt.show()
plot_multi([(i, generate_tree(model, start, temperature=0.8)) for i in range(8)])
```



```
def generate_batch(model, start, n):
    return [generate_tree(model, start, temperature=t) for t in [0.5, 0.8, 1.0, 1.2]]
batch = generate_batch(model, start, 4)
print(f"Generated {len(batch)} trees")
```

```
Generated 4 trees
```

```
from keras.models import load_model
model = load_model('tree-generator.keras')
start = np.array([0.500, 0.900, 0])
```

```
WARNING:tensorflow:Compiled the loaded model, but the state of the optimizer is not restored.
```

```
tree = generate_tree(model, start, temperature=0.8)
print(f"Generated {len(tree)} points")
```

```
Generated 14 points
```

```
plot_tree(tree, "Generated tree")
```



```
batch = generate_batch(model, start, 4)
plot_multi([(i, batch[i]) for i in range(4)])
```



```
for t in [0.3, 0.5, 0.8, 1.0, 1.2, 1.5]:
    tree = generate_tree(model, start, temperature=t)
    print(f"temp={t}: {len(tree)} points")
```

```
temp=0.3: 14 points
temp=0.5: 14 points
temp=0.8: 14 points
temp=1.0: 14 points
temp=1.2: 14 points
temp=1.5: 14 points
```

```
# Save the generated dataset
np.save('generated_trees.npy', all_trees)
print("Saved generated_trees.npy")
```

```
Saved generated_trees.npy
```

```
# Load it back to verify
loaded = np.load('generated_trees.npy')
print(f"Loaded {len(loaded)} trees")
```

```
Loaded 4 trees
```

```
# Plot the saved dataset
plot_multi([(i, loaded[i]) for i in range(len(loaded))])
```



```
# Compare different temperatures
for t in [0.3, 0.8, 1.2]:
    trees = [(i, generate_tree(model, start, temperature=t)) for i in range(4)]
    plot_multi(trees)
    print(f"temperature={t}")
```



```
temperature=0.3
```

```
temperature=0.8
```



```
temperature=1.2
```

```
# Generate 100 trees
all_trees = [generate_tree(model, start, temperature=0.8) for _ in range(100)]
print(f"Generated {len(all_trees)} trees")
```

```
Generated 100 trees
```

```
# Save
np.save('tree-dataset.npy', all_trees)
```

```
# Load and verify
ds = np.load('tree-dataset.npy')
print(f"Shape: {ds.shape}")
print(f"Average points: {sum(len(t) for t in ds)/len(ds):.1f}")
```

```
Shape: (100, 14, 3)
Average points: 14.0
```

```
# Plot 8 random trees
import random
plot_multi([(i, ds[i]) for i in random.sample(range(len(ds)), 8)])
```



```
# Stats
print(f"Min points: {min(len(t) for t in ds)}")
print(f"Max points: {max(len(t) for t in ds)}")
```

```
Min points: 14
Max points: 14
```

```
# Save a few for later
for i in range(3):
    np.save(f'tree-{i}.npy', ds[i])
print("Saved tree-0.npy, tree-1.npy, tree-2.npy")
```

```
Saved tree-0.npy, tree-1.npy, tree-2.npy
```

```
# Load one and plot
t = np.load('tree-0.npy')
plot_tree(t, "Loaded tree")
```



```
# Load all three and plot
trees = [(i, np.load(f'tree-{i}.npy')) for i in range(3)]
plot_multi(trees)
```



```
# Generate 50 more
more = [generate_tree(model, start, temperature=0.8) for _ in range(50)]
print(f"Generated {len(more)} trees")
```

```
Generated 50 trees
```

```
# Save the full dataset
np.save('tree-dataset-v2.npy', more)
```

```
# Compare two datasets
ds = np.load('tree-dataset.npy')
ds2 = np.load('tree-dataset-v2.npy')
print(f"Dataset 1: {len(ds)} trees, avg {sum(len(t) for t in ds)/len(ds):.1f} points")
print(f"Dataset 2: {len(ds2)} trees, avg {sum(len(t) for t in ds2)/len(ds2):.1f} points")
```

```
Dataset 1: 100 trees, avg 14.0 points
Dataset 2: 50 trees, avg 14.0 points
```

```
# Plot both datasets side by side
plot_multi([('A', ds[0]), ('B', ds2[0])])
```



```
# Combine all
combined = ds.tolist() + ds2.tolist()
print(f"Combined: {len(combined)} trees")
```

```
Combined: 150 trees
```

```
# Save combined
np.save('tree-dataset-combined.npy', combined)
```

```
# Load and verify
c = np.load('tree-dataset-combined.npy')
print(f"Shape: {c.shape}")
```

```
Shape: (150, 14, 3)
```

```
# Random 8 from combined
plot_multi([(i, c[i]) for i in random.sample(range(len(c)), 8)])
```



```
# Stats
print(f"Min points: {min(len(t) for t in c)}")
print(f"Max points: {max(len(t) for t in c)}")
```

```
Min points: 14
Max points: 14
```

```
# Save a few for later
for i in range(3):
    np.save(f'tree-{i}.npy', c
```