# Speech Generation through RNN

**Anahita Esfandiaryfard**                                    A.ESFANDIARYFARD@STUDIUM.UNICT

**Gianluigi Mazzaglia**[1]                                     UNI310589@STUDIUM.UNICT.IT

## 1. Model Description

We describe here **Speech Generation**, a deep model that use RNN to learn from a TV show script and generate for us a new scene related, to create an RNN model we need to input 3 dimensions which are batch size as we set to 8, the batch sequence length that we tried between numbers of 16, 32, 64 to see which one gives us a better result and in the end we chose 64 because of lower loss function and better prediction, for SceneLength we first counted the actual length of the scene based on the script to have an idea of the average of actual length then we choose a number that is close to the actual length and also is multiple of our batch seq length which in this 384. (see Fig. 1).
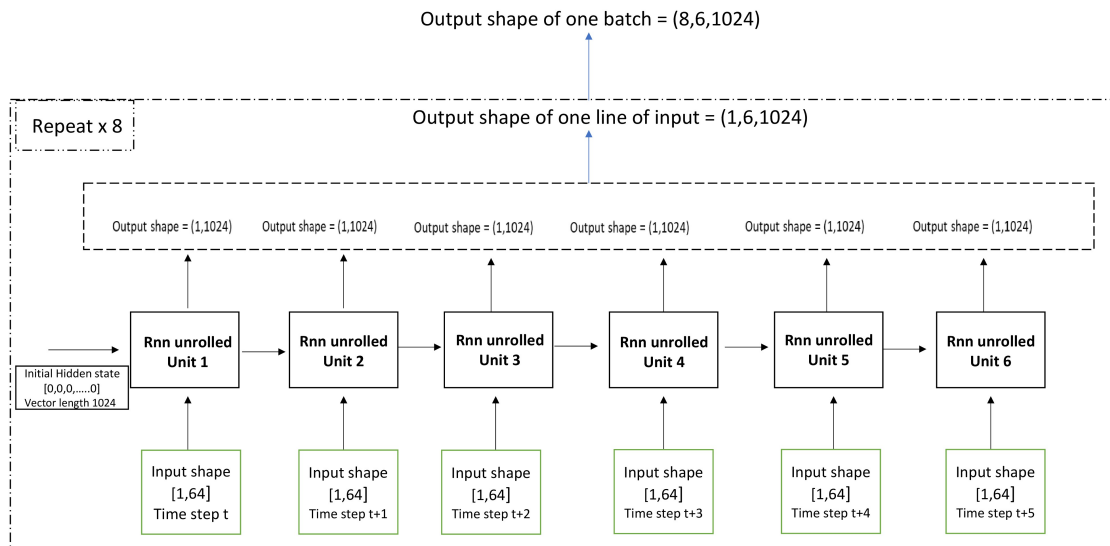


Figure 1: A simple representation of the proposed model.

## 2. Dataset

The dataset is provided by kaggle names: "Friends TV Show Script" This dataset contains a .txt file that contains scripts of all the episodes of F.R.I.E.N.D.S. aired through 1994 to 2004. The text file consists section written like : "[Scene: Central Perk, Chandler, Joey, Phoebe, and Monica are there.]" so we used this pattern to count number of scenes in the text.

## 3. Training procedure

For this models we first created a dictionary for words in the text file by assigning a number to each word. Then we used regex to find scenes in the text file and by dividing the total length of the file to number of scenes we found the average length of each scene. For our RNN model we decided to have batch size of 8, we set batch sequence length to 64 and we decide to set size of rnn to 1024. Based on the calculation and the size of the average scene length we defined number 384 for scene length which is the number near actual scene length and its a multiple of 64. To create the rnn model we used lstm because it can handle the information in memory for the long period of time as compare to RNN. LSTM networks are a type of RNN that uses special units in addition to standard units. LSTM units include a 'memory cell' that can maintain information in memory for long periods of time. This memory cell lets them learn longer-term dependencies. LSTMs deal with vanishing and exploding gradient problem by introducing new gates, such as input and forget gates, which allow for a better control over the gradient flow and enable better preservation of "long-range dependencies". Next We used Adam optimizer and run the model in 25 epochs because it seems after 25 epochs loss value start to increase.

## 4. Experimental Results

In the first picture you can see the result of the network with 32 batches sequence length and 35 epochs as you can see the loss value is 1.69



Figure 2: RNN result with batch sequence length of 32 and 35 epochs

2

But as you can see in epoch 20 we had less loss value 1.56 which we see is less than 35 epochs so for the next attempt we decided to go with 20 epochs and also we increased batch sequence length 64. As we see here the loss value is improving and the loss value is 1.37



Figure 3: The result with 20 epochs and 64 batch sequence length

As you can see in the following our model in the beginning just writing "Phoebe: I can't believe", it generated :

**"Phoebe: I can't believe this trashy bang Jamie amazed SCOUT Crash 30 Yup Kung Hoyt wonderfully interns hitchhike Alan-ness Chinese Soaps Mississipily fling Scuse Changing"**

As we can see it's something meaningless but after training the model the generated test appeared much more meaningful:

```
Phoebe: I can't believe this Wow have believe you . I can't believe you did this .
Monica: But what if it is , what happened ?
Ross: He was staring . This is where we're living under your wedding .
Erica: Wow ! Maybe , maybe she wouldnt mind about the little bit of our office .
Ross: Yeah , I mean , if you moved there , you get a lot of guys , you know , when I was younger . . . I don't know what you've got going this squirmy quality that I wrote in that
Rachel: Oh , well , that's right . I blocked that out .
Ross: Oh , no , no , no , no . . . I love you .
Rachel: Oh , I can't believe you guys are leaving a little child .
Ross: Oh my God .
Chandler: Yeah , I mean , you know , you know , you don't have a friend because you needed to make a special Valentines Day .
Ross: Oh , no !
Rachel: Well , I can't believe you did this .
Ross: I know , I know , I know . . . I know . . . I know . . . I know . . . I know . . . I know . . . I know . . . I know . . . I know . . . I know . . . I know . . .
Rachel: ( to Ross ) So , are you gonna change the last question so that I can make up ?
Ross: I don't know .
Rachel: You were so happy .
Ross: I know , I know . . . I know . . . I know . . . I know . . . I know . . . I know . .
```

Figure 4: Final Generated Text