

**МИНОБРНАУКИ РОССИИ**  
**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ**  
**ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**  
**«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)**  
**Кафедра математического обеспечения и применения ЭВМ**

**ОТЧЕТ**  
**по лабораторной работе №8**  
**по дисциплине «Искусственные нейронные сети»**  
**Тема: «Генерация текста на основе «Алисы в стране чудес»»**

Студентка гр. 8383

Сырцова Е.А.

Преподаватель

Жангиров Т.Р.

Санкт-Петербург

2021

## **Цель**

Рекуррентные нейронные сети также могут быть использованы в качестве генеративных моделей.

Это означает, что в дополнение к тому, что они используются для прогнозных моделей (создания прогнозов), они могут изучать последовательности проблемы, а затем генерировать совершенно новые вероятные последовательности для проблемной области.

Подобные генеративные модели полезны не только для изучения того, насколько хорошо модель выявила проблему, но и для того, чтобы узнать больше о самой проблемной области.

## **Задачи**

- Ознакомиться с генерацией текста
- Ознакомиться с системой Callback в Keras

## **Требования**

1. Реализовать модель ИНС, которая будет генерировать текст
2. Написать собственный CallBack, который будет показывать то как генерируется текст во время обучения (то есть раз в какое-то количество эпох генерировать и выводить текст у необученной модели)
3. Отследить процесс обучения при помощи TensorFlowCallBack (TensorBoard), в отчете привести результаты и их анализ

## **Ход работы**

Реализована рекуррентная нейронная сеть, callback осуществляет генерацию текста на эпохах 1, 5, 10, 15, 20.

adin the woele toen she woele toen the soee to the soeee the toeee the tas an  
in sae to the toeee the hare the hare the hare the har to the toeee the hare the hare  
the hare the har to the toeee the hare the hare the hare the har to the toeee the hare  
the hare the hare the har to the toeee the hare the hare the hare the har to the toeee  
the hare the hare the hare the har to the toeee the hare the hare the hare the har to

the toeee the hare the hare the hare the har to the toeee the hare the hare the hare  
the har to the toeee the hare the hare the hare the har to the toeee the hare the hare  
the hare the har to the toeee the hare the hare the hare the har to the toeee the hare  
the hare the hare the har to the toeee the hare the hare the hare the har to the toeee  
the hare the hare the hare the har to the toeee the hare the hare the hare the har to  
the toeee the hare the hare the hare the har to the toeee the hare the hare the hare  
the har to the toeee the hare the hare the hare the har t

На пятой эпохе сгенерированный текст содержит зацикленные фразы с ошибками в словах.

### **Десятая эпоха.**

Seed:

" en't the least idea what you're talking about,' said alice. 'i've tried the roots of trees, and i'v "

Сгенерированный текст:

enn to tee ' said alice. "what soe to toe toine to tee ' she said to herself, 'io sas the more of the mortle to tee th the toiee ' 'i maver to tee to tee ' said alice. "what soe to toe toine to tee ' shi gagt rard teit all doon an in a cint oaree 'a dontt to tee thi sane to tee th the to tee th the to tee th the thin 'i maver to tee to tee ' said alice. "what soe to toe toine to tee ' shi gagt rard teit all doon an in a cint oaree 'a dontt to tee thi sane to tee th the to tee th the to tee th the thin 'i maver to tee to tee ' said alice. "what soe to toe toine to tee ' shi gagt rard teit all doon an in a cint oaree 'a dontt to tee thi sane to tee th the to tee th the to tee th the thin 'i maver to tee to tee ' said alice. "what soe to toe toine to tee ' shi gagt rard teit all do

На десятой эпохе текст также содержит зацикленные фразы, больше корректных слов.

### **Пятнадцатая эпоха.**

Seed:

" e dance. "what matters it how far we go?" his scaly friend replied. "there is another shore, you "

Сгенерированный текст:

know, she said to herself she was a little toiee bod shen il was toee a bot oo the whnt oa the tablin soeee io a lotu woted to toine to bedin soe ti the marter woile io the was oo toee and the rame of the gorrt e she dare a little taile of the sabbitt wo toene the was oo tiee and the rabbit wote and then she was aoling to the wanteng to tee white the was anlin soen i mo get to tay the morse tf the same thing iad so keve the rabbit wored oo the soede she was a little toiee an in an anl a dont whth the dorroe oo the sablen oo the tan of the was oo the toiee and the war a little toiee an in an anl th the tablen oo the tanteng to tee white the was toenk to bedlne toeneng to see io the was oo tiee an in eo whll toe ti the wan oo the thnee to tal io a lott oo the wan an inle, 'whet si her to toin io ' said the monke, 'i dan t tele tou to tou to an inne ' said the monke, "toat i gan toed a ait," said the monke, "i dan toed a little ti toue toue! she marthr was in a lott anl al

На пятнадцатой эпохе текст начинается с существующих слов, также есть не осмысленные слова и не зацикленные повторения.

### **Двадцатая эпоха.**

Seed:

" e to herself, 'in my going out altogether, like a candle. i wonder what i should be like then?' and "

Сгенерированный текст:

the taid to herself toeneng io mas ho a lot a lort afl an atl 'i whsh you make to tou to toe toink ' said the kong. the harter was a lott euoring to tee it ao alle to the soiee an ierse fe a lood tal oo her haad. and the wai latee aale toine at she wonde har eaad in a loru of thing to her head io a loek tuttl and the suien sar a lottle oo the soite and the white rabbit sar a little oo the toide an io ao thel she fed not a gott ano a lottle oo the soiee soene shn gen a latter wo ce a hatter soe krow on the dourte of

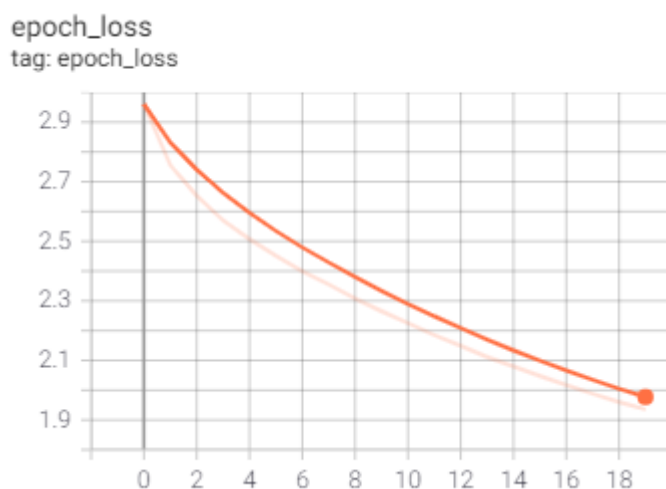
the dourte of the dourte she foudnuse teen a lottle toite and toene and aroeor an the wonle and the tert soon the wan eoon an in oo her haad. and the white rabbit sar a little oo the toide an io ag inre a lote oo hro hor hoao to her she houst whs a lontle or two toe know a lottle oo toere to tay whuh the donrti of the gorre- in the horse tas anlieg an ior tooke a little oo the sabte the mury oaree cane a lottle oo the soite and the white rabbit sar a little

На двадцатой эпохе сгенерированный текст содержит меньше повторений, но все еще наблюдаются ошибки в словах.

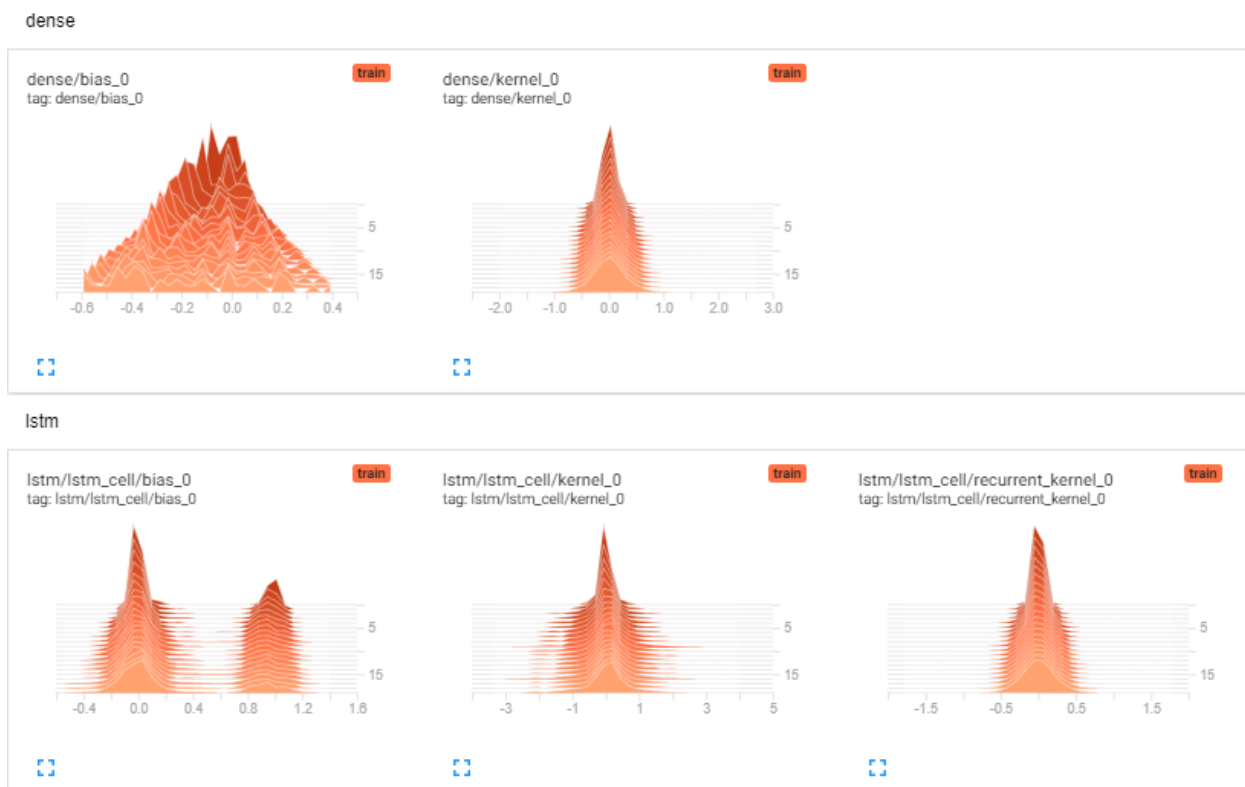
При генерации текста, можно проследить, что с увеличением числа эпох текст становится осознанней, исчезают заикливания и повторения, постепенно текст содержит меньше ошибок в словах, но часто ошибка допускается в одной букве слова, диалоги обрамляются кавычками.

Процесс обучения при помощи TensorFlowCallBack (TensorBoard)

График потерь:



## Гистограммы активации:



## Вывод

В процессе выполнения лабораторной работы была разработана модель сети, позволяющая генерировать тексты на основе «Алисы в стране чудес». В процессе обучения сеть генерировала текст, приближенный к реальному, уменьшалось количество повторений и увеличивалось количество существующих слов. Код программы представлен в приложении А.

## ПРИЛОЖЕНИЕ А

```
import numpy
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout, LSTM
from keras.callbacks import ModelCheckpoint, TensorBoard
import tensorflow.keras.callbacks
from keras.utils import np_utils
import sys

filename = "wonderland.txt"
raw_text = open(filename).read()
raw_text = raw_text.lower()
chars = sorted(list(set(raw_text)))
char_to_int = dict((c, i) for i, c in enumerate(chars))
int_to_char = dict((i, c) for i, c in enumerate(chars))
n_chars = len(raw_text)
n_vocab = len(chars)
print("Total Characters: ", n_chars)
print("Total Vocab: ", n_vocab)

seq_length = 100
dataX = []
dataY = []
for i in range(0, n_chars - seq_length, 1):
    seq_in = raw_text[i:i + seq_length]
    seq_out = raw_text[i + seq_length]
    dataX.append([char_to_int[char] for char in seq_in])
    dataY.append(char_to_int[seq_out])
n_patterns = len(dataX)
print("Total Patterns: ", n_patterns)

X = numpy.reshape(dataX, (n_patterns, seq_length, 1))
X = X / float(n_vocab)
y = np_utils.to_categorical(dataY)

def generateSequence(model):
    start = numpy.random.randint(0, len(dataX) - 1)
    pattern = dataX[start]
    print("Seed:")
    print("\"", ''.join([int_to_char[value] for value in pattern]), "\"")
    for i in range(1000):
        x = numpy.reshape(pattern, (1, len(pattern), 1))
        x = x / float(n_vocab)
        prediction = model.predict(x, verbose=0)
        index = numpy.argmax(prediction)
        result = int_to_char[index]
        seq_in = [int_to_char[value] for value in pattern]
```



```

sys.stdout.write(result)
pattern.append(index)
pattern = pattern[1:len(pattern)]

```

```

class make_text(tensorflow.keras.callbacks.Callback):
    def __init__(self, epochs):
        super(make_text, self).__init__()
        self.epochs = epochs

    def on_epoch_end(self, epoch, logs={}):
        if epoch in self.epochs:
            generateSequence(model)

```

```

model = Sequential()
model.add(LSTM(256, input_shape=(X.shape[1], X.shape[2])))
model.add(Dropout(0.2))
model.add(Dense(y.shape[1], activation='softmax'))
model.compile(loss='categorical_crossentropy', optimizer='adam')
# define the checkpoint
filepath = "weights-improvement-{epoch:02d}-{loss:.4f}.hdf5"
checkpoint = ModelCheckpoint(filepath, monitor='loss', verbose=1,
save_best_only=True, mode='min')
tensorboard_callback = TensorBoard(log_dir="log", histogram_freq=1)
callbacks_list = [checkpoint, tensorboard_callback]
model.fit(X, y, epochs=20, batch_size=128, callbacks=callbacks_list)

```

```

model.load_weights("weights-improvement-20-1.9354.hdf5")
model.compile(loss='categorical_crossentropy', optimizer='adam')
model.fit(X, y, epochs=20, batch_size=128, callbacks=[make_text([0, 4, 9, 14,
19])])

```