

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра МО ЭВМ

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

Студент гр. 8383

Колмыков В.Д.

Преподаватель

Санкт-Петербург

2021

Цель работы.

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.

Ход работы.

Была реализована программа из методических указаний. Пример генерирующегося текста:

«eee oo the sabbit and the toide to be i vele the was a lott eoon tha whsh the war oo the pooe of the sabb to the white wabbit so tee that she was so tie whrt ou to toe kad io tae the was oo the tabb to tee white rabbit so tee that she was so tie pooe tf the was oo the tab it tas an the corrt, and the white rabbit sase the pabb to the whide whsh the whrt oatee th the was oo the tabbit oate the was so the whrt on the whst oo the whst oo the tabbit and the toide to be i vele the war a lott eo an toued oo the sooe, and the white rabbit sase thet she was so tie whrt on the whst oo the whst oo the tabbit and the toide to be i vele the war a lott eo an toued oo the sooe, and the white rabbit sase thet she was so tie whrt on the whst oo the whst oo the tabbit and the toide to be i vele the war a lott eo an toued oo the sooe, and the white rabbit sase thet she was so tie whrt on the whst oo the whst oo the tabbit and the toide to be i vele the war a lott eo an toued oo the sooe, and the wh»

Сгенерированный тест не содержит в себе много смысла. Лучшего результата можно было бы добиться, обучая нейросеть на словах.

Был реализован коллбэк для генерации текста раз в 5 эпох во время обучения модели:

```
class TextGenerationCallBack(Callback):
    def on_epoch_end(self, epoch, logs=None):
        if epoch % 5 == 0:
            start = numpy.random.randint(0, len(dataX) - 1)
            pattern = dataX[start]
            print("Seed:")
            print("\n", ''.join([int_to_char[value] for value in
pattern])), "\n")
            for i in range(100):
                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)]
            print("\nDone.")
```

Результат после первой эпохи:

«o the to»

После шестой:

«and the kan to the toeee 'nh thu to tee toen the toee to the toee to the toee to the toee to the t»

После 16:

«e to tea io the san the dar soite in the rooeo of the sabee bnd the matte care woine she was anl t»

Как видно, с обучением нейросети кол-во повторений в тексте уменьшается.

Также был добавлен коллбэк TensorBoard, добавляющий логгирование для возможности визуального наблюдения за происходящим внутри модели во время обучения. Например, можно просмотреть изменение потерь во время обучения, приведенную на рис. 1.



Рисунок 1 – Гистограмма потерь нейросети

Из графика видно, что во время обучения потери снижаются до ~2.

Также можно наблюдать гистограмму активаций для слоев нейросети, приведенную на рис. 2.:

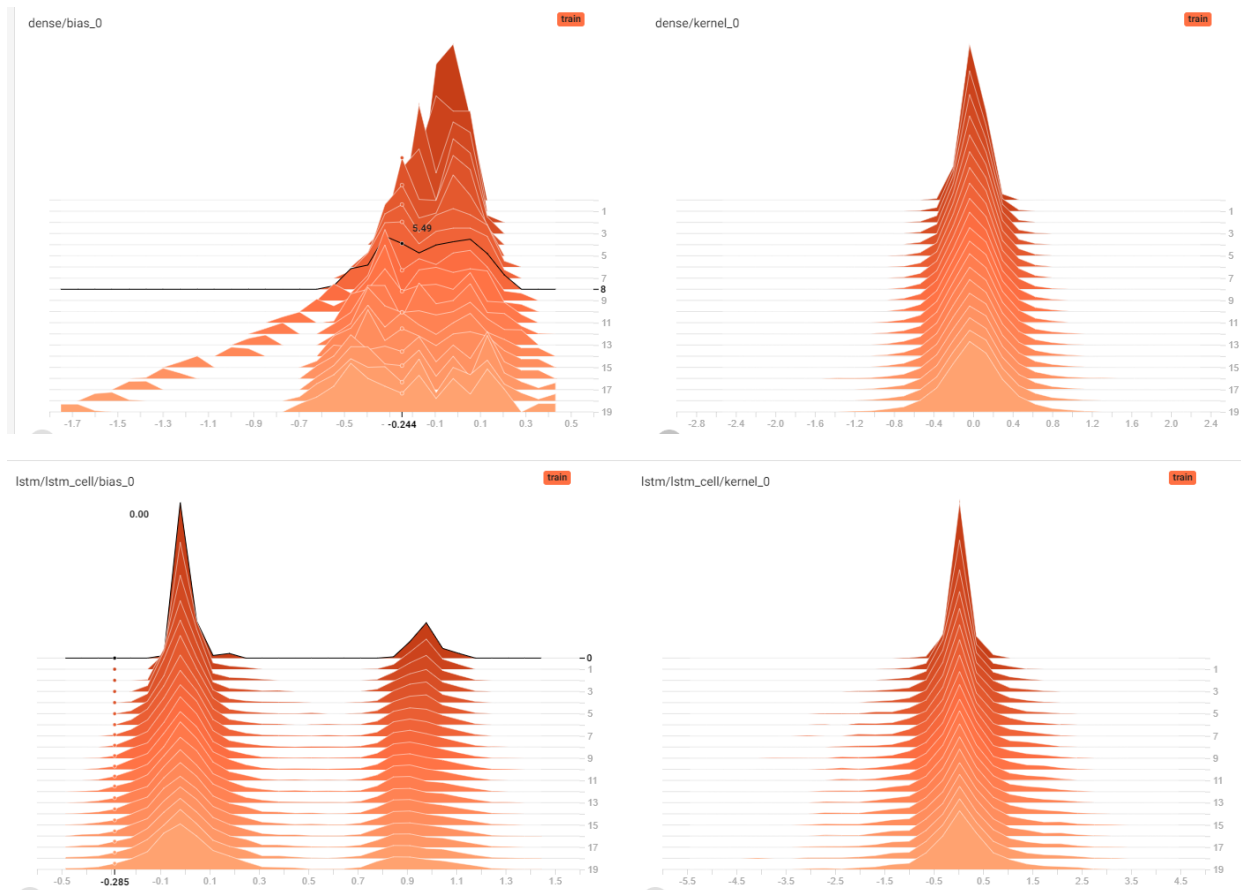


Рисунок 2 – Гистограмма активаций для слоев нейросети

На ней можно увидеть значения весов, которые были установлены во время обучения.

Выводы.

В ходе выполнения работы был реализована нейросеть для генерации текста на основе «Алисы в стране чудес». Было проведено ознакомление с генерацией текста, системой коллбэков в Keras а также с TensorBoard.