

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

ОТЧЕТ
по лабораторной работе №6
по дисциплине «Искусственные нейронные сети»
Тема: «Прогноз успеха фильмов по обзорам»

Студентка гр. 8382

Звегинцева Е.Н.

Преподаватель

Жангиров Т.Р.

Санкт-Петербург

2021

Цель работы.

Прогноз успеха фильмов по обзорам (Predict Sentiment From Movie Reviews).

Задачи.

- Ознакомиться с задачей регрессии
- Изучить способы представления текста для передачи в ИНС
- Достигнуть точность прогноза не менее 95%

Требования.

- Построить и обучить нейронную сеть для обработки текста
- Исследовать результаты при различном размере вектора представления текста
- Написать функцию, которая позволяет ввести пользовательский текст (в отчете привести пример работы сети на пользовательском тексте)

Ход работы.

Была создана и обучена модель искусственной нейронной сети в соответствии с условиями.

```
model = Sequential()
model.add(layers.Dense(60, activation="relu",
input_shape=(10000,)))

model.add(layers.Dropout(0.5, noise_shape=None, seed=None))
model.add(layers.Dense(56, activation="relu"))
model.add(layers.Dropout(0.2, noise_shape=None, seed=None))
model.add(layers.Dense(56, activation="relu"))

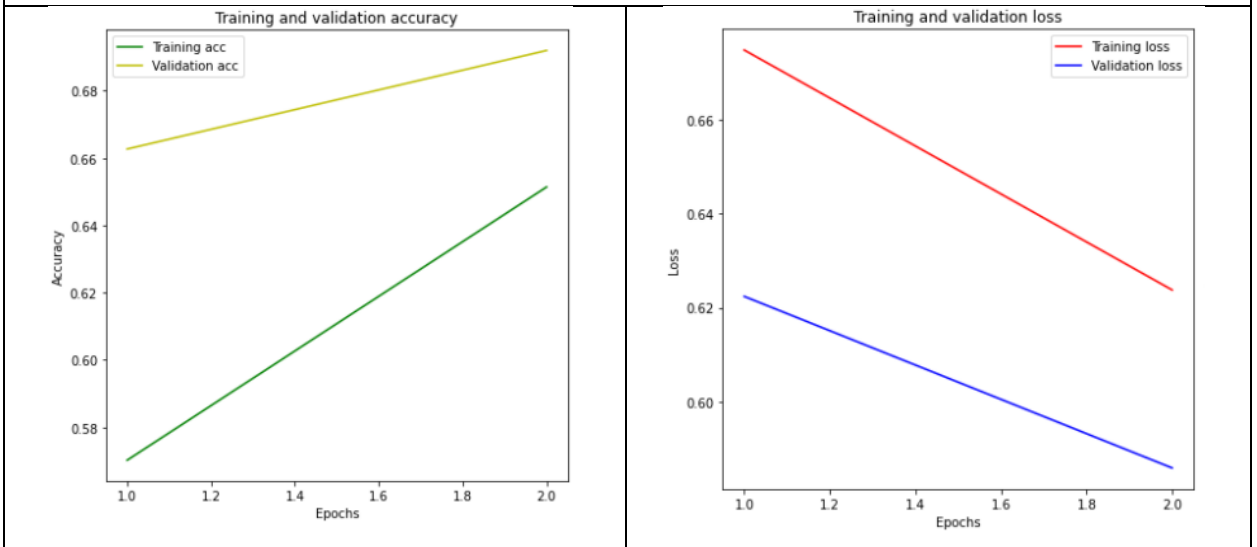
model.add(layers.Dense(1, activation="sigmoid"))
model.summary()
```

При тестировании которой, получили следующие результаты. Точность достигла 89,5% на валидационных данных.

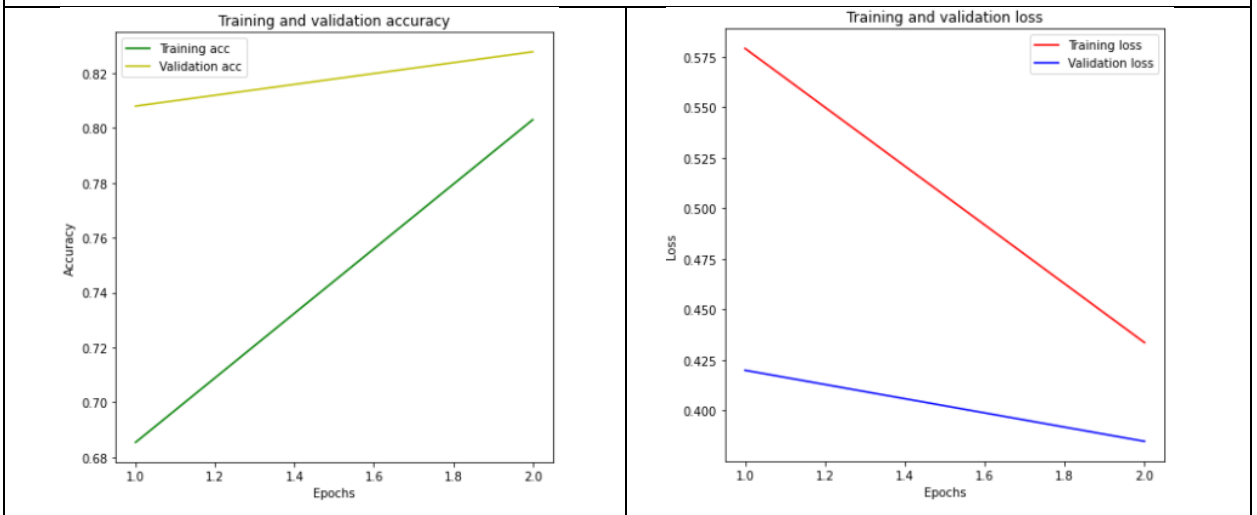
```
Epoch 1/2
80/80 [=====] - 2s 18ms/step -
loss: 0.5654 - accuracy: 0.6859 - val_loss: 0.2640 -
val_accuracy: 0.8939
Epoch 2/2
80/80 [=====] - 1s 15ms/step -
loss: 0.2306 - accuracy: 0.9112 - val_loss: 0.2579 -
val_accuracy: 0.8968
0.8953499794006348
```

Протестировано поведение при варьирующемся размере вектора представления текста.

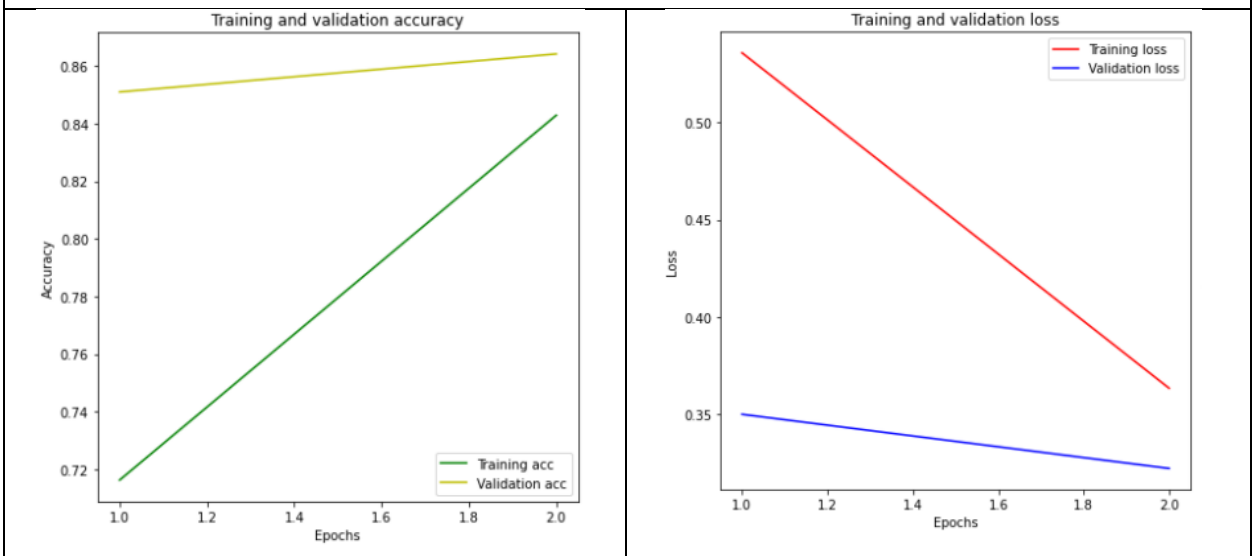
50

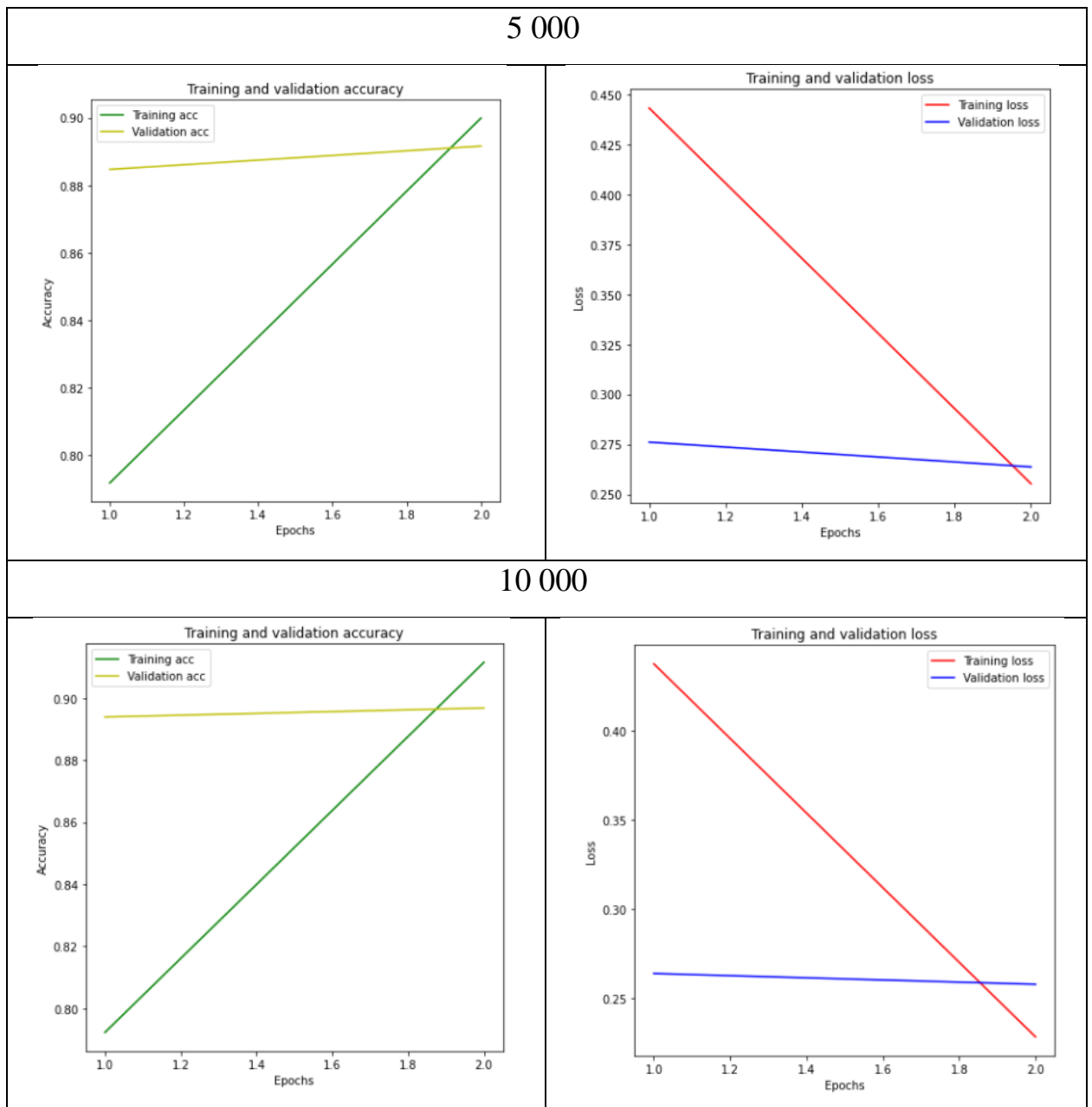


500



1 000





Как мы видим, наибольшая точность была достигнута на наибольшем размере вектора, но после 5000 она лишь незначительно увеличивается (выбрано 10000 в связи с ближайшему к 90% значению).

Была написана функция `user_load` для загрузки пользовательского текста и прогнозирования успеха фильма по этому тексту.

Пример работы приведен ниже.

```
Input string:
Great movie with a great story
[[20, 19, 6, 87, 65]]
[[0.793855]]
```

```
Input string:  
I'm crying that I spent my time on it.  
[[2578, 15, 1084, 61, 58, 23, 12]]  
[[0.409451]]
```

Нейронная сеть работает корректно.

Вывод.

Была построена сеть, прогнозирующая оценку фильма по обзорам. Было рассмотрено преобразование текста в формат, с которым может работать нейронная сеть. Было исследовано влияние размера вектора представления текста и выявлено, что наибольшей точностью обладает сеть с максимальным размером вектора, равным 10000. Была написана функция прогнозирования оценки по пользовательскому тексту.