

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

Студент гр. 8383

Мирсков А. А.

Преподаватель

Жангиров Т. Р.

Санкт-Петербург

2021

Цель работы.

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

Задачи:

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

Выполнение работы.

Для решения поставленной задачи была реализована следующая нейросеть:

```
model = Sequential()
```

```
# Input - Layer
```

```
model.add(layers.Dense(50, activation = "relu", input_shape=(vec_size, )))
```

```
# Hidden - Layers
```

```
model.add(layers.Dropout(0.3, noise_shape=None, seed=None))
```

```
model.add(layers.Dense(50, activation = "relu"))
```

```
model.add(layers.Dropout(0.2, noise_shape=None, seed=None))
```

```
model.add(layers.Dense(50, activation = "relu"))
```

```
# Output- Layer
```

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

Данная нейросеть имеет точность прогноза около 90%. Результат выполнения программы представлен ниже.

```
40000/40000 [=====] - 11s 265us/step - loss: 0.4027 - accuracy: 0.8238 - val_loss: 0.2600 - val_accuracy: 0.8981
Epoch 2/2
40000/40000 [=====] - 5s 116us/step - loss: 0.2151 - accuracy: 0.9172 - val_loss: 0.2698 - val_accuracy: 0.8908
```

Далее была написана функция обработки пользовательского текста.

```
def user_text_vectorize(text):
    for i in punctuation:
        text = text.replace(i, ' ')
    text = text.lower().split()
    index = imdb.get_word_index()
    coded_text = [index.get(i)+3 for i in text]
    vectorized_text = vectorize(np.asarray([coded_text]))
    return vectorized_text
```

Функция была протестирована на двух пользовательских отзывах.

Данный отзыв программа классифицировала как положительный: The film is very good. I will definitely watch the second part.

Данный отзыв программа классифицировала как отрицательный: The film is very bad. I do not recomend you to watch.

Далее был изменён размер вектора входных данных на 12000. Результат представлен ниже.

```
40000/40000 [=====] - 56s 1ms/step - loss: 0.4213 - accuracy: 0.8084 - val_loss: 0.2629 - val_accuracy: 0.8959
Epoch 2/2
40000/40000 [=====] - 263s 7ms/step - loss: 0.2091 - accuracy: 0.9209 - val_loss: 0.2709 - val_accuracy: 0.8919
```

Точность незначительно увеличилась.

Выводы.

В ходе выполнения лабораторной работы была написана программа для обучения нейросети для прогноза успеха фильмов по обзорам. Были изучены способы представления текста для передачи в ИНС.