

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. 8383

Аверина О.С.

Преподаватель

Жангиров Т.Р.

Санкт-Петербург

2021

Цель.

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

Выполнения работы.

1. Был загружен датасет IMDb, встроенный в Keras.

```
(training_data, training_targets), (testing_data,
testing_targets) = imdb.load_data(num_words=10000)

data = np.concatenate((training_data, testing_data), axis=0)

targets = np.concatenate((training_targets, testing_targets),
axis=0)
```

2. Были подготовлены данные. Каждый обзор был векторизован и заполнен нулями, чтобы вектор содержал 10000 числе.

```
def vectorize(sequences, dimension = 10000):

    results = np.zeros((len(sequences), dimension))

    for i, sequence in enumerate(sequences):

        results[i, sequence] = 1

    return results

data = vectorize(data)

targets = np.array(targets).astype("float32")
```

3. Датасет был разделен на обучающий и тестировочный наборы. Обучающий набор будет состоять из 40 000 обзоров, тестировочный — из 10 000.

```
test_x = data[:10000]

test_y = targets[:10000]
```

```
train_x = data[10000:]
train_y = targets[10000:]
```

4. Была создана нейронная сеть. Была выбрана последовательная модель и добавлены входные, скрытые и выходные слои. В качестве входного и скрытых слоев выбраны полносвязные слои из 50 нейронов с функцией активации 'Relu'. Для предотвращения переобучения между слоями добавлены два слоя 'dropout' с вероятностями исключения 0.3 и 0.2. В качестве выходного также выбран полносвязный слой, но с функцией активации 'sigmoid' и размерностью 1.

```
model = Sequential()
# Input - Layer
model.add(layers.Dense(50, activation="relu",
input_shape=(10000,)))
# 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"))
model.summary()
```

Layer (type)	Output Shape	Param #
=====		
dense_1 (Dense)	(None, 50)	500050

dropout_1 (Dropout)	(None, 50)	0

dense_2 (Dense)	(None, 50)	2550

dropout_2 (Dropout)	(None, 50)	0
dense_3 (Dense)	(None, 50)	2550
dense_4 (Dense)	(None, 1)	51
=====		
Total params: 505,201		
Trainable params: 505,201		
Non-trainable params: 0		

5. Модель была скомпилирована. Оптимизатором выбран ‘Adam’, в качестве функции потерь - бинарная кросс-энтропия, метрика оценки - точность.

```
model.compile(optimizer="adam", loss="binary_crossentropy",
metrics=["accuracy"])
```

6. Модель была обучена. Размер партии - 500, количество эпох - 2.

```
results = model.fit( train_x, train_y, epochs=2, batch_size=500,
validation_data =(test_x, test_y))
```

7. Модель достигла точности на тестовых данных 89,6%. Также модель была протестирована с помощью функции для ввода пользовательского текста на рецензиях, взятых из интернета. Результат тестирования представлен на рис. Тексты рецензий представлены в приложении А.

```

Epoch 2/2
80/80 [=====] - 1s 14ms/step - loss: 0.2147 - accuracy: 0.9201 - val_loss: 0.2591 - val_accuracy: 0.8968
0.8956999778747559
0.99826056
1
It is positive feedback
Answer is true.

0.39568087
0
It is negative feedback
Answer is true.

0.018774688
0
It is negative feedback
Answer is true.

0.98752767
1
It is positive feedback
Answer is true.

```

8. Исследуем результаты при различных размерах вектора представления текста.

При размере вектора 5000:

Epoch 1/2

```

80/80 [=====] - 1s 14ms/step - loss:
0.5508 - accuracy: 0.7046 - val_loss: 0.2746 - val_accuracy:
0.8869

```

Epoch 2/2

```

80/80 [=====] - 1s 9ms/step - loss:
0.2485 - accuracy: 0.9040 - val_loss: 0.2657 - val_accuracy:
0.8914

```

```

0.8891499936580658

```

При размере вектора 10000:

Epoch 1/2

```

80/80 [=====] - 2s 20ms/step - loss:
0.5257 - accuracy: 0.7361 - val_loss: 0.2626 - val_accuracy:
0.8946

```

Epoch 2/2

```
80/80 [=====] - 1s 14ms/step - loss:
0.2147 - accuracy: 0.9201 - val_loss: 0.2591 - val_accuracy:
0.8968

0.8956999778747559
```

При размере вектора 15000:

Epoch 1/2

```
80/80 [=====] - 3s 29ms/step - loss:
0.5072 - accuracy: 0.7530 - val_loss: 0.2526 - val_accuracy:
0.9001
```

Epoch 2/2

```
80/80 [=====] - 2s 21ms/step - loss:
0.1963 - accuracy: 0.9271 - val_loss: 0.2673 - val_accuracy:
0.8934
```

0.8967500030994415

При размере вектора 17000:

Epoch 1/2

```
80/80 [=====] - 3s 34ms/step - loss:
0.5225 - accuracy: 0.7363 - val_loss: 0.2542 - val_accuracy:
0.8985
```

Epoch 2/2

```
80/80 [=====] - 2s 25ms/step - loss:
0.1883 - accuracy: 0.9316 - val_loss: 0.2679 - val_accuracy:
0.8956
```

0.8970500230789185

Анализируя результаты исследования, можно сделать вывод, что при изменении размера вектора точность модели практически не меняется, но при увеличении размера значительно вырастает размер используемой оперативной памяти.

Вывод.

В ходе лабораторной работы была построена нейронная сеть, прогнозирующая успех фильмов по обзорам. Была разработан функционал, позволяющий считывать пользовательский текст из файла. Сеть была протестирована на пользовательском тексте отзыва.

Приложение А.

Текст рецензий.

Рецензия 1.

Позитивная рецензия.

The film is simply gorgeous! Best movie I've ever seen. Everything is great. And the script, and directing, and special effects, well, the music... Hans Zimmer did his best (standing ovation)! The actors also played well... Everything is gorgeous. The only thing Nolan picked up a little bit... I immediately noticed this (although of course it is clear that this was done for the "piquancy" of the script and the logic of what is happening)... In a dream, time goes not slower, but faster than in reality. Well, for example. How much do you sleep at night?... Maybe 7-8 hours?... But you do not see a dream for 15-16 hours, but only a few minutes (usually 2-3 minutes), or even less than a minute! In the film, the hero DiCaprio explains to the girl that 5 minutes of sleep in reality is equal to several hours in a dream! And, consequently, the more levels of sleep, the slower time goes (who watched-will understand what I mean)! It's almost like this... When the hero DiCaprio with his beloved creates more and more levels of sleep, then in the end they would spend together not 50 years (!), as they say in the film, but only a few seconds (!). But I understand Nolan. The logic of the script and all that... I'd do that too... A little more than that...

Рецензия 2.

Негативная рецензия.

The story is about a daring challenge that a brave woman threw to fate. No, no, she didn't climb Mount Everest in her tights. And she didn't sink to the bottom of the Mariana Trench in a cosmetic mask. Still more serious. She decided to give birth. It is said that billions of women have done this before her, but this is not accurate. It remains unknown what inspired her to take such a risky step. It must be a thirst for achievement. But it is unlikely that a relationship with a vulnerable partner, whose sobs and sobs make up a noticeable part of the soundtrack. 9 months of asceticism remained behind the scenes. There are 30 minutes of final

naturalism in the frame. In the thriller genre. Not every fan of horror stories-scarecrows will bear it. I'm not an amateur, and it was easier for me to turn away. It is foolish to rely on the aesthetic reaction of the public, showing the suffering of the b-ka in real time. After all, in real life, such cases do not involve passive observation. Alas ,the "experience" failed. ADHD. I.e., the reasons are unclear. The world collapsed. Psychological problems piled up. We must somehow complete the Gestalt. It would be nice to put someone in jail. At least a midwife. Negligence? Incompetence? Medical error? Malice? I didn't understand what they were trying to accuse her of. There were no indications for preservation, the presentation was correct, no intervention was required, the child was born alive, an ambulance was called just in case... Many months passed before the heroine's faded gaze fell on the sprouted seeds and, oh joy, the Gestalt closed itself. The woman opened herself to a new experience. The story of overcoming it, obviously, should inspire. But no. The demotivator from the "childfree" generation still worked. To repeat this, in any case, "without special training", you will not wish the enemy. It seems that in the absence of war, sapiens loses the scale of grief and suffering. Ceases to consider childbirth and death an integral part of life. My grandmother gave birth and raised six children. I don't remember her ever thinking of herself as a heroine. As, indeed, most of the billions of mothers who lived before her. Maybe they were all hiding something?

Рецензия 3.

Негативная рецензия.

"The budget is visible in everything, and this concerns artistic implementation, not technical capabilities. The film certainly does not cause wild disgust, it is quite watchable, dynamic, with more or less lively actors (who really play, and do not utter pathetic speeches with a fixed expression), with streams of blood and what is most interesting, comical and parodic. And at the same time, it is absolutely not a successful film adaptation. The most important element is missing, without which everything is meaningless - there is no tournament, there are just chaotic clashes of opponents

(although the existence of the outer and earthly worlds is described in detail and the rules of the game are announced) and there is no convincing atmosphere and faith in the reality of what is happening (which was actually famous for the old trilogy). A monotonous cold color, no colors and everything faded, terrible heavy costumes, some "garbage" scenery-the action is in barns, in some mines, or even in some landfill.

Рецензия 4.

Положительная рецензия.

Recently, while typing in a search engine something like "existential comedies", I came across the series "High with delivery". Despite the high rating among film critics, it is not at all popular with the Russian-speaking audience. Maybe because of the translation (it sounds like another stupid comedy), or maybe because of the fact that viewing requires some tolerance, liberality in views (n-r, on the use of marijuana). But anyway, the series turned out to be absolutely wonderful. The plot is simple — the main character, "the guy with the weed" delivers marijuana to customers, each episode is the story of a new client. The movie is both sad and funny, without complicated ideas and pretentious messages. About ordinary people and about life that just happens. A kaleidoscope of life stories. The main character is a very real character (he, by the way, is played by the director himself), well, just a guy from the next street. He often gets into a little ridiculous situations — you look and do not know how to react: laughter and sadness, sympathy and Spanish shame — and often all at the same time. And his bike rides around New York City to music are a meditation that deserves special attention. I wonder why I was so hooked on the show? First of all, the mood. There is something Japanese about it, some elusive zen, a reminder-from which you want to stop, cancel all urgent matters and go for a walk in the nearest park, enjoy nature, a walk and pleasant company-simple joys.

I like that the author does not have a clear message for the viewer, I think everyone will have their own message. For me, this is about celebrating life, about the versatility of life stories, about beauty in the ordinary.

