

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. 8382

Преподаватель

Мирончик П.Д.

Жангиров Т.Р.

Санкт-Петербург

2021

ЦЕЛЬ

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

ЗАДАЧИ

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

Требования:

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

ХОД РАБОТЫ

Данные, передаваемые в нейронную сеть, представляют собой массив из 0 и 1 размерностью $idim$, где i -й индекс показывает наличие или отсутствие слова с индексом i в тексте.

Была реализована ИНС со следующей структурой:

```
Dense(50, activation="relu", input_shape=(idim,))
Dropout(0.5)
Dense(50, activation="relu")
Dropout(0.45)
Dense(50, activation="relu")
Dropout(0.3)
Dense(1, activation="sigmoid")
```

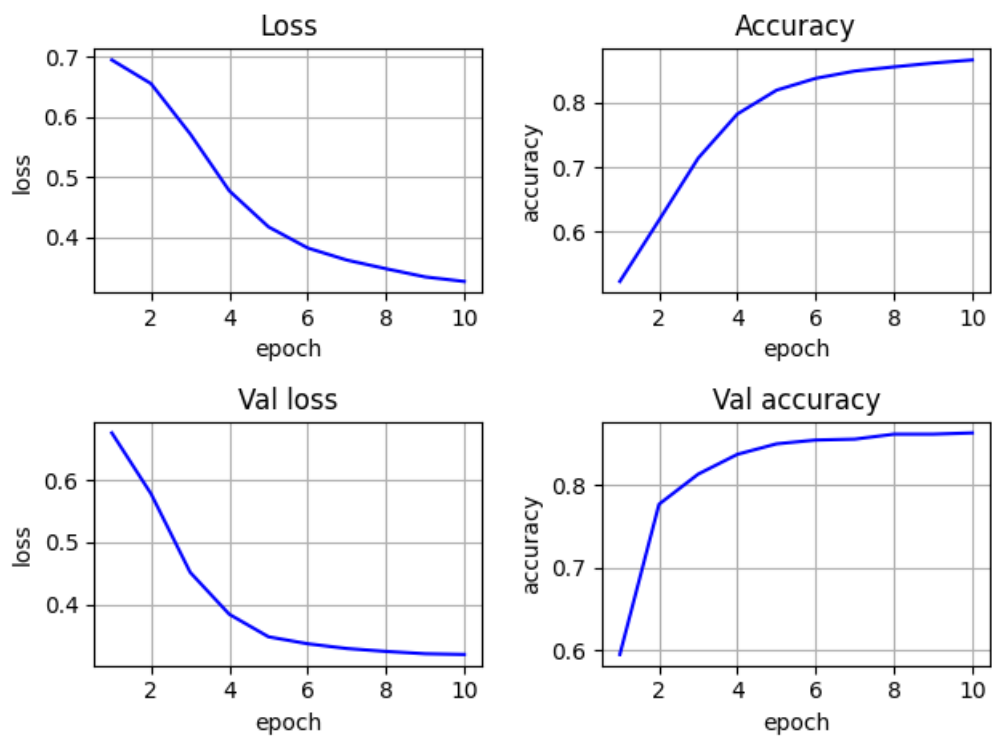
В модели используется оптимизатор `adam`, функция потерь `binary_crossentropy` и метрика `accuracy`.

Рассмотрим результаты обучения модели на разном количестве слов.

batch_size: 3000

epochs: 10

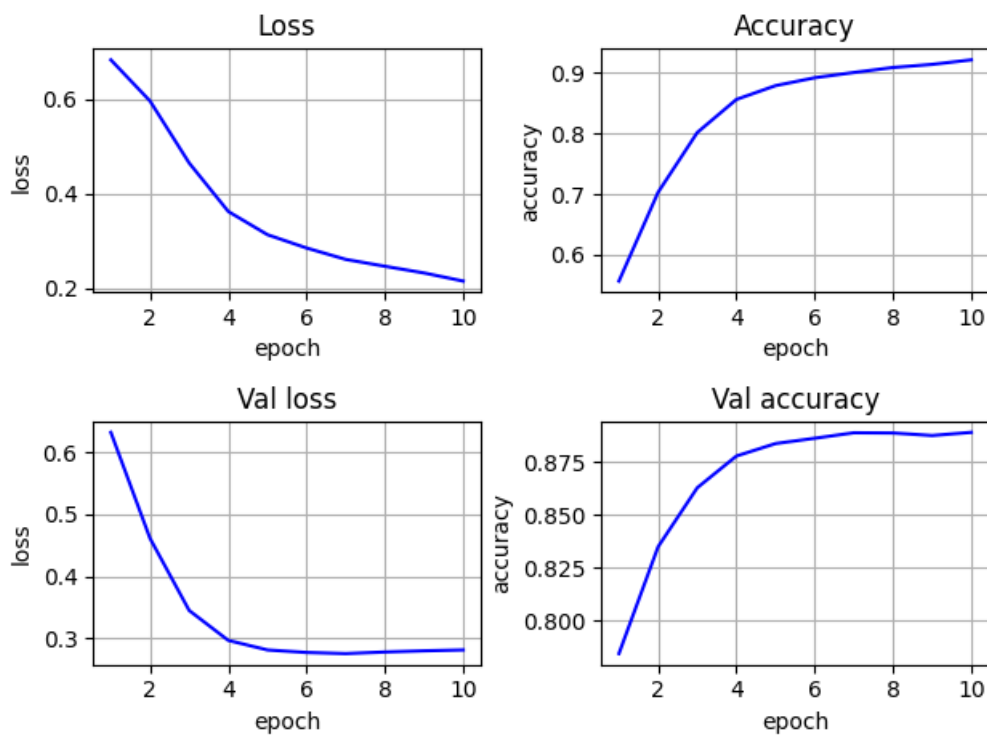
1. 1000 слов



Потери: 0.3202

Точность: 0.863

2. 3000 слов

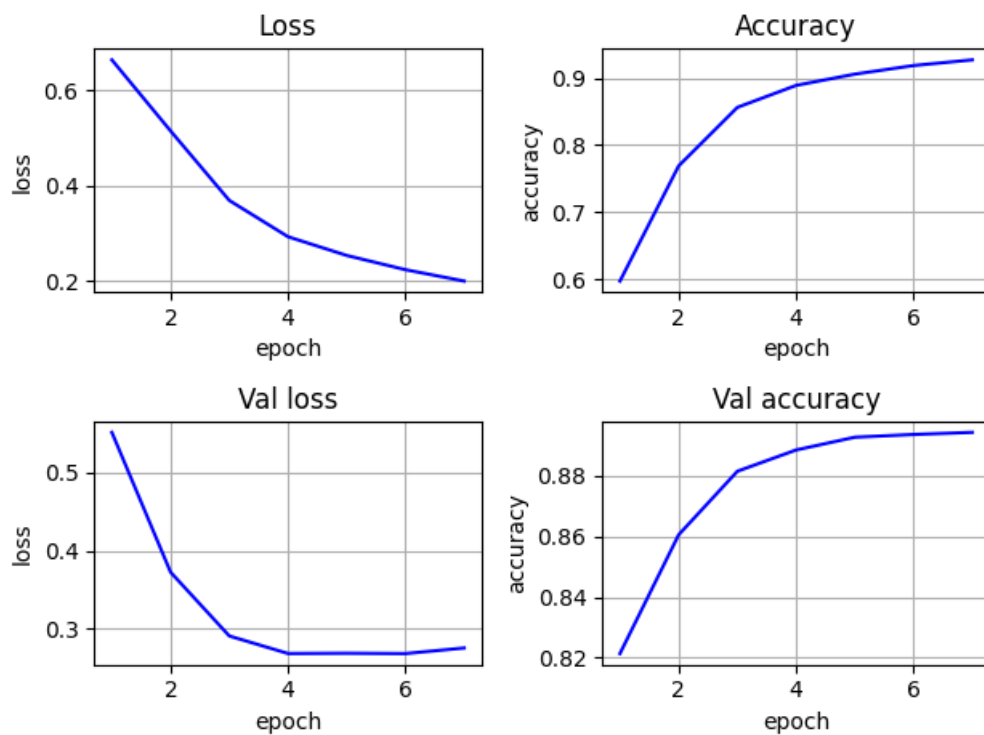


Потери: 0.28077802062034607

Точность: 0.88919997215271

Из графиков видно, что модель начинает переобучаться примерно к 7 эпохе. Уменьшим количество эпох до 7.

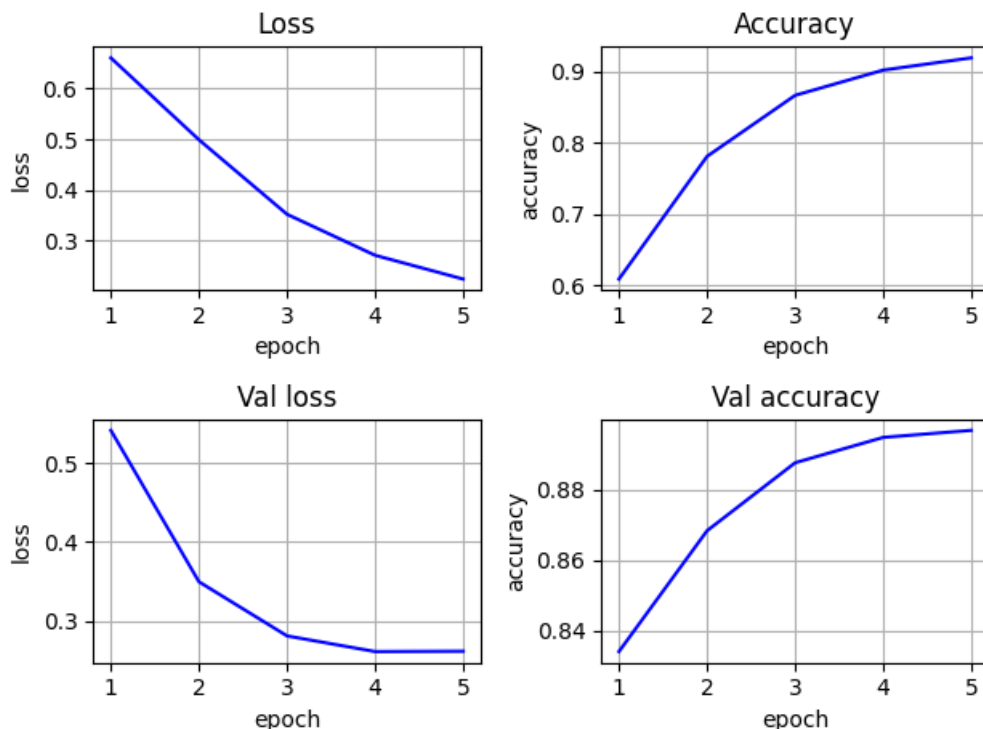
3. 6000 слов.



Потери: 0.27568045258522034

Точность: 0.8942000269889832

4. 10000 слов.



Потери: 0.26129764318466187

Точность: 0.8967999815940857

Видно, что с повышением количества слов точность растёт незначительно, и при этом модель обучается значительно быстрее. При этом эпоха, после которой начинается переобучение модели, наступает также раньше.

Также были взяты 2 реальных отзыва из онлайн кинотеатров.

good/1.txt

In the bins of your sideboards, for sure, there are plates from the service bought once by your grandmother or even great-grandmother. The plates do not fit the rest of your dishes at all, and by rearranging them from place to place when cleaning, you annoy yourself that you cannot have the courage to throw out unnecessary trash.

Now imagine that these plates are the only thing that connects you to the past. With parents, childhood, home. A home that no longer exists and most likely never will.

Fern and her husband lived in the small town of Empire. There was nothing remarkable about it - only one drywall factory, around which the city grew. But during the crisis, the company went bankrupt, Fern's husband died of an illness, and the woman

was suddenly left without a home, without a job and a livelihood.

Fern has no choice but to load a few things dear to his heart in the car and hit the road. A path that has no destination.

On the way, she will meet many people - good and different. She will go through difficult trials and discover new truths. And we will watch this calm, unhurried, somewhere philosophical action without any action, special effects and drama. I can't take my eyes off.

It would seem that the American theme of the film will surely find a response in the heart of the domestic audience to the tips of the nails. And we have many small towns with one city-forming enterprise, the closure of which means the collapse of life, at least the old one. But the mentality and fear of the future will not allow most of us to go so boldly towards the new and unknown. And this is more bad than good, in my opinion.

The film is just great. It allows you to think about a lot, and - to look differently at seemingly familiar things, people who are circling. Learn to appreciate them a little more.

Результат: 0.64357394

bad/1.txt

Like a movie, 'Crimea' is a lazy, clumsy hack. From the point of view of shooting, everything is bad: the scale of events in Ukraine is not felt because of about 10 locations for the entire film, the camera angles are as banal and inventive as possible, as if a first-year journalist was filming. The editor in his skills did not go far from the operator: the transition between scenes is carried out only by darkening, color correction in different scenes, sometimes even at the same location, is not combined with each other, and the frame is flat separately, as if the frame was not painted at all. The director's disregard is clearly seen in the ridiculous scene with the attack of the revolutionaries on the Berkut and in the choreography of the battle in general. The writers lazily prescribed the heroes of the film, sometimes without giving motivation for the actions shown in the film, a huge number of plot holes and illogicalities, and the dialogues were written by the forty-year-old loader Vasily, there are so many cliches that can be found in the federal media. As a result, the actors do not know who to play and do their job poorly. A waste of money for ordinary citizens who want to see a movie.

Результат проверки: 0.03804237

ВЫВОДЫ

В ходе выполнения лабораторной работы было изучено решение задачи классификации отзывов на фильмы для определения отношения

пользователя к фильму. Рассмотрен способ передачи текстовых данных в нейронную сеть, влияние слов-индикаторов на результат работы ИНС. Обучена нейронная сеть, осуществляющая предсказания о настроении отзыва с точностью 89%.