МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«РОССИЙСКИЙ БИОТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ (РОСБИОТЕХ)»

Институт промышленной инженерии, информационных технологий и мехатроники

Кафедра «Информатика и вычислительная техника пищевых производств»

Направление:

**ЛАБОРАТОРНАЯ РАБОТА № 19-21**

*на тему:*

«СЕНТИМЕНТ-АНАЛИЗ ВЫСКАЗЫВАНИЙ

С ПОМОЩЬЮ РЕКУРРЕНТНЫХ LSTM»

Вариант № 7

|  |  |  |
| --- | --- | --- |
| Выполнил: |  | Студент 4 курса, гр. 21-ИУ-3  Морозова М.П. |
| Проверил: |  | Ящун Т.В. |

**Цель работы:** изучить теорию рекуррентных нейронных сетей, способы предобработки текстовой информации для задачи сентимент-анализа, а также научиться строить и обучать рекуррентные сети, в том числе, на базе ячеек LSTM.

**ЗАДАНИЕ 1.** СЕНТИМЕНТ-АНАЛИЗ ВЫСКАЗЫВАНИЙ С ПОМОЩЬЮ РЕКУРРЕНТНЫХ LSTM

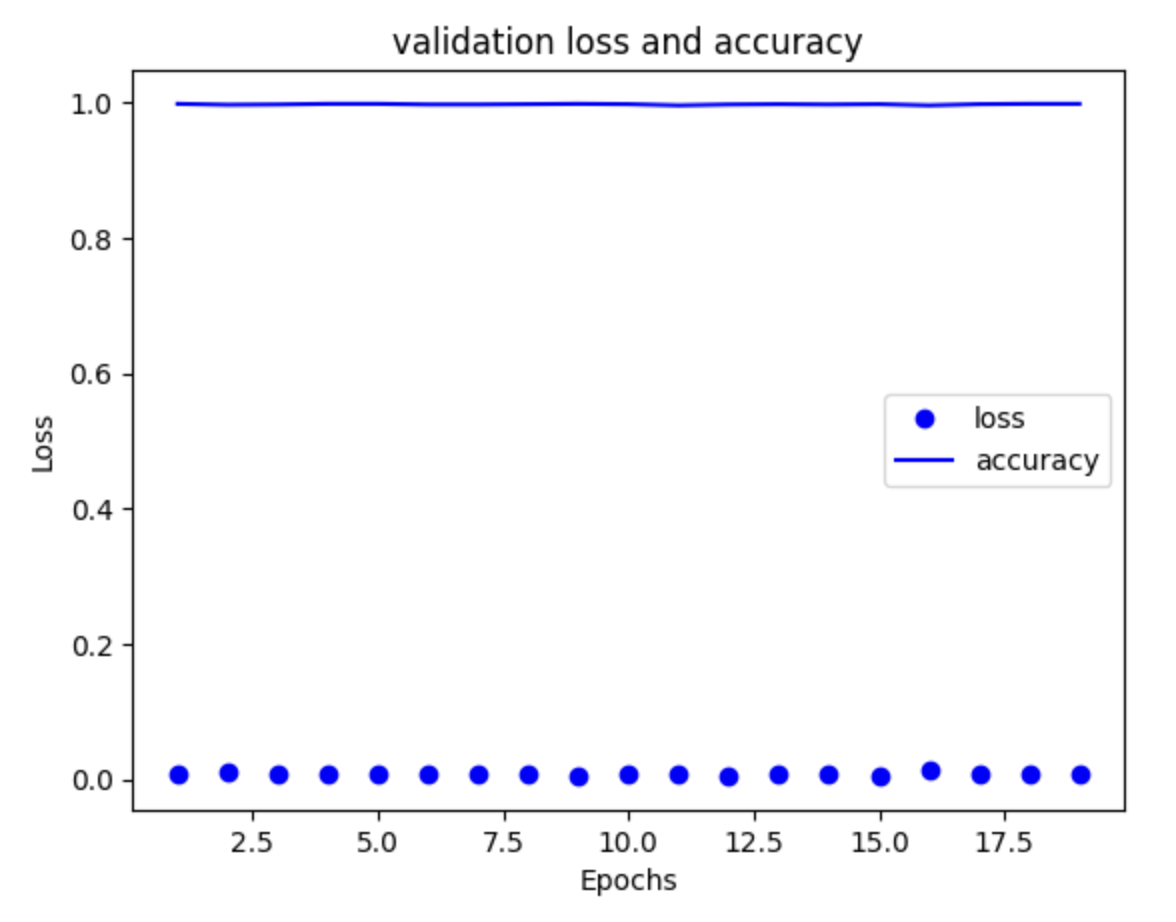
Сравнительная таблица показателей точности работы рекуррентной нейронной сети



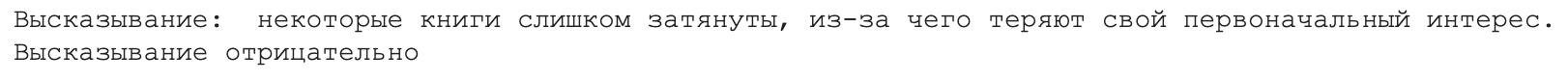
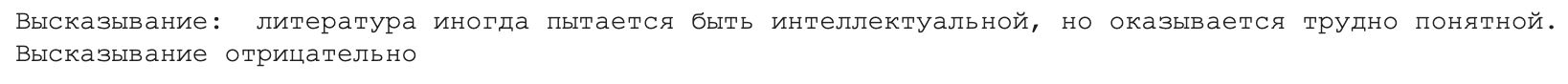
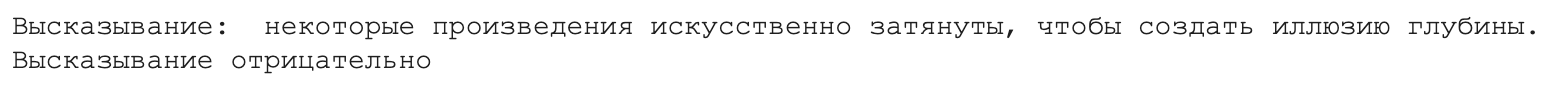
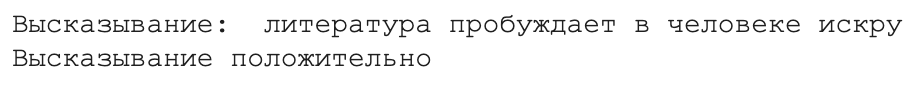
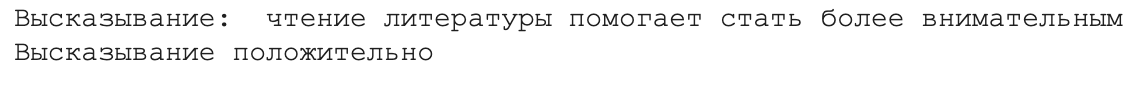
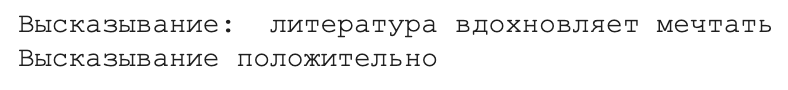
**Выводы:** наличие слоя Dropout не сильно сказалось на результаты работы рнс, больше всего на них повлияли число эпох и размер батча (большое число эпох сказалось негативно, а уменьшение размера батча сказалось положительно).

**Ссылка на GoogleCollab:** <https://colab.research.google.com/drive/1GR5u35YZ3EFQ0RbxGK-UbOBsMHsmn7Cg?usp=sharing>

Графическое отображение потерь и качества прогноза на обучающей выборке для модели с максимальной точностью прогноза:



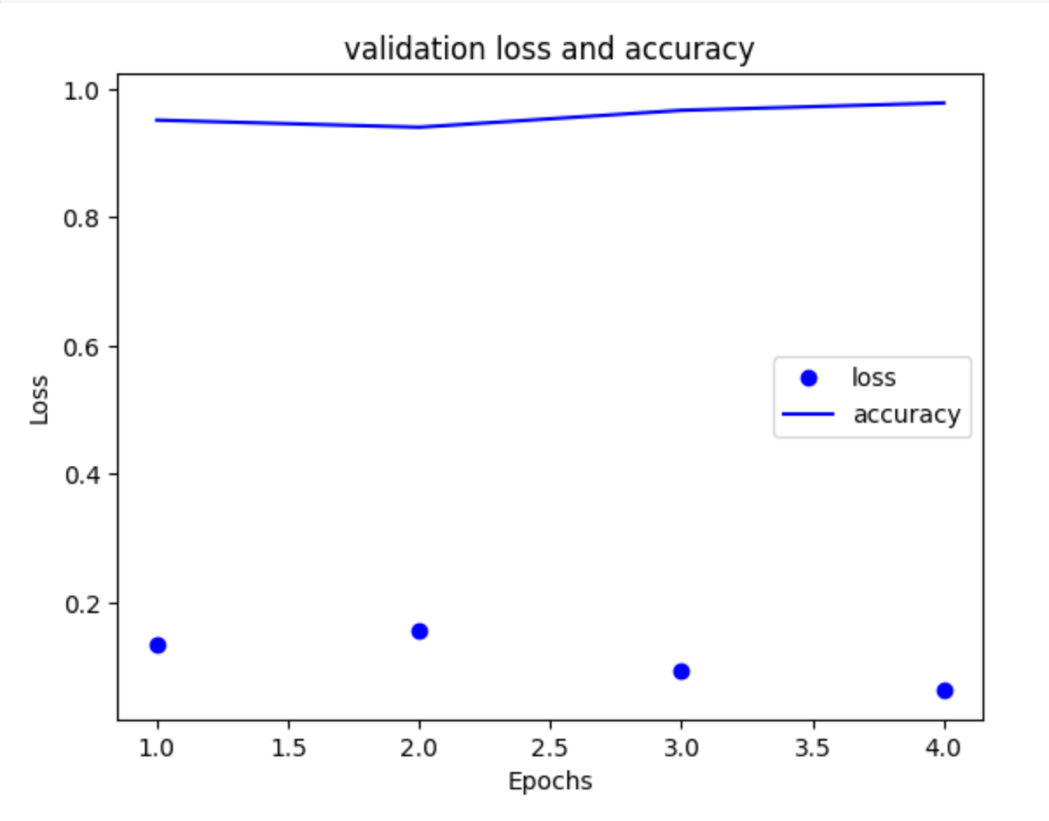
Результаты:



**ЗАДАНИЕ 2.** ПРОГНОЗ УСПЕХА ФИЛЬМОВ ПО ОБЗОРАМ

Сравнительная таблица показателей точности работы рекуррентной нейронной сети

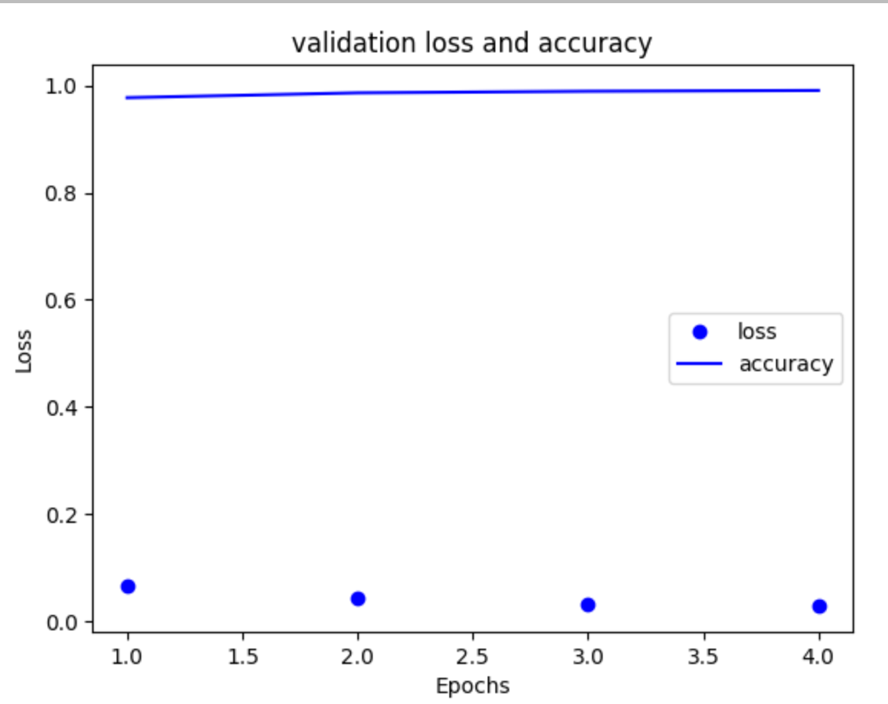




**Выводы:** при уменьшении размера батча и увеличении числа эпох качество обучения как правило возрастает, однако в данном случае подобные изменения необходимо делать с осторожностью, т.к. слишком сильное уменьшение размера батча привело к ухудшению результата.

Сравнительная таблица показателей точности работы альтернативной модели RNN





**Выводы:** использование дополнительного сверточного слоя при тех же параметрах привело к значительному увеличению точности работы и сильно уменьшило количество необходимого времени обучения.

**Ссылка на GoogleCollab:** <https://colab.research.google.com/drive/1nSyV3f5xrybLawHpBZx7DJxAtvtuENOm?usp=sharing>

Сравнение результатов моделирования с помощью двух рассмотренных моделей

|  |  |  |
| --- | --- | --- |
| **Название файла** | **Результаты первой модели** | **Результаты второй модели** |
| Good1 | 0.38446435 | 0.17107698 |
| Good2 | 0.88960457 | 0.6067898 |
| Good3 | 0.5165762 | 0.13151531 |
| Bad1 | 0.99051255 | 0.6078753 |
| Bad2 | 0.95677453 | 0.24136592 |
| Bad3 | 0.98434776 | 0.14800358 |

**Выводы:** вторая модель показывает лучшие результаты по сравнению со первой.