

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

ОТЧЕТ
по лабораторной работе №7
по дисциплине «Искусственные нейронные сети»
Тема: Классификация обзоров фильмов

Студентка гр. 7381

Лящевская А.П.

Преподаватель

Жукова Н.А.

Санкт-Петербург

2020

Цель работы:

Реализовать прогнозирование успеха фильмов по обзорам (Predict Sentiment From Movie Reviews)

Задачи.

1. Ознакомиться с рекуррентными нейронными сетями
2. Изучить способы классификации текста
3. Ознакомиться с ансамблированием сетей
4. Построить ансамбль сетей, который позволит получать точность не менее 97%

Ход работы.

1. Были созданы и обучены две модели искусственной нейронной сети, решающей задачу определения настроения обзора. Первая нейронная сеть рекуррентная с добавлением полносвязных слоев и слоев разреживания. Её архитектура представлена на рис. 1.

Модель представлена на рис. 1.

```
model = Sequential()  
model.add(Embedding(TEST_DIMENSIONS, embedding_vector_length, input_length=max_review_length))  
model.add(LSTM(100))  
model.add(Dense(1, activation='sigmoid'))
```

Рисунок 1 – Модель первой сети

Вторая нейронная сеть рекуррентная с добавлением слоя свертки. Её архитектура представлена на рис. 2.

```

model = Sequential()
model.add(Embedding(TEST_DIMENSIONS, embedding_vector_length, input_length=max_review_length))
model.add(Conv1D(filters=32, kernel_size=3, padding='same', activation='relu'))
model.add(MaxPooling1D(pool_size=2))
model.add(LSTM(100))
model.add(Dense(1, activation='sigmoid'))

```

Рисунок 2 - Модель второй сети

При обучении моделей использовался оптимизатор Adam и функция потерь бинарная кросс энтропия. Результаты обучения первой сети представлены на рис. 3-4. Результаты обучения второй сети представлены на рис. 5-6.

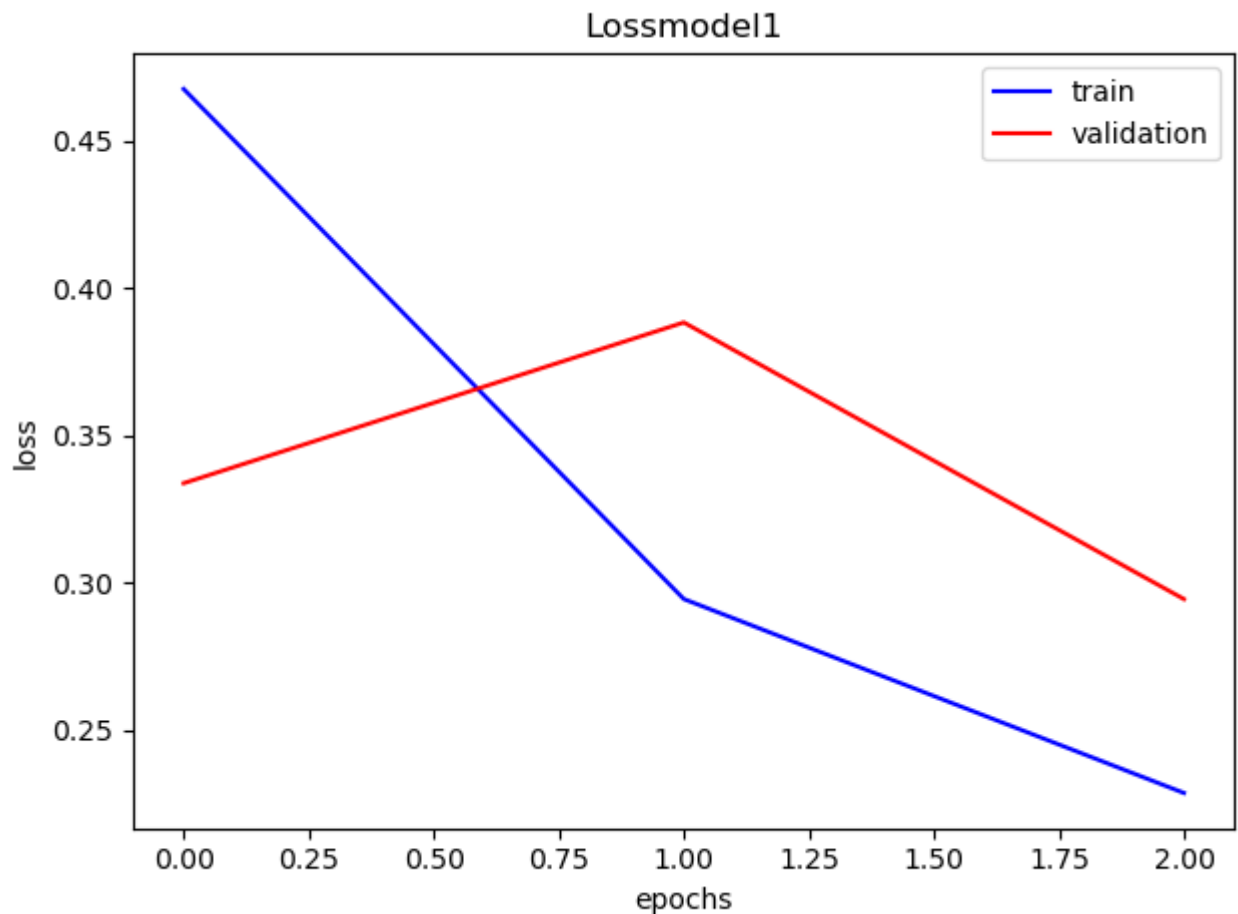


Рисунок 3 - Потери первой сети

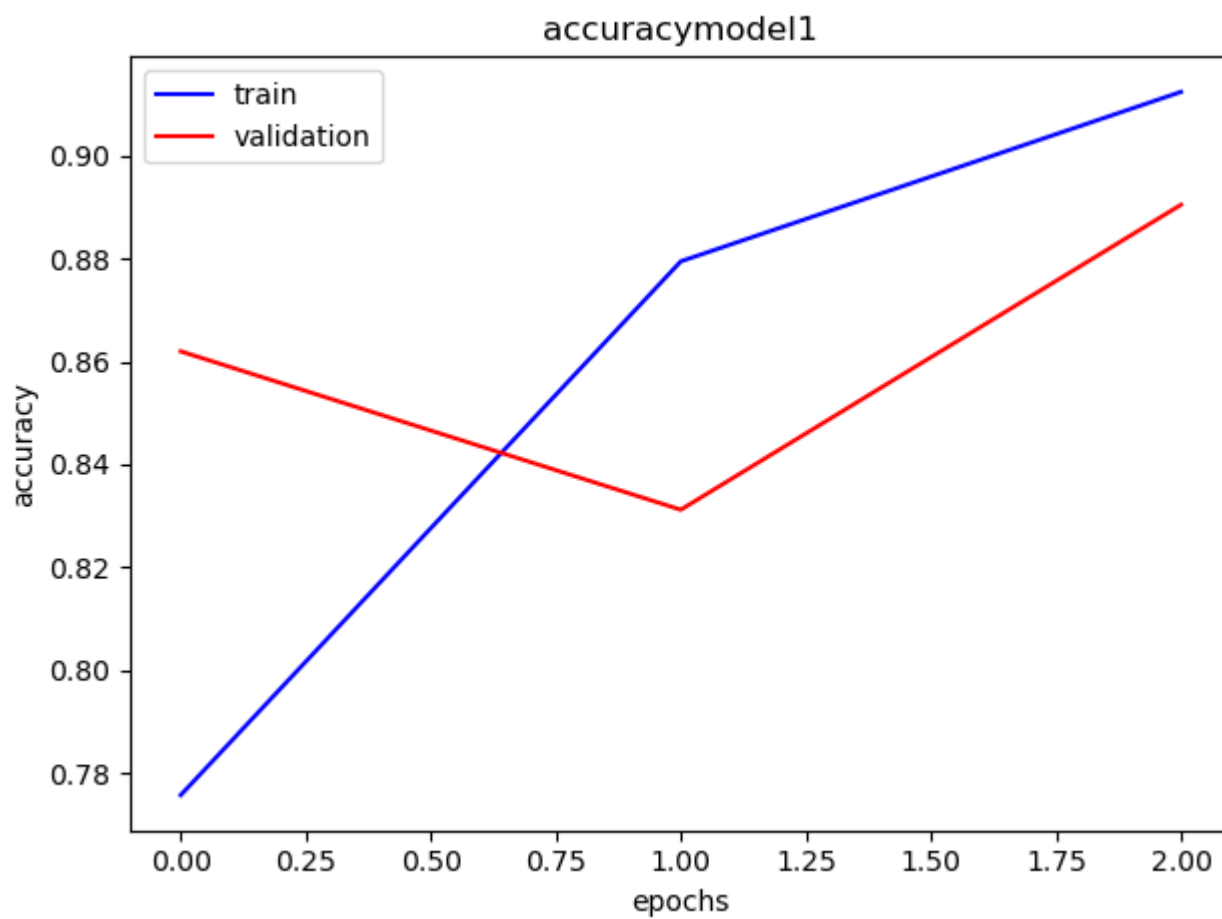


Рисунок 4 - Точность первой сети

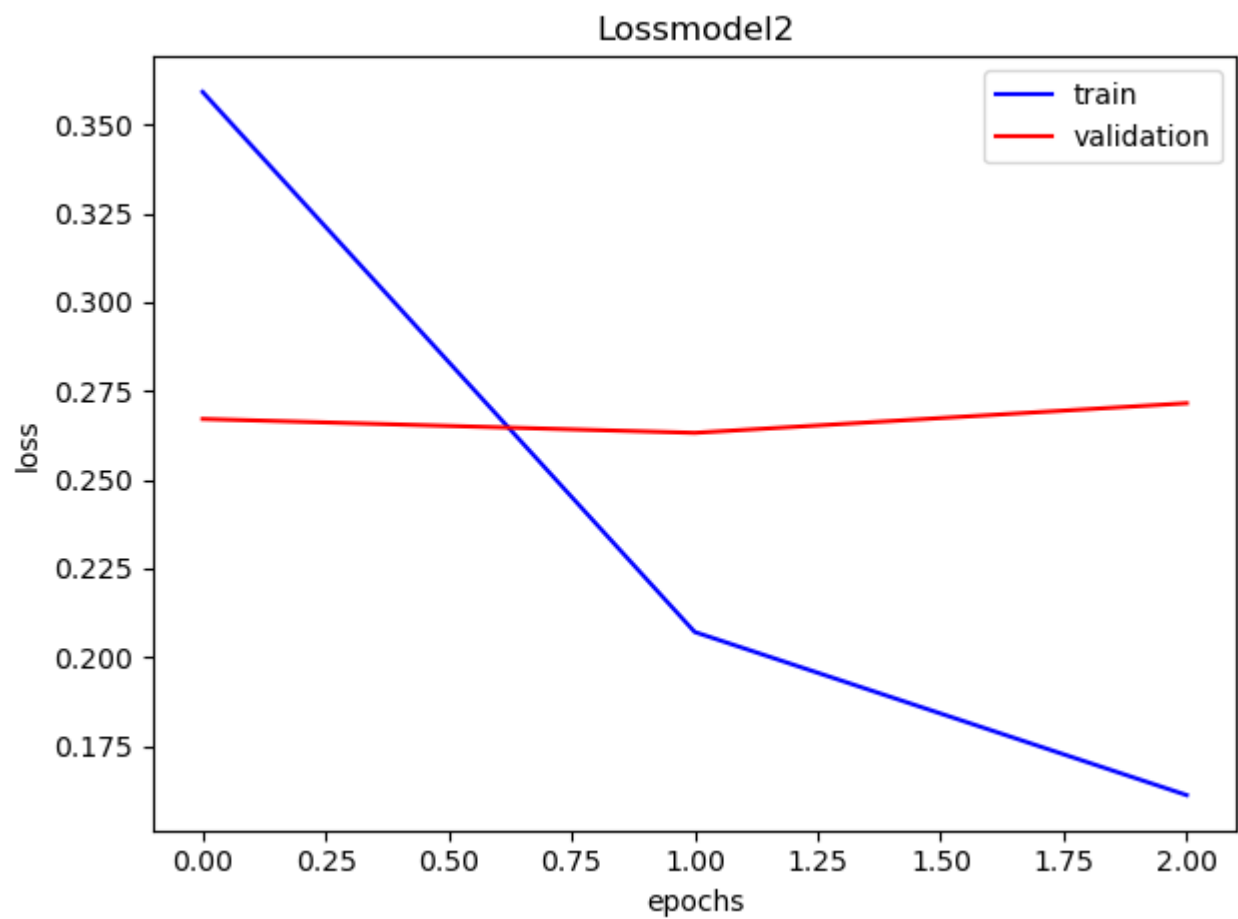


Рисунок 5 - Потери второй сети

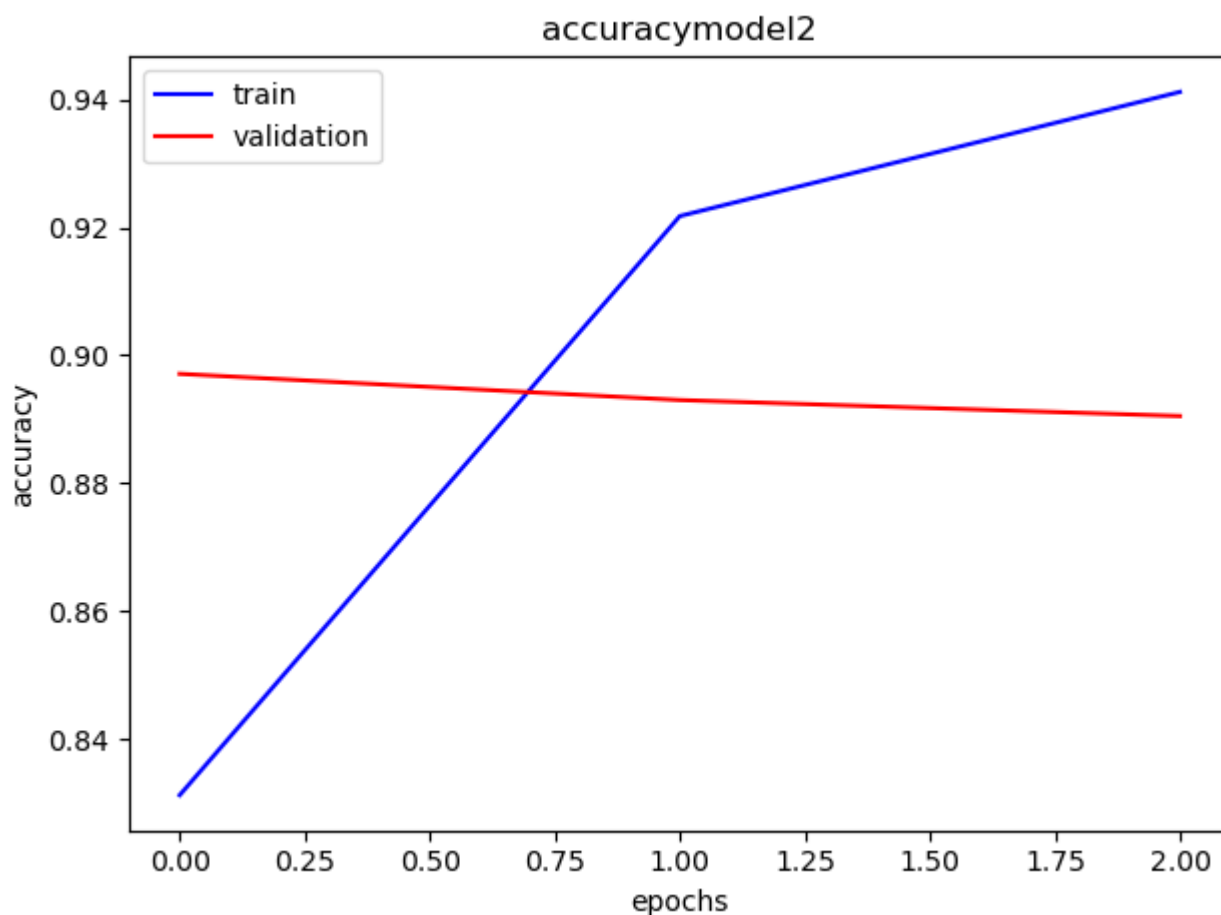


Рисунок 6 - Точность второй сети

- Для ансамблирования моделей была написана функция `ensembling_models`, объединение результатов работы сетей – среднее арифметическое результатов обработки каждого обзора.

Результат работы ансамблирования сетей представлен на рис. 7

```
Accuracy of ensembling models is 0.88724
```

Рисунок 7 - Ансамблирование сетей

- Была написана функция `test_my_text` для загрузки пользовательского текста и прогнозирования успеха фильма по этому тексту. Точность прогнозирования обзоров моделями обеих сетей по отдельности и ансамблированием этих моделей представлена на рис. 8.

```
Validation accuracy of 1st model is 1.0  
Validation accuracy of 2nd model is 0.8333333134651184  
Validation accuracy of ensembling models is 1.0
```

Рисунок 8 - Точность на пользовательском тексте

Как видно по точности, первая сеть показывает лучшие результаты прогнозирования, в отличие от точности моделей на тестовых образах, где все наоборот. Так же видно, что точность ансамблирования модели получилась достаточно высокой.

Выводы:

В ходе выполнения данной лабораторной работы были построены модели сетей, прогнозирующих оценку фильма по обзорам, и проведено ансамблирование этих моделей. Также была написана функция прогнозирования оценки по пользовательскому тексту с помощью ансамблированных моделей.