

**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»**

Студент гр. 8383

Переверзев Д.Е.

Преподаватель

Жангиров Т.Р.

Санкт-Петербург

2021

## Цель работы

Классификация последовательностей - это проблема прогнозирующего моделирования, когда у вас есть некоторая последовательность входных данных в пространстве или времени, и задача состоит в том, чтобы предсказать категорию для последовательности.

Проблема усложняется тем, что последовательности могут различаться по длине, состоять из очень большого словарного запаса входных символов и могут потребовать от модели изучения долгосрочного контекста или зависимостей между символами во входной последовательности.

В данной лабораторной работе также будет использоваться датасет IMDb, однако обучение будет проводиться с помощью рекуррентной нейронной сети.

## Задачи

- Ознакомиться с рекуррентными нейронными сетями
- Изучить способы классификации текста
- Ознакомиться с ансамблированием сетей
- Построить ансамбль сетей, который позволит получать точность не менее 97%

## Ход работы

Для выполнения работы использовались модели двух типов: первый тип содержит сверточный слой, второй такового не содержит.

Первая модель:

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 500, 32)	320000
conv1d (Conv1D)	(None, 500, 32)	3104
max_pooling1d (MaxPooling1D)	(None, 250, 32)	0
dropout (Dropout)	(None, 250, 32)	0
lstm (LSTM)	(None, 64)	24832
dense (Dense)	(None, 1)	65

Total params: 348,001

Trainable params: 348,001

Non-trainable params: 0

## Вторая модель:

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 500, 32)	320000
lstm_1 (LSTM)	(None, 500, 64)	24832
lstm_2 (LSTM)	(None, 32)	12416
dense_1 (Dense)	(None, 1)	33
Total params: 357,281		
Trainable params: 357,281		
Non-trainable params: 0		

## Проверка сети на собственных данных

1. Текст из файла «1»  
оценка: *отрицательная*  
Результат: отрицательный
  2. Текст из файла «2»  
оценка: *положительная*  
Результат: положительный
  3. Текст из файла «3»  
оценка: *отрицательная*  
Результат: отрицательный
  4. Текст из файла «4» ,  
оценка: *положительная*  
Результат: положительный
- Результаты правильные, сеть работает правильно.

## **Выводы**

В ходе лабораторной работы был реализован прогноз успеха фильмов по обзорам. Также был изучен и применен метод ансамблирования моделей для более точного семантического анализа текста.