

**МИНОБРНАУКИ РОССИИ**  
**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ**  
**ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**  
**«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)**  
**Кафедра Математического обеспечения и применения ЭВМ**

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

Студентка гр. 8382

Рочева А.К.

Преподаватель

Жангиров Т.Р.

Санкт-Петербург

2021

## **Цель.**

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

## **Задание.**

- Ознакомиться с рекуррентными нейронными сетями
- Изучить способы классификации текста
- Ознакомиться с ансамблированием сетей
- Построить ансамбль сетей, который позволит получать точность не менее 97%
- Найти набор оптимальных ИНС для классификации текста
- Провести ансамблирование моделей
- Написать функцию/функции, которые позволят загружать текст и получать результат ансамбля сетей
- Провести тестирование сетей на своих текстах (привести в отчете)

## **Ход работы.**

1. Были созданы следующие модели нейронных сетей: рекуррентная нейронная сеть (RNN), сверточная нейронная сеть (CNN) и сверточная рекуррентная нейронная сеть (CRNN). Код программы представлен в приложении А.

2. Обучим три модели при одинаковых параметрах обучения и проведем их ансамблирование. Оценим точность нейронных сетей и их ансамблей. Графики точностей для моделей представлены на рис. 1.

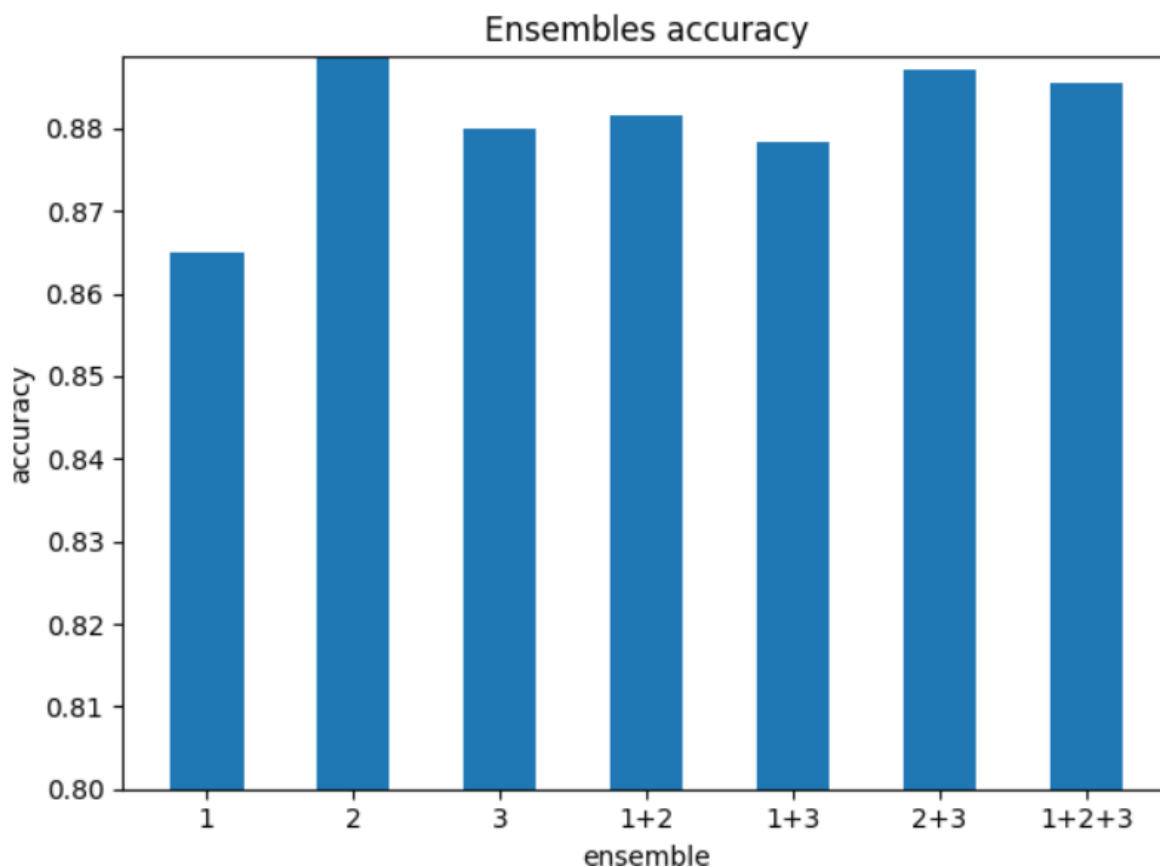


Рисунок 1 – График точностей моделей и их ансамблей

Значения точностей из графика: 86.5% RNN, 88.9% CNN, 88.0% CRNN, 88.2% RNN + CNN, 87.8% RNN + CRNN, 88.7% CNN + CRNN, 88.6% RNN + CNN + CRNN. Все модели и их ансамбли дают примерно одинаковую точность, однако, наибольшую точность более 0.88 дает CNN модель и ансамбль CNN и CRNN моделей.

3. Была написана функция, которая позволяет загружать пользовательский текст из файла. Получим результат ансамбля сетей, загрузив следующий отзыв: The acting was amazing and the film was good overall but I think 'masterpiece' and 'film of the year' are a bit overused throughout the reviews. In no way did I dislike

this film, I thought it was really good, just overrated. Ответ ансамбля сетей: 0.7085517, что говорит о том, что данный отзыв относится к положительному.

### **Выводы.**

В результате выполнения данной работы были созданы несколько моделей нейронных сетей и проведено их ансамблирование. Также были оценены точности каждой из моделей и ансамблей и получен результат оценки пользовательского отзыва ансамблем.