**ВСТУПЛЕНИЕ**

Здравствуйте, меня зовут Мышкин Артём и тема моей дипломной работы «Распознавание речи с помощью нейронных сетей», мои научные руководители Долбещенкова Наталья Вячеславовна и Кожекина Елена Николаевна.

**СЛАЙД 1**

Целью моей работы было исследование параметров рекуррентной нейронной сети для распознавания человека по голосу и преобразования в текст

Объект исследования была рекуррентная нейронная сеть для распознавания человека по голосу

Предмет исследования выступает модель искусственной нейронной сети для распознавания речи человека

Актуальность данной работы заключается в том, что распознавание речи может быть широко использовано в таких областях, как автоматизация офиса, связь, робототехника, поиск информации, а так же может применяться в работе спецслужб.

**СЛАЙД 2**

**Были поставлены такие задачи как**

1) Выбор метода и модели для распознавания человека по голосу и преобразования речевого сообщения в текст

2) Разработка автоматизированной системы обнаружения с преобразованием в текст распознанных голосов

3) Тестирование на реальных голосах, исследование качества распознавания в зависимости от параметров

Актуальностью данной работы заключается в том, что системы распознавания речи в последние несколько лет переживают бурное развитие. Распознавание речи может быть широко использовано в таких областях, как автоматизация офиса, связь, робототехника, поиск информации и спецслужб

**СЛАЙД 3**

Для данной работы была выбрана рекуррентную нейронная сеть, т.к она является универсальной. Она одновременно хорошо обрабатывает текст, звук, видео и изображение

Если описывать схематично, слой RNN использует цикл for для итерации по упорядоченной по времени последовательности, храня при этом во внутреннем состоянии, закодированную информацию о шагах, которые он уже видел.

**СЛАЙД 4**  
В данной работе были использованы библиотеки спич рекогнейшн, сис вебброус и птсХ.

Функция Ток отвечает за озвучивание, а функция Команд - за преобразование речи в текст.

**СЛАЙД 5**

Вы можете видеть на слайде результаты работы моей программы. На данной картинке мы видим значок микрофона на панели уведомлений, что свидетельствует об активации микрофона. Программа распознаёт то, что ей говорят, после чего повторно озвучивает команду, для уточнения правильности распознания. Далее можно увидеть работы программы по поиску в браузере по запросу «Картинки лягушки».

**СЛАЙД 6**

* Далее была написана программа для распознавания речи.

В программе используется рекуррентная нейронная сеть и библиотеки PyTorch. Программа сравнивает и распознаёт переданные ей файлы, содержащие записи голоса. Вначале файлы обрабатываются для удаления шумов из записи. Прописывается рекуррентная нейронная сеть. Кросс-энтропия измеряет расхождение между двумя вероятностными распределениями.

В данной работе используется усовершенствованный оптимизатор Адам, запускающий адаптивную оценку моментов.

Задаём значение коэффициента, масштабирующего дельту перед применением к параметрам. После чего программа выводит результаты своей работы

**СЛАЙД 7**

Для сравнения были выбраны основные методы оптимизатора Adadelta, Adagrad, Adam, AdamW, Adamax, ASGD. На графиках представлена зависимость процента успеха от количества циклов для тестовых и тренировочных значений.

Оптимизатор Adadelta показал медленный рост процента оттренированных нейронов и нуждается в большем количестве циклов.

Adagrad Работает стабильно, однако, имеет низкий процент успеха для тестовых и тренировочных значений.

Оптимизатор AdamDef достиг 100% успеха, однако, в середине работы происходит переобучение.

**СЛАЙД 8**  
AdamW имеет выброс в начале работы, после чего его работа стабильная, но имеет средние показатели.

Adamax- Имеет высокий показатель оптимизации, но через каждые 400 циклов происходит переобучение, что было подтверждено при увеличении их общего количества.

ASGD имеет самые низкие показатели оптимизации. Он не подходит для решения данных задач.

**СЛАЙД 9**  
AdamME-

На основании тестовых экспериментов, было определено, что для работы с этим видом данных больше всего подходит оптимизатор Адам. С помощью изменения значение коэффициента, масштабирующего дельту, было устранено переобучение в начале работы, за счет чего удалось достигнуть высоких и стабильных показателей.   
На слайде можно увидеть график сравнения лучшей и худшей модели оптимизатора, а также таблицу с их результатами.

**Слайд 10**  
Вывод

1. Разработана и реализована автоматизированная система распознавания голоса
2. Реализована модель нейронной сети позволяющая распознать человека по голосу
3. Тестовые эксперименты с использованием различных оптимизаторов показали что, наилучшие показатели были у усовершенствованной модели Адам, наихудшее ASGD