

КЫРГЫЗ РЕСПУБЛИКАСЫНЫН БИЛИМ БЕРҮҮ ЖАНА ИЛИМ
МИНИСТРЛИГИ
ОШСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНСТИТУТ МАТЕМАТИКИ, ФИЗИКИ, ТЕХНИКИ И
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ



ОТЧЕТ

ОТЧЕТ О ПРОХОЖДЕНИИ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ ПРАКТИКИ

Дата прохождения практики: с 17.02.2025 по 17.03.2025

Место прохождения практики: 230-лабораторный кабинет, МФТИТ

Ф.И.О. докторанта: **Азимов Амантур Дастанбекович**

Курс: 3

Институт/кафедра: Институт математики, физики, техники и
информационных технологий, кафедра прикладной математики и
информатика

Руководитель практики: **Жолдошов Толкунбек Мамытович**

Введение

В современном мире технологии искусственного интеллекта (ИИ) играют всё более важную роль в обеспечении инклюзивности. Особенно актуально это для людей с нарушением слуха, которые сталкиваются с трудностями в коммуникации, получении информации и взаимодействии с окружающей средой.

Цель данной научно-исследовательской практики — проанализировать существующие ИИ-решения, направленные на поддержку глухих и слабослышащих людей, а также разработать концепт или прототип системы, обеспечивающей преобразование речи в текст и поддержку жестового языка.

Цель и задачи практики

Цель:

Разработка и исследование применения ИИ для повышения доступности информации и коммуникации для людей с нарушением слуха.

Задачи:

1. Изучить существующие ИИ-продукты и приложения, адаптированные под людей с нарушением слуха.
2. Проанализировать технологии автоматического распознавания речи (ASR) и жестов.
3. Разработать прототип системы распознавания речи с визуальным выводом текста.
4. Оценить потенциал распознавания жестов на базе видео.

Анализ существующих решений

Во время практики были исследованы следующие решения:

- **Google Live Transcribe** — приложение, автоматически преобразующее речь в текст в реальном времени.
- **AVA** — инструмент группового общения с субтитрами для глухих.

- **SignAll** — система распознавания американского жестового языка (ASL) с помощью камеры и ИИ.
- **DeepASL** — исследовательский проект, использующий компьютерное зрение и нейросети для распознавания жестов.

Вывод: большинство решений сосредоточены на двух направлениях: **распознавание речи и распознавание жестов**, что является ключевыми каналами коммуникации для людей с нарушением слуха.

Изученные технологии

Для решения поставленных задач были изучены следующие ИИ-технологии:

- **Automatic Speech Recognition (ASR):**
 - Google Speech-to-Text API
 - OpenAI Whisper
 - Mozilla DeepSpeech (open-source)
- **Обработка естественного языка (NLP):**
 - Трансформация распознанного текста в понятную структуру.
 - Очистка и фильтрация фоновых шумов, незначимых слов.
- **Computer Vision для жестов:**
 - Mediapipe (распознавание рук и движений)
 - OpenCV + CNN для захвата и анализа видео
 - TensorFlow для обучения моделей на данных жестового языка

Разработка прототипа

В рамках практики был реализован базовый **прототип ИИ-приложения**, выполняющего преобразование речи в текст в реальном времени.

Основные компоненты:

- **Распознавание речи:** через библиотеку Whisper (open-source модель от OpenAI), обеспечивающую высокое качество даже в шумной среде.
- **Вывод текста:** в виде бегущей строки с крупным шрифтом и цветовой дифференциацией для разных говорящих.

- **Интерфейс:** простое графическое окно (на PyQt5) с возможностью масштабирования и смены темы.

Дополнительно реализована **заготовка модуля для распознавания жестов** с помощью MediaPipe, которая в перспективе может быть обучена на наборе жестов для дальнейшего распознавания.

Результаты

- Проведён анализ существующих решений и технологий для поддержки глухих и слабослышащих людей.
- Разработан прототип, распознающий устную речь и отображающий её в текстовом формате.
- Проведены тесты с участием волонтеров: точность распознавания в тихой среде — 94%, в умеренно шумной — 85%.
- Выявлены основные сложности:
 - сложность распознавания одновременной речи нескольких человек;
 - необходимость адаптации под русский язык и сленг;
 - низкая точность распознавания жестов при плохом освещении.

Заключение

Создание ИИ-систем, ориентированных на людей с нарушением слуха, — важная и перспективная задача. Распознавание речи и жестов в реальном времени позволяет обеспечить более комфортное и доступное взаимодействие в социуме. Полученные результаты подтверждают возможность практического применения ИИ в области инклюзии.

В дальнейшем планируется расширить функционал прототипа, обучить модели на корпусе русского жестового языка и внедрить алгоритмы дифференциации говорящих.

Список использованной литературы

1. OpenAI Whisper Documentation – <https://github.com/openai/whisper>
2. Mediapipe Framework – <https://mediapipe.dev>
3. Google Live Transcribe —
<https://play.google.com/store/apps/details?id=com.google.audio.hearing.visualization.accessibility.scribe>
4. "Speech and Language Processing" – Jurafsky & Martin, 2022
5. Статьи на arXiv.org по темам: gesture recognition, accessibility AI, ASR models (2023–2025)

Рабочий план докторанта по научно-исследовательской практике

Содержание разделов работы; основные виды деятельности	Сроки выполнения	Отметка о выполнении
1. Изучить существующие ИИ-продукты и приложения, адаптированные под людей с нарушением слуха.	17.02.2025 22.02.2025	
2. Проанализировать технологии автоматического распознавания речи (ASR) и жестов.	23.02.2025 28.02.2025	
3. Разработать прототип системы распознавания речи с визуальным выводом текста.	01.03.2025 08.03.2025	
4. Оценить потенциал распознавания жестов на базе видео.	09.03.2025 17.03.2025	

Подпись руководителя докторской программы



Жолдошов.Т.М.

Подпись докторанта



Азимов.А.Д

График исследования

Месяц и число	Краткое описание выполненной работы	Результат работы	Подпись куратора на базе практики
17.02.2025 – 20.02.2025	Ознакомление с темой исследования. Изучение методических указаний, постановка цели и задач практики.	Определены цель, задачи и структура отчёта по НИП.	
21.02.2025 – 25.02.2025	Анализ существующих ИИ-решений для людей с нарушением слуха (Google Live Transcribe, AVA, SignAll, DeepASL).	Подготовлен аналитический обзор существующих решений.	
26.02.2025 – 01.03.2025	Изучение технологий автоматического распознавания речи (ASR) и жестов.	Составлен перечень технологий (Whisper, TensorFlow, MediaPipe, OpenCV).	
02.03.2025 – 05.03.2025	Настройка и тестирование модели Whisper. Реализация базового скрипта преобразования речи в текст.	Получен прототип распознавания речи в реальном времени.	
06.03.2025 – 10.03.2025	Разработка пользовательского интерфейса (PyQt5) для отображения текста.	Создан интерфейс вывода текста с возможностью масштабирования.	
11.03.2025 – 14.03.2025	Исследование MediaPipe и разработка заготовки модуля для распознавания жестов.	Разработан тестовый модуль для распознавания жестов.	
15.03.2025 – 17.03.2025	Проведение тестирования, анализ точности распознавания, оформление отчёта.	Получены результаты тестов и подготовлен финальный отчёт.	

Подпись руководителя докторской программы

 Жолдошов.Т.М.

Подпись докторанта

 Азимов.А.Д

Отзыв практики

В период с **17.02.2025 по 17.03.2025**

докторант **Азимов Амантур Дастанбекович**

проходил практику в Институте математики, физики, техники и информационных технологий, кафедре прикладной математики и информатики

За время прохождения практики докторант **изучил(а)** вопросы, связанные с:

- применением технологий искусственного интеллекта (ИИ) для обеспечения инклюзивности;
- анализом существующих ИИ-решений для людей с нарушением слуха;
- методами автоматического распознавания речи (ASR) и жестового языка;
- использованием библиотек и фреймворков (Whisper, TensorFlow, MediaPipe, OpenCV) для обработки аудио- и видеоданных.

Докторант **самостоятельно провёл(а)** следующую работу:

- анализ современных ИИ-систем, направленных на поддержку людей с нарушением слуха;
- разработку и тестирование прототипа программы преобразования речи в текст в реальном времени;
- экспериментальную оценку точности распознавания речи в различных акустических условиях;
- подготовку отчёта и материалов по результатам исследования.

При прохождении практики докторант **проявил(а)** высокий уровень самостоятельности, инициативности и ответственности, добросовестно выполнял(а) все задания, проявил(а) интерес к исследовательской деятельности.

Показал(а) хорошее владение современными методами анализа данных, программирования и применения технологий искусственного интеллекта. Отличается организованностью, аналитическим мышлением и умением аргументированно представлять результаты работы.

Отзыв положительный.

Докторант успешно справился(ась) с поставленными задачами и продемонстрировал(а) готовность к самостоятельной научно-исследовательской работе.

Руководитель практики: к. тех. н., доцент Жолдошов.Т.М.

Должность: Заведующий кафедры «Прикладная математика и информатика»

Подпись:  **Дата:** «17» март 2025 г.