

Перехват и
обработка
текста
с удаленного
микрофона



НАД ПРОЕКТОМ РАБОТАЛА

Наша команда



Артем Саратовцев

Высший магистр



Александра Земцова

Высший маг



Роберт Битлев

Специалист по Arduino (гений)



Арина Бурдина

Эксперт по шифрованию



Шиянов Александр

Менеджер



Мотивация

В наше время судопроизводство является затратным на ресурсы. Наш проект берется за решение этой проблемы посредством создания устройств для записи и анализа текста. Помимо судов, устройства данного типа могут использоваться во многих других областях, таких как продажи и медицина.



Цели и задачи



Цель проекта

Реализовать перехват аудио с последующей его обработкой и преобразованием в текст с помощью алгоритмов машинного обучения.



Задачи проекта

- Разработать и собрать двух устройств, оснащенных микрофонами для передачи и приема звуковых сигналов;
- Реализовать шифрования передаваемой информации;
- Интегрировать алгоритмов машинного обучения для преобразования полученной информации в текст.

КАК ЭТО МОЖНО ПРИМЕНИТЬ

Гипотетические применения



В суде



В работе операторов

Для оценки KPI сотрудников.



В здравоохранении

Для записи и анализа
медицинских консультаций.



В полиции



В образовательных учреждениях

Для записи выступлений при сдаче
дипломных и проектных работ.



В бизнесе

Для автоматизации процессов и
улучшения коммуникации между
сотрудниками.





Две ключевые части



Arduino и компоненты

Создание MVP устройств, для приема и передачи звука.



Python

Преобразование полученных звуковых данных с помощью алгоритмов ML в текст и последующая его шифрование и обработка.

Этапы работы с Arduino



1. Подключение и настройка микрофона. Устранение ошибок;
2. Подключение и калибровка радио-модуля;
3. Добавление шифрования;
4. Доработка и калибровка шифрования.

Микрофон МАХ9814

В ходе работы использовался микрофон МАХ9814. Данный микрофон преобразует звуковые сигналы в аналоговые с достаточно большой амплитудой, что позволяет точно считывать звук.

В ходе записи в .wav файл возникает следующая проблема: стандартная частота дискретизации .wav файла составляет 44.1 кГц, тогда как частота получения сигнала с Arduino – 10 кГц. Это означает, что за одну секунду можно считать в 4.1 раз меньше данных, чем нужно для качественной записи. Эту проблему удастся решить, если дополнить уже имеющиеся данные недостающими. Эти недостающие данные мы «восстанавливаем», линейно экстраполируя зависимость громкости между двумя соседними считанными значениями. Например, если у нас были считаны значения громкости a и b , и нам нужно дополнить их еще тремя дополнительными, то их значения будут следующими: $\frac{3a+b}{4}$, $\frac{a+b}{2}$, $\frac{a+3b}{4}$. Таким образом, нам удастся получить все необходимые значения громкости, и записать полноценный .wav файл.



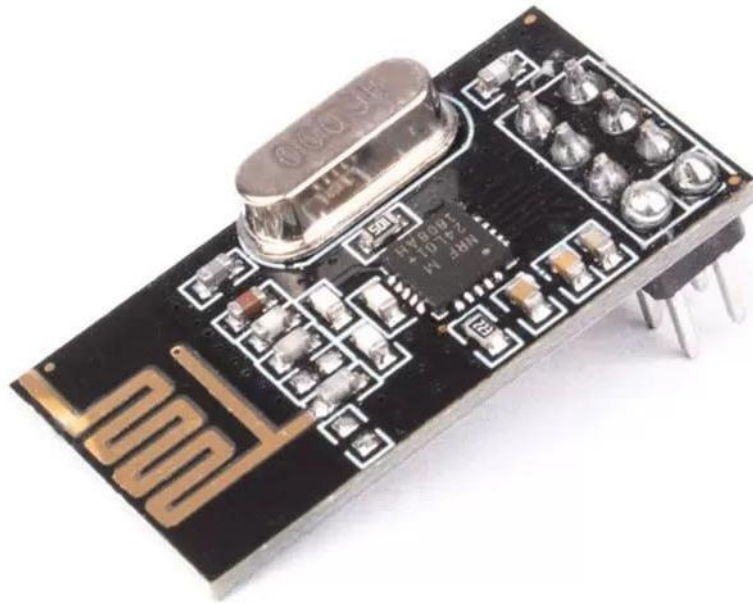
Микрофон

Радио-модуль

Радио-модуль NRF24L01

Радио-модуль работает на максимальном расстоянии в 100 метров, чего более чем достаточно в нашем случае, т.к. устройства в основном будут находиться в одном здании.

Несмотря на трудности в работе, нам удалось откалибровать данные модули и настроить их на передачу достаточного количества данных быстро и точно, без каких-либо потерь.



Микрофон

Радио-модуль

Шифрование

В среде разработки Arduino IDE доступна библиотека Arduino Cryptography Library, которая предоставляет набор инструментов для реализации криптографических алгоритмов. Однако, эта библиотека использует аппаратное ускорение ARM CryptoCell CC310, которое доступно только на nRF52840. Поскольку наш проект основан на использовании nRF24l01, **данная библиотека не может быть применена в нашем случае.**

В настоящее время, этап реализации шифрования и дешифрования сигнала находится в процессе разработки. Мы реализовали процедуру шифрования, основанную на **комбинации циклического сдвига, инверсии определенных битов и других стандартных битовых операций.** Это позволяет нам обеспечить достаточный уровень безопасности для передаваемых данных.



Шифрование

Из речи в текст

Преобразование речи в текст

Whisper - это алгоритм, который использует нейронные сети для распознавания речи и преобразования ее в текст. Он обладает высокой точностью, но может быть медленным при обработке данных.

CMUSphinx - это система автоматического распознавания речи, которая основана на методах статистического распознавания речи. Она хорошо работает с разными языками и может быть использована в автономном режиме, но иногда страдает от низкой точности.

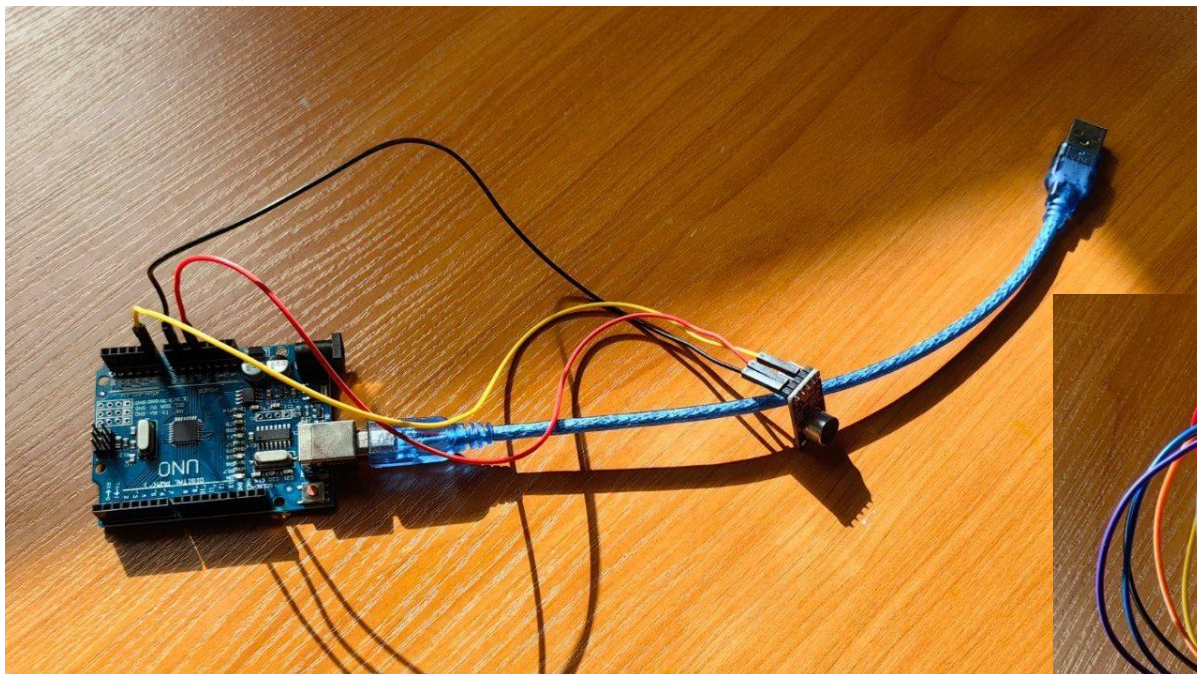
Google API - это набор инструментов, который обеспечивает высокую скорость обработки данных и приемлемую точность, но имеет существенный недостаток: использование сервисов Google требует постоянного подключения к Интернету. Тем не менее, Google API отвечает всем нашим требованиям: является быстрой, точной, простой в использовании и не требующей большого расхода памяти, что особенно важно при работе с устройствами с ограниченными ресурсами.



Шифрование

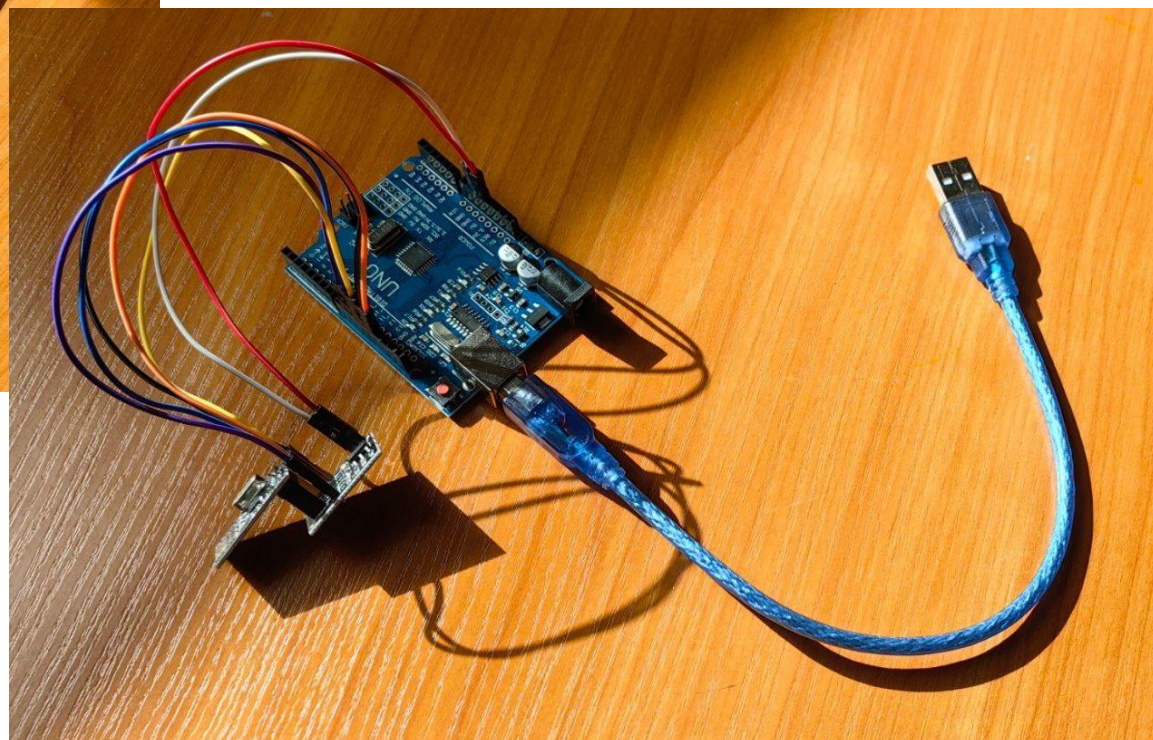
Из речи в текст

Итог работы



Микрофон

Радио-приемник



Перспективы проекта

Наша команда видит весьма обширный спектр аспектов проекта, которые можно бы было улучшить при дальнейшей работе над проектом.

Мы планируем развиваться в следующих направлениях:

- Более сложный алгоритм шифрования
- Обобщение и анализ смыслового содержания текста с помощью машинного обучения
- Улучшение имеющегося оборудования для более высокого качества записи
- Внедрение локально работающей модели, не требующей подключения к интернету и обращения к сторонним сервисам.



