



ACTIVE
BUSINESS
CONSULT

Распознавание речи

Система позволяет преобразовывать в текст разговорную речь. Сервис дополнен нейросетевыми моделями отсеивания шумов, распознавания эмоций, определения пола и возраста абонента



Преимущества

- + Нейросетевая модель отсеивания шумов
- + Распознавание эмоций
- + Определение пола
- + Определение возраста
- + Гибкие программные интерфейсы

ПРЕИМУЩЕСТВА



ВОЗМОЖНОСТИ



- Online и offline-преобразование в текст (транскрибирование) спонтанной разговорной речи
- Быстрая адаптация под любые задачи заказчика
- Интеграция с коммуникационной платформой заказчика
- Обучение модели распознавания речи для любой сферы
- Поддержка форматов аудиофайлов WAV, MP3, AIFF, WMA, FLAC, OGG и др.
- Расшифровка записи телефонных разговоров низкого качества
- Обработка многоканальных аудио- и видеофайлов



Цифры

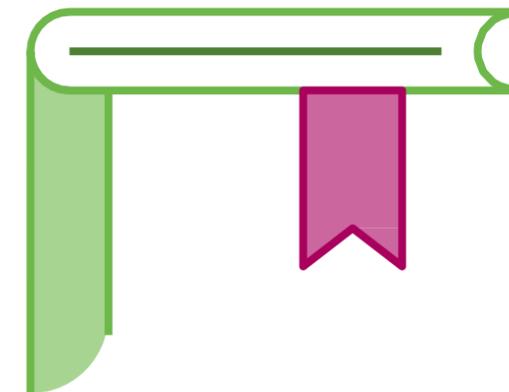


Скорость распознавания
1 секунда речи за 0,3 секунды
процессорного времени

Система распознавания речи
обучена на **1000+ часов**
реальных разговоров



Тематические словари общим объемом
более 2 000 000 слов
с возможностью пополнения



ИУМБЕР



Сферы применения

Протоколирование



Расшифровка телефонных обращений



Голосовое заполнение документов



Call-центры



Smart-гаджеты



Создание субтитров для аудио- и видеоконтента



Интеллектуальный агент
голосовой робот, помощник оператора,
голосовой офисный ассистент



Как это работает



ШАГ 1

Речь абонента поступает в сервер распознавания

ШАГ 2

Для обработки речь разбивается на фреймы по 10 миллисекунд



ШАГ 3

Фреймы с шумом отсеиваются при помощи рекуррентной нейросети



ШАГ 4

Очищенная речь абонента поступает в акустическую модель, которая преобразует последовательность звуковых импульсов в фонемы



ШАГ 5

Поток фонем проходит через лингвистическую модель, представляющую собой конечный автомат для преобразования последовательностей. Таким образом речь преобразуется в текст фразы

ШАГ 6

Для уточнения результатов распознавания используется дополнительная лингвистическая модель, основанная на нейросетевой архитектуре RNNLM



Точность работы



Используя свёрточную нейронную сеть, система способна с точностью до 97% определять пол абонента и его возраст со средней погрешностью в 6 лет



Демонстрация

На демо сервисе АБК можно опробовать наш продукт



Загрузите аудиофайл
с записью голоса



Или сделайте запись
своего голоса

Попробовать

speech-recog-demo.activebc.ru



Под ключ

от 1 месяца

- Оценка объемов речевых данных и параметров звуковых потоков
- Подготовка и настройка оборудования на стороне Заказчика
- Интеграция с коммуникационной платформой заказчика
- Адаптация модели распознавания речи на основе разметки 100-200 часов разговоров
- Интеграция с CRM
- Запуск в эксплуатацию
- Гарантийная и постгарантийная поддержка



Аренда

от 3 дней до 1 месяца

предоставление API

- Оценка объемов речевых данных и параметров звуковых потоков
- Настройка оборудования на стороне АБК
- Предоставление доступа к API (интерфейс)
- Адаптация модели распознавания речи на основе разметки 100—200 часов разговора
- Сопровождение и поддержка

Благодарим за внимание

activebc.ru