Техническое задание

Модуль распознавания голосовых сообщений на смешанном казахско-русском языке с извлечением товаров и количеств

Nov 24, 2026

1. Общие сведения

Модуль встраивается в ИИ-ассистента для продаж, который ведёт переписку с клиентами в WhatsApp, рассчитывает заказ по прайсу продавца и формирует коммерческое предложение. Модуль обрабатывает голосовые сообщения клиентов и передаёт ассистенту структурированный заказ.

2. Цель

Сократить долю голосовых сообщений, которые передаются менеджеру из-за ошибок распознавания, за счёт точного распознавания смешанной казахско-русской речи и извлечения позиций заказа.

3. Функциональные требования

1. Распознавание речи на казахском, русском и смешанном языке, в том числе в пределах одной фразы.
2. Корректное распознавание чисел (цифрами и словами на обоих языках), единиц измерения и сокращений: «уп», «рул», «м²», «пачка», «дана».
3. Извлечение позиций заказа: товар, количество, единица измерения, характеристики (размер, толщина, марка, цвет).
4. Сопоставление позиций с прайсом продавца с учётом синонимов и разговорных названий (например, «вата» → «Минвата 50 мм»).
5. Расчёт степени уверенности для каждой позиции. При уверенности ниже порога модуль возвращает уточняющий вопрос для клиента на языке сообщения.
6. Определение языка сообщения для ответа клиенту.

4. Входные и выходные данные

Вход: аудиофайл (OGG/Opus, до 2 минут), прайс продавца (JSON или XLSX: артикул, наименование, единица, синонимы).

Выход (JSON):

```
{
  "text": "маған 24 пачка вата керек, плюс пароизоляция екі рулон",
  "language": "kk-ru",
  "items": [
    {"sku": "UT-0150", "name": "Минвата 50 мм", "qty": 24, "unit": "уп.",
"confidence": 0.94},
    {"sku": "IZ-0075", "name": "Пароизоляция 75 м²", "qty": 2, "unit":
"рул.", "confidence": 0.91}
  ],
  "clarification": null
}
```

5. Нефункциональные требования

Параметр	Значение
Интерфейс	REST API, JSON
Время обработки	до 5 секунд на сообщение длительностью 30 секунд
Точность извлечения товаров и количеств	не ниже 90% на тестовом наборе заказчика
Нагрузка	не менее 20 одновременных запросов
Размещение	на территории РК или on-premise; аудио не передаётся за пределы РК
Хранение данных	аудио удаляется после обработки, логи без персональных данных
Документация	описание API, инструкция по развёртыванию

6. Что предоставляет заказчик

- обезличенный набор голосовых сообщений с разметкой;
- тестовые прайсы;
- доступ к тестовой среде ассистента;
- консультации команды по сценариям использования.

7. Этапы работ

Этап	Результат	Срок
1. Анализ и прототип	Прототип распознавания, первые замеры точности	3 недели
2. Извлечение и сопоставление с прайсом	API с выдачей позиций и уверенности	4 недели
3. Интеграция и пилот	Подключение к тестовой среде, пилот на реальных сообщениях	4 недели
4. Доработка и приёмка	Отчёт о точности, документация	1 неделя

8. Приёмка

Решение принимается, если на тестовом наборе заказчика (не менее 500 сообщений) выполнены требования раздела 5 по точности и скорости, а API и документация переданы заказчику.