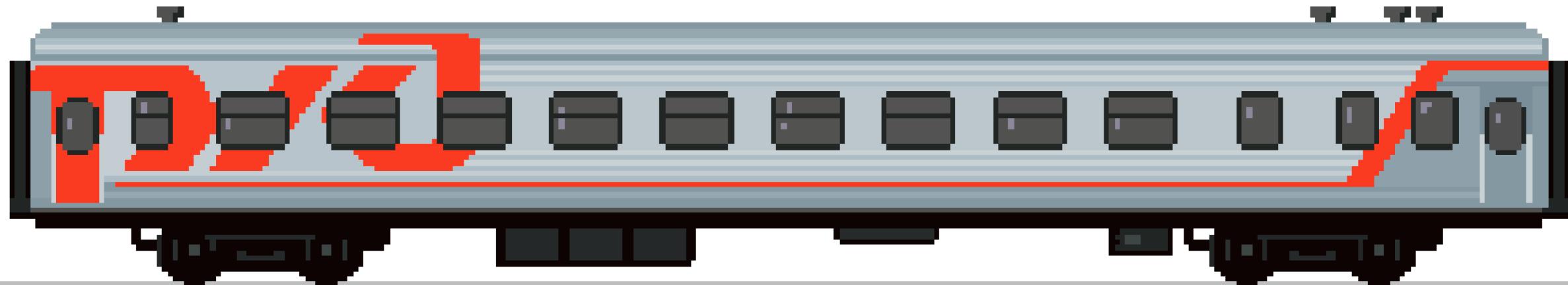


</shamps>

Интеллектуальный пульт составителя с голосовым управлением

🔥 Цифровой прорыв



Ключевые задачи



Защита от шумов

Учет высокого уровня фоновых шумов в голосовой команде



Защита от случайных действий

Исключение случайного ввода команд при ошибочном нажатии на ПСЭ



Валидация команд

Фильтрация и отказ от невалидных команд, не входящих в допустимый перечень



Гибкость в распознавании

Учет небольших вариаций в окончаниях слов, чтобы команды оставались валидными



Работа оффлайн

Автономная работа без необходимости интернет-соединения

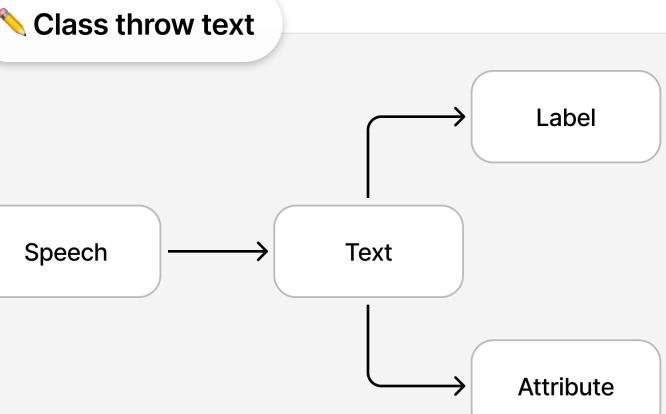


Масштабируемость

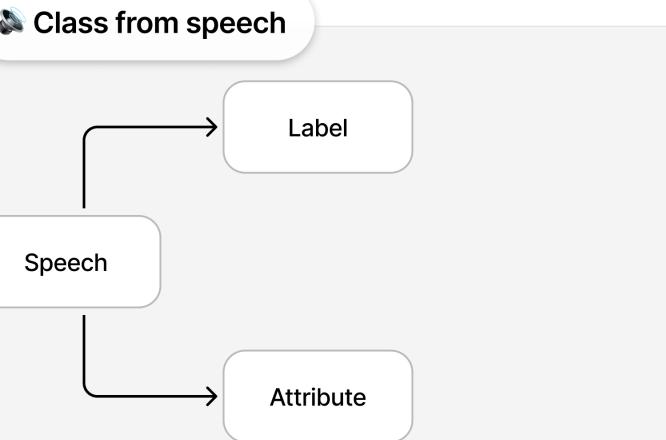
Легкая адаптация под изменения в составе команд

Гипотезы

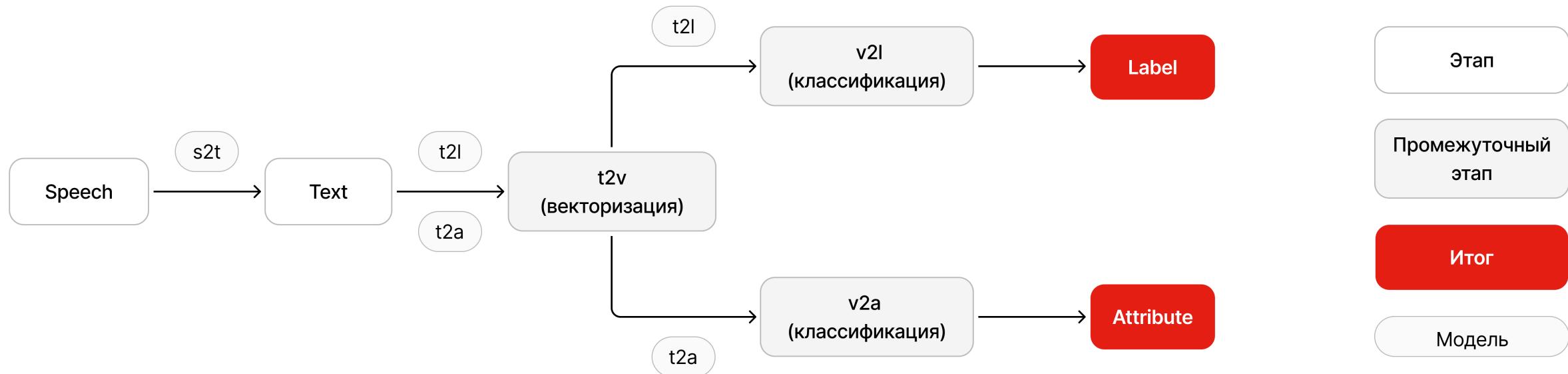
Class throw text



Class from speech



Алгоритм работы решения



Оценка качества

F1

0.81

WER

0.13

Оценка производительности

Оборудование 1

13th Gen Intel i7-13700H (20) @ 4.800GHz

Latency

0.3s

Peak RAM

4GB

Оборудование 2

Intel Xeon Gold 6240R @ 2.4GHz

Latency

3.0s

Peak RAM

860Mb

Disk

7.7G

Используемые модели

s2t

Готовая модель

Предобученная нейронная сеть

Наша модель

Неэффективна из-за малого объема данных
при обучении**t2l + t2a**

Выбор алгоритма обучения

result

CountVectorizer + SVC

F1 = 0.81
WER = 0.13

Hash + RndForest

F1 = 0.76
WER = 0.14

Hash + SVC

F1 = 0.72
WER = 0.18

TF-IDF + MultinomialNB

F1 = 0.54
WER = 0.35

Обоснование выбора технологий



Короткие команды



Ограниченнное количество команд



Оценка метриками

Беспрививочное обновление ПСЭ



Необходимо поддерживать актуальный состав команд без нарушения работы устройств



ПСЭ работает без постоянного подключения к интернету



Невозможно удаленное обновление системы



Сотрудники жд часто работают в отдаленных регионах



Ограничено взаимодействие с IT поддержкой



Возможно изменение или расширение списка команд



Необходимость актуализации системы ПСЭ



Обновление перечня команд без правок исходного кода и без внесения изменений в архитектуру клиентской части ПСЭ.

Повышение точности оценки предобработанной речи



Возможность неправильного распознавания команды



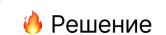
Запись команды может быть инициирована случайно



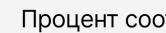
Речь может не совпадать ни с одной из команд в списке, но быть ошибочно воспринята как команда



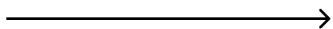
Речь может быть неправильно преобразована в текст из-за очень частичного соответствия команде



Использование метода оценки речи по проценту соответствия в модели Text2Label



Процент соответствия ниже установленного порога



Команда
не подтверждается



Несколько команд имеют схожие процентные значения с незначительными отличиями





Анализ

Решение

Технологии

Киллер-фичи

Showcase

Преимущества

Команда

Преимущества решения



Защищено от случайных команд



Масштабируемо



Совместимо с С и С++



Записи недостаточной или излишней продолжительности отбрасываются



Изменение состава команд не требует изменений в коде — достаточно обновить dataset и переобучить модель



Процесс обучения и использования модели разделен: модель сохраняется в файл и может использоваться независимо от платформы обучения



Оценка записи по проценту соответствия команде: если % ниже установленного порога или есть команды с малыми процентными отличиями, то запись не приводится к команде



Клиентская часть ПСЭ остается неизменной при изменении состава команд — достаточно загрузить новую модель на устройство



Параметризация весов классических моделей ML упрощает совместимость с С и С++



Ваня Тарасов

ML Engineer



Мишель де Джоффрой

ML Engineer



Олег Селянта

ML Engineer



Настя Богданова

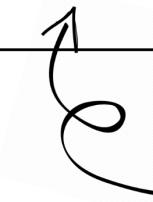
BI Analyst



Вика Кулешова

UX/UI designer

</shamps>



Наш тг-канал 😊