

Кабанчик Колокольчик

21 сентября 2019 г.

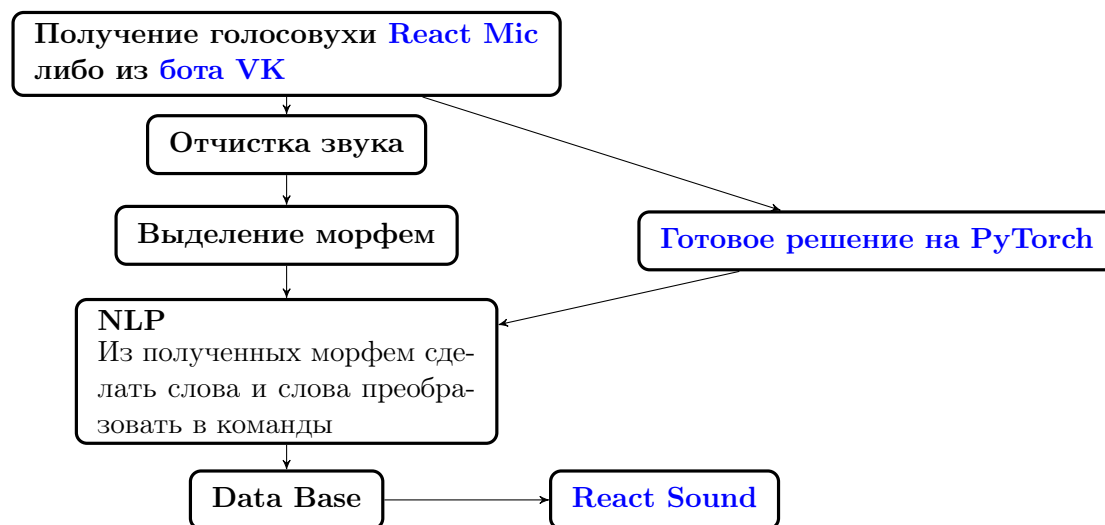
Планы на подготовку

Быстрый старт с [VK Mini App](#). Также всем необходимо стать админами в [VK Mini App](#) и форкнуть [GitHub](#)

1 Распознавание голоса

Задача распознавания голоса является одной из наиболее сложных в проекте, поэтому желательно максимально подготовиться и посмотреть как можно больше методов. В идеале хочется, чтобы на все подзадачи уже имелись готовые решения и датасеты.

Пайплайн будет следующий:



Рассмотрим каждый из пунктов подробнее:

1. В первую очередь необходимо с помощью фронтенда научиться получать голосовуху и отправлять ее на сервер. На данном этапе получилось сделать интерфейс для записи голоса. К сожалению на данный момент запись из VK Mini App не работает, но работает на всех других платформах – это связано с какими-то внутренними ограничениями VK Mini App

В связи с возникшими трудностями было принято решение написать бота VK, этим займется Матвей. Ответом бота будет ссылка на VK Mini App с необходимыми props

2. Голосовое из диалога с ботом необходимо преобразовать к правильному формату, далее разложить сигнал в ДФТ, после чего мы получим картинку, (подробнее в пункте 3. в секции полезные ссылки). Необходимо обучить модель, которая сможет максимально отчистить картинку от шума, довольно подробно о процессе обучения сказано в пункте 2. в секции полезные ссылки. Этим также лучше заняться до хакатона и разберется с этим Саша

3. Отчищенный спектр будет подаваться на вход большой и сложной машине, которая выделяет морфемы из спектрограммы, как это делается описано в пункте 2. в секции полезные ссылки. Очень хотелось бы найти готовое решение этой задачи

4. Кажется, что на вход системе будет подаваться ограниченное число команд, выделяются следующие интенды:

- (a) Как мне добраться до название картины / выхода из музея / туалета / раздевалки
- (b) Расскажи мне про название картины
- (c) Когда ближайшая экскурсия
- (d) Какие сейчас работают выставки

Основная проблема заключается в том, чтобы распознать до какой именно картины хочет добраться пользователь. Ответ на этот вопрос частично дали на ODS. Этим займутся Булат и Саша, при этом очень прошу Булата извлечь максимально сути из дискуссии на ODS и переписать эту суть сюда.

5. **Булат** Переписывает эту суть сюда: Можно попробовать из массивов готовых частей и ифов собрать грамматику составляющих (которая как раз учитывает синтаксическую структуру), которая будет парситься каким-то автоматом. Перед парсингом незнакомые грамматике слова можно конвертировать в знакомые как раз при помощи семантической близости (расстояние между эмбедингами слов). Почему грамматика не хуже других подходов:
- (а) ею можно описать то же множество фраз, что и регулярками, более компактно
 - (б) если хотим машиннообученный классификатор, а данных нет или очень мало, то из этой же грамматики их можно нагенерировать
 - (с) грамматикой можно не только классифицировать тексты, но и собственно парсить их (выделять слоты и всё такое).

Основной минус грамматики: её нужно писать. **Руками**. Ну то есть можно и машинно обучить, если данных много, но если данных много, то грамматика не очень и нужна. Вот тут есть учебный [примерчик](#) бота, основанного на грамматиках.

6. Вот тут то, что касается аппроксимации близких слот-филлов: Для простых случаев подойдут простые текстовые меры близости того, что ввёл пользователь, и эталона (например, 0.8 косинусной близости мешков лемм + 0.2 расстояния левенштейна). Для более сложных случаев понадобится датасет синонимов/переформулировок, специфичный для вашей задачи - скорее всего, основное время уйдёт на его сбор, ну а дальше можно обучать любую модельку, используя эмбединги из предобученных моделей и всё те же текстовые близости как фичи.

Важные ссылки:

1. [ODS dlcourse.ai](#) Не особо подробная лекция, не особо внятный лектор
2. [Выступление человека из МФТИ](#) также в описании к ролику приложена ссылка на [github](#). Необходимо разобраться и потестить как работает их решение
3. [Лекции ФИВТ](#) 21, 23, 25. Довольно подробно разбирается теория, также можно посмотреть домашки по курсу доступные в описании под видео

2 Backend, Матвей

1. Расчеты нейронки будут запускаться при вызове метода `api`, который будет callback-ом возвращать расчеты и результат. Это все сделает Матвей на голлом flask. Если кому-то интересно понимать, что будет происходить на бэке [Курс грустного мужика](#)
2. Асинхронность
3. Карта музея и путь от картины А до картины В. Могут возникнуть следующие трудности:
 - (a) Как передать карту в Front и в каком формате необходимо ее хранить
 - (b) Как отрисовать путь на карте и отслеживать положение пользователя на этом пути.

До хакатона нужно найти решение и желательно попробовать его как-то реализовать

3 Карты

Хотелось бы написать свой путеводитель и под имеющуюся карту Пушкинского музея. Необходимо сделать следующие вещи:

1. Рендерить карты на стороне сервера и отправлять их в VK Mini App
2. Отправлять пользователя в другие приложения из карт, если он хочет уточнить маршрут или еще что-то
3. Получать координаты пользователя
4. Отправлять пользователю Push, когда пользователь проходит рядом с чем-то интересным

Есть пример [реализации карт](#) в vk mini apps, в нем написано как решаются почти все из этих задач, кроме рендеринга собственных карт на сервере

4 Frontend

1. Сделать анимацию ожидания при обработке запроса сервером
2. Экран формы для прохождения теста за стикеры
3. Всплывающая панелька информации о картинах с возможностью прослушать аудио, полистать фоточки и почитать текст
4. Чатик с ботом, где есть две кнопки: записать аудио и сфотографировать QR с помощью [VK UI Connect](#)

5 Design + Презентация

1. Подумать над тем как это все будет выглядеть и нарисовать это в какую-нибудь презентацию.

6 Стикеры

Если найдутся люди, готовые помочь, добавляйте их сюда

1. [Софья-Мария](#)

7 Сервер

1. Найти сервер минимальной стоимости и запустить на нем что-нибудь простое

Успехи

1. Получаем QR с фоточки
2. Попробовали React Mic, работает везде, кроме vk mini apps
3. Попробовал React Sound, работает
4. Сделали всплывающий плеер для управления музыкой
5. Получилось играть поставленную музыку на всех страницах, пока что работает с одним багом на github pages и почти не работает в VK Mini Apps