



Мы Users.

Кейс: ИИ для анализа пользовательских ответов

Проблема

Отображение результатов в сервисе “МТС Линк” есть,
но не всегда информативно



Наше решение. Коротко.

- Кросс-платформенный веб-сервис
- Open-source инструменты
- Высокая точность на CPU
- Масштабируемое благодаря мультипроцессингу
- Фильтрация нерелевантных ответов
- Автодополнение слов
- Альтернативный наглядный способ визуализации

Стек технологий



Transformers



TweetNLP



UNIFORM MANIFOLD
UMAP
APPROXIMATION & PROJECTION



django


Streamlit

Преимущества решения

Пост-анализ результатов.
Получив прогноз,
пользователь может узнать,
какие факторы повлияли
на полученный результат

Универсальный ввод данных.
Можно загрузить
информацию по проекту
как в формате CSV,
так и вручную через форму

Использование
28 различных моделей.
Каждая предназначена
для своего типа задач

Ошибки, эмодзи и новые слова

«ФФ»

«🤖 замечательно»

«ЧИЧМЧ»

«;!”№;34»

«123ww»

«<blink>hello world</blink>»

«повышение зп»

«орг-ть встречу кол-ва с дир-ом»

«признание, благодарность, делеги»

← 30 символов →

Препроцессинг

Пунктуация

Лемматизация

Транслитерация

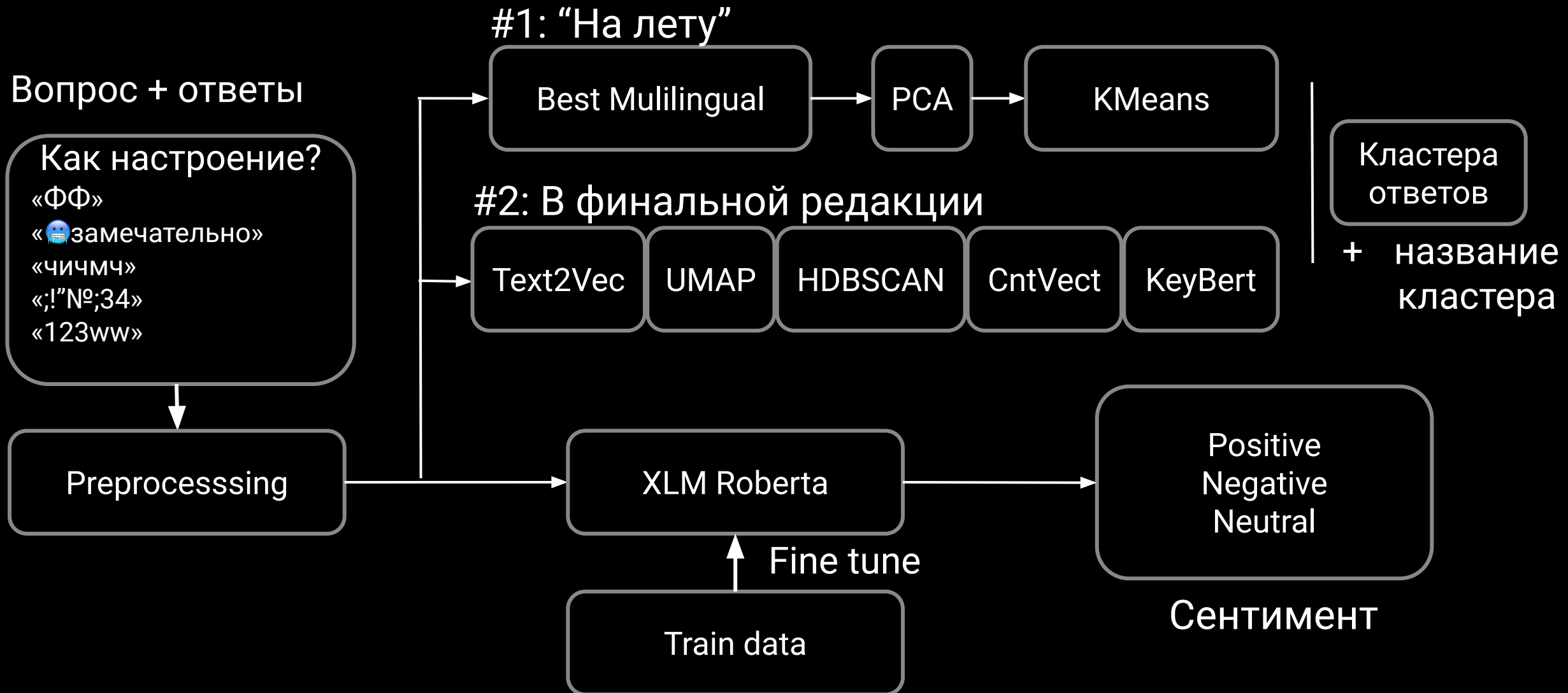
Удаление стоп-слов

Удаление html-тегов

Пайплайн

	F1 Macro	Time
NaiveModel каждый ответ - это новый кластер	10 ms	0.81
LevensteinSimilarityModel > 63%	0.87	102 ms
LevensteinSimilarity + Processing Lemmmatization, delete punct	0.89	300 ms
SelfClusterModel#1 + SentimentTransformer (Bert-Multilingual + PCA + KMeans) + (TweetNLP + xlm-roberta-multilingual)	0.92	2100 ms
SelfClusterModel#2	0.94	1300 ms
SelfClusterModel#2 + SentimentTransformer	0.97	1850 ms

Процесс обработки. Настоящая магия



Демонстрация решения



Коротко о результатах

Генерируем название
СМЫСЛОВЫХ ТЕМАТИК ОТВЕТОВ

Работаем с **10 языками**
de, en, es, fr, it, nl, pl, ru, zh

Понимаем **не только текст** 🤗

Два режима:
“На лету” и в финальной редакции

Быть, а не казаться.



Post /api/qaitems

Отправление
ответов пользователей



Get /api/qaitems

Получение
кластеров и сентиментов



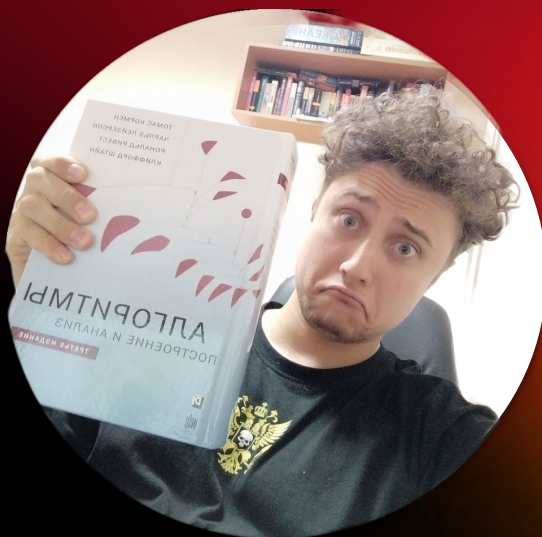
Возможные улучшения

Новые форматы
ввода данных

Несколько сценариев
в прогнозе

Учет других
факторов риска

Команда Users



Вадим
Оленников



Артем
Тарасов



Вероника
Мурзина



Кристина
Корнеева

Поделитесь с нами обратной связью,
и мы, Users, ответим на вопросы

