



приоритет2030^
лидерами становятся



Анализ результатов EXIT-опроса

Команда “Геймоваторы”

17-20 октября 2024

Анна Николаева
Артём Носов
Александр Архангельский
Яна Кузьмина

К чему и зачем идем?

Проблема

Сотрудники увольняются не только из-за зарплаты и карьерных перспектив, но и по менее очевидным причинам. Текущие методы анализа анкет не выявляют эти причины эффективно, что затрудняет снижение текучести кадров.

Задачи:

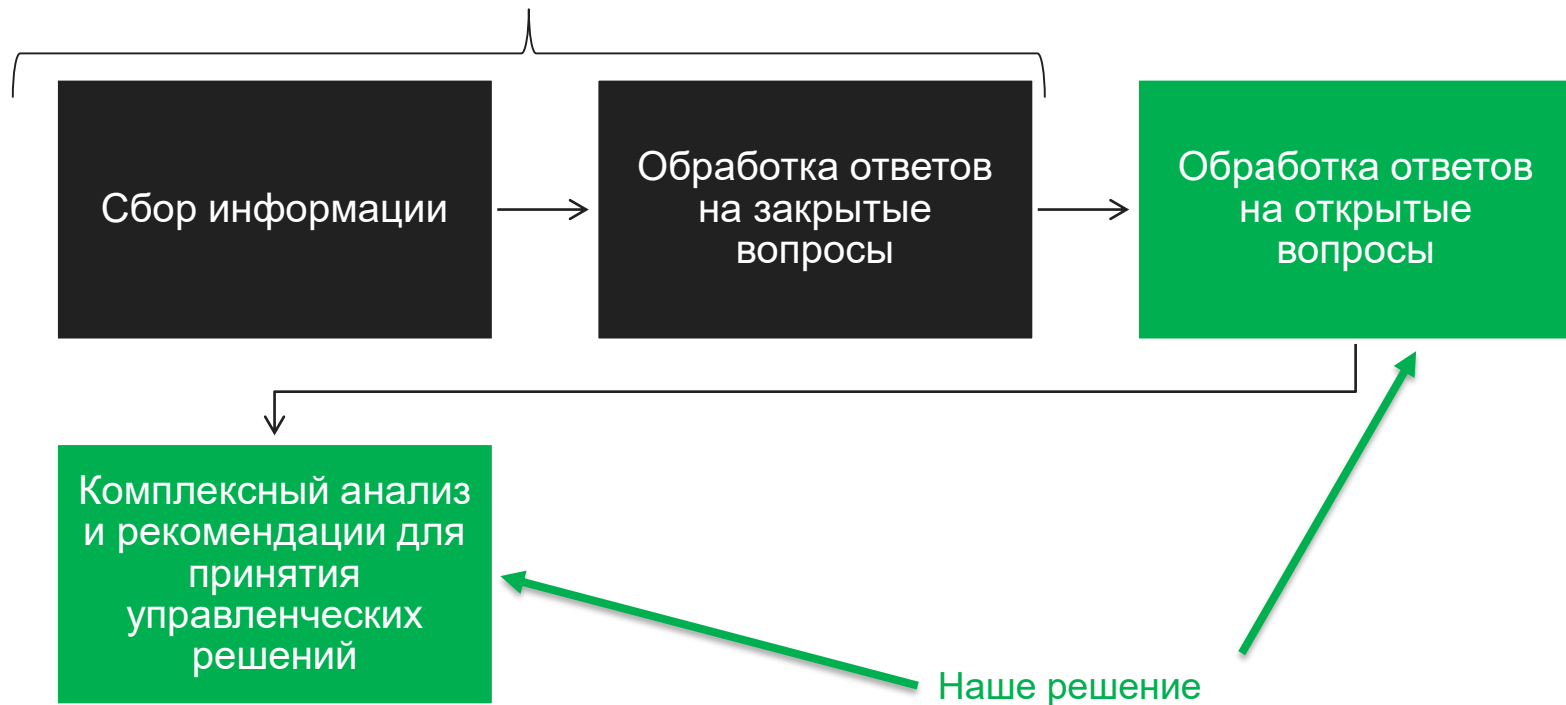
- Импорт данных и очистка текстов
- Разделение ответов на группы по причинам увольнения.
- Диаграммы и графики по результатам кластеризации.
- Автоматическая генерация PDF-отчёта.
- Проверка корректности и улучшение программы.

Основная цель

За 24 часа разработать программу для автоматического анализа и кластеризации открытых ответов из выходных анкет, визуализировать данные и сгенерировать отчёт.

Процесс интервью

KitBot, Hrmessenger, др



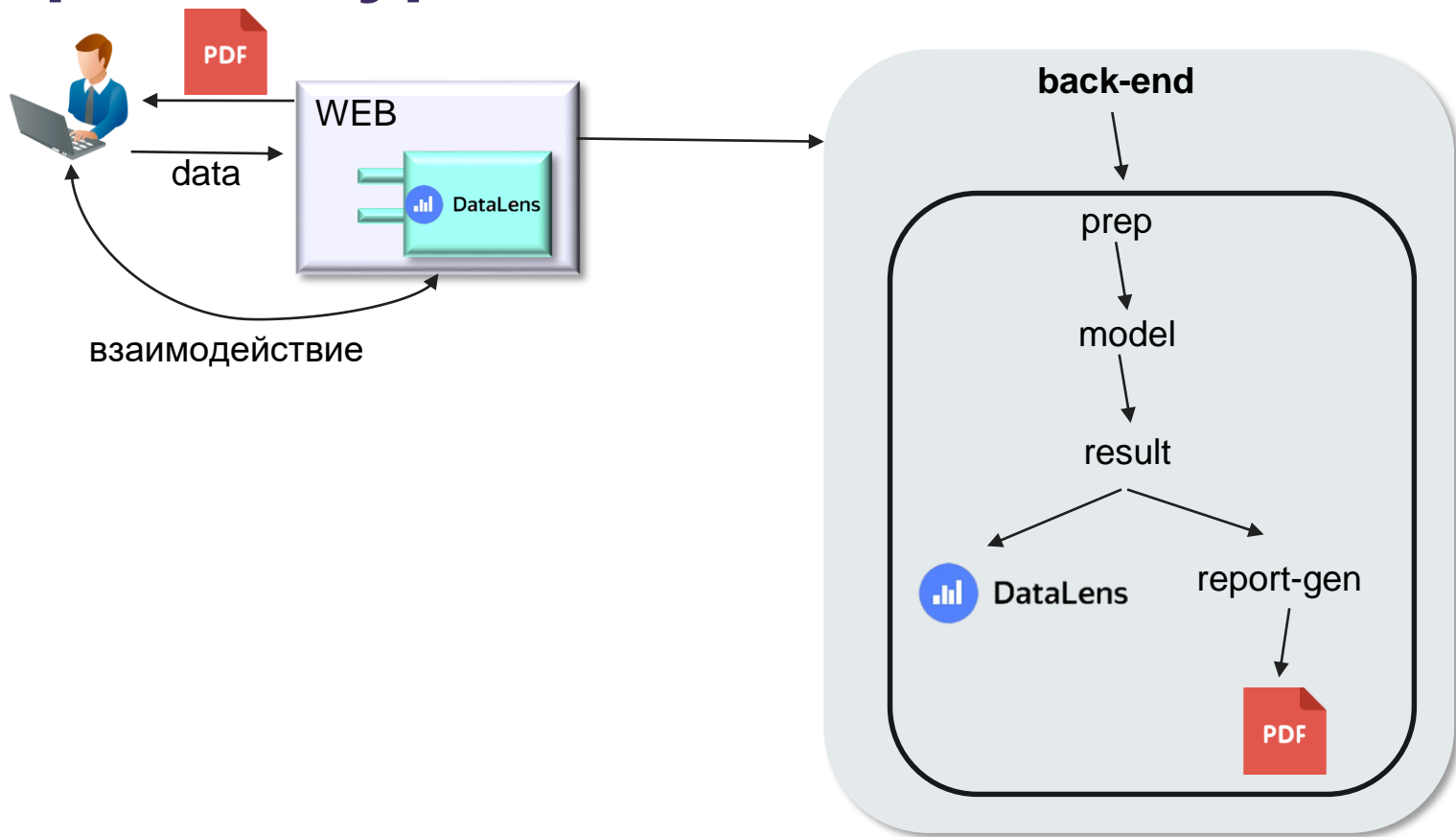
Концепция системы (1/2)



Концепция системы (2/2)

Этап	Инструмент	Описание	Результат
Сбор данных	Чат-бот в Telegram или форма на платформе компании	Сотрудник заполняет анкету, бот задаёт вопросы анкеты по заранее подготовленным темам	Ответы записываются в структурированную таблицу
Обработка и предобработка данных	Python (pandas, NLTK, Spacy)	Очистка текстов Разбиение на предложения или ключевые слова для более точного анализа	Чистые и подготовленные для анализа текстовые данные
Анализ данных	Машинное обучение и NLP	Кластеризация текстов Анализ тональности	Кластеры с классифицированными причинами увольнения и аналитика по эмоциональной окраске
Генерация отчёта	Python и Yandex DataLens	Создание сводного отчёта с визуализацией данных Включение текста, описывающего ключевые выводы	Интерактивный дашборд, содержащий анализ данных, визуализации и текстовые выводы
Рекомендации на основе данных	ML-модель	На основе собранных данных система генерирует персонализированные рекомендации для HR-отдела	HR-специалисты получают список рекомендаций по каждому кластеру причин увольнения
Сопровождение	Автоматизация через интеграцию с HR-системами (API)	Регулярное обновление данных и автоматическая генерация отчётов по мере поступления новых анкет	Система автоматически обновляется и отправляет свежие отчёты без участия пользователя

Архитектура системы



Алгоритмы и ИИ



Схема работы скриптов

+

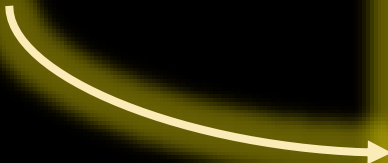


Модель

Оценка модели

Метрики модели для анализа
тональности

rubert-sentimental-analysis
(руберт для анализа
тональности)



	neutral	positive	negative	macro avg	weighted avg
precision	0.7	0.84	0.74	0.76	0.75
recall	0.74	0.83	0.69	0.75	0.75
f1-score	0.72	0.83	0.71	0.75	0.75
auc-roc	0.85	0.95	0.91	0.9	0.9
support	5196	3831	3599	12626	12626

Визуализации по тону

Позитивный

Наработки работой Руководитель проекты Лучшая **Уход** было Нервная Усиление Интересные Интересная Удаленность Спасибо Всё хорошо **пенсию** **Компания** ГК **Росатом** Процветание **Благодарность** Жалоб Бывший Быт Косяки Команда режима **работа** от коллега карьерный рост **Доволен** На Загруженность Документация

Негативный

Подчиненные ТопДедаины Безответственность **Пир** Пароли **ОКР** Обучение Поставщики Сотрудники Оскорбления **Неуважение** границы Проект Руководитель рост Улучшение **Отсутствие** Система Карьерный Личные Инициативы **увольнение** Мнение отдела Коллеги Конкуренция **Непрофессионализм** Руководство **Обязанности**

Динамический свод с фильтрами

Вопрос ?

Комментарий к вопросу 1. Какие причины (факторы) сформировали ва... ▾

Крастеры ответов сотрудников

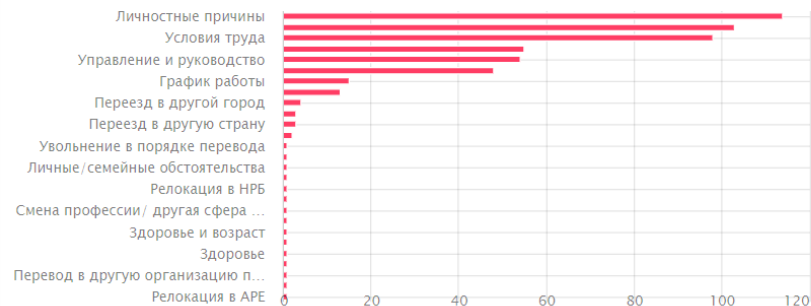
...



Аналитика по конкретным вопросам

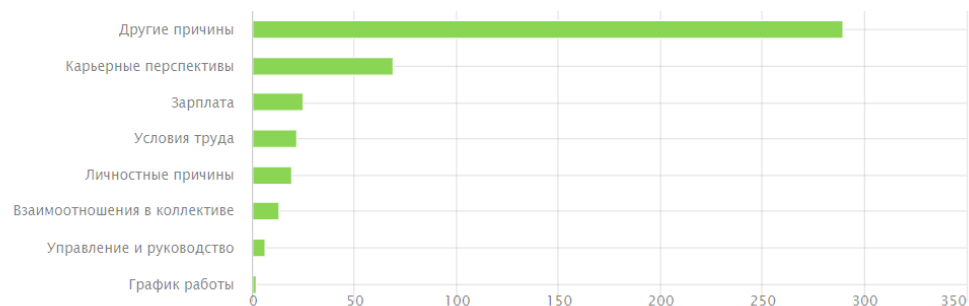
Основная причина увольнения

...



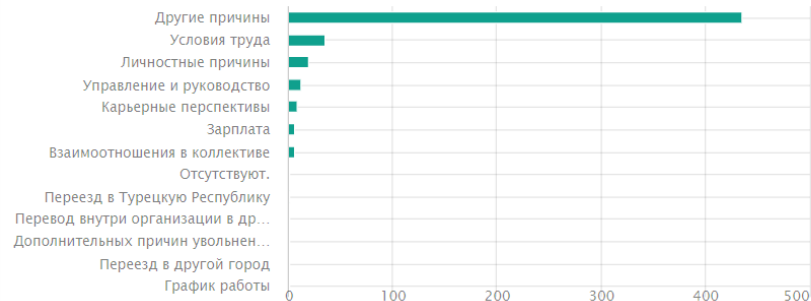
Почему сотрудники думают остаться или перевестись внутри отрасли

...



Дополнительная причина увольнения

...



Дорожная карта реализации

Этап 1: Анализ требований
Сбор данных по ключевым метрикам, определение
целей и функциональности

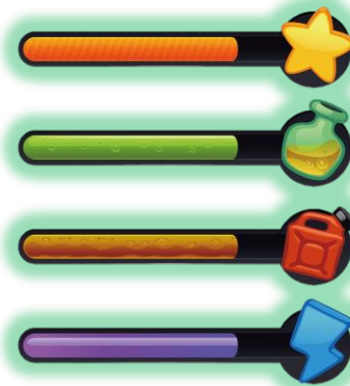
Готово на 100%

Этап 2: Разработка MVP
Разработка основных модулей, тестирование

Готово на 100%

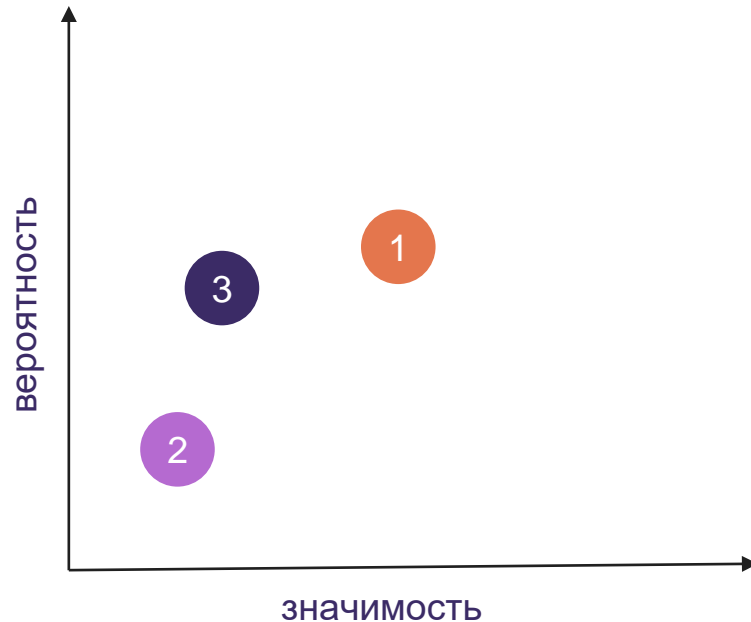
Этап 3: Сопровождение
Интеграция с формами сбора информации,
расширение функционала

+ 2 нед



Риски и пути их минимизации

- 1 Неправильная кластеризация или неверный анализ тональности -> обучение на реальных данных, обратная связь
- 2 Некорректные ответы или пропуски -> проверки на обязательность, предупреждение о некорректных данных
- 3 Технические трудности при интеграции с платформами -> Использование стандартных API для интеграции, тестирование на ранних этапах, A/B тесты



Стратегия развития системы



Интеграция с 1С /
платформами сбора
информации



Система **отслеживания**
отзывов сотрудников:
через регулярную анкету
удовлетворенности



Анализ корректности
введенного **ответа при**
заполнении анкеты

Наш процесс работы через игру

❏ Сериус

- Исходные данные
- Анализ конкурентов
- Кластеризация
- Архитектура
- Визуализация

Проекты / ithack

ithack

Поиск



Эпик

Метка

ГРУППЫ

К ВЫПОЛНЕНИЮ 10

Оценка качества кластеров
КЛАСТЕРИЗАЦИЯ ОТВЕТОВ
✓ НАСК-172

Предварительная обработка данных
ОБРАБОТКА ДАННЫХ АНКЕТ
✓ НАСК-173

Выделение ключевых тем
КЛАСТЕРИЗАЦИЯ ОТВЕТОВ
✓ НАСК-178

Подсчет количества ответов в каждом кластере
АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ДАННЫХ
✓ НАСК-179

Визуализация результатов
АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ДАННЫХ
✓ НАСК-181

В РАБОТЕ 1

Анализ тональности и содержания
АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ДАННЫХ
✓ НАСК-180

+ Создать задачу

РЕВЬЮ 2

Реализовать загрузку данных из файлов Excel
ОБРАБОТКА ДАННЫХ АНКЕТ
✓ НАСК-170

Кластеризация ответов
КЛАСТЕРИЗАЦИЯ ОТВЕТОВ
✓ НАСК-171

ГОТОВО 67 ✓

Сделать ИИ для помощи выбора задач
ИИ не_в_хакалоне
✓ НАСК-52

Сбор и анализ данных о релизах и развертываниях
ОТСЛЕЖИВАНИЕ ЭФФЕКТИВНОСТИ ЧЕ... не_в_хакалоне
✓ НАСК-113

Оптимизация процессов взаимодействия через встречи
МЕТРИКА ВОВЛЕЧЕННОСТИ: КОЛИЧЕС... не_в_хакалоне
✓ НАСК-112

Сбор и анализ данных о встречах сотрудников
МЕТРИКА ВОВЛЕЧЕННОСТИ: КОЛИЧЕС... не_в_хакалоне
✓ НАСК-111

Команда



Анна Николаева

Экономика

Лид команды и внешних коммуникаций
Финалист конкурса "Лидеры России" ЦФО и
поддержка команды



Артём Носов

Программная инженерия

Лид разработки

Работает в IT-компании, 2 года лидирует команды в
хакатонах и просто упорный человек



Александр Архангельский

Электроника и автоматика физических установок
Инженер-программист и живая Википедия



Яна Кузьмина

Прикладная математика и информатика

Эксперт по глубокому и машинному обучению, подружила
продукт с ИИ

Ссылки



Дашборд



Отчет PDF