



ЛИДЕРЫ ЦИФРОВОЙ
ТРАНСФОРМАЦИИ



АДМИНИСТРАЦИЯ
ГОРОДА НЕРЮНГРИ



КОЛМАР
УГЛЕДОБЫВАЮЩАЯ КОМПАНИЯ

ТРЕК: Сервис прогнозирования и увольнения на основе вовлеченности сотрудника

Администрация г. Нерюнгри и Колмар

Команда: **K**

Содержание

Что мы вам покажем

Представление команды	3	Структура сервиса	9
Постановка задачи	4	Описание сервиса	10
Используемые библиотеки	5	Телеграмм Бот	15
Описание данных	6	Основные преимущества	16
EDA	7	Перспектива развития	18
ML модель	8	Ссылка на сервис	19

Команда K



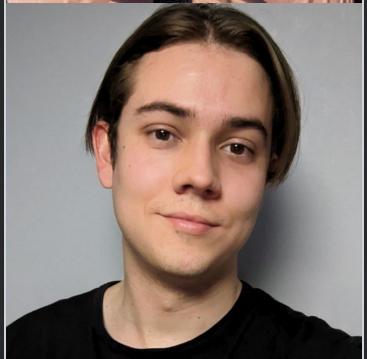
Алексей Верт-Миллер

Россия г. Архангельск

Data Science

Текущий стек: Python, Pandas, Sklearn, Catboost.

@alexwert3
+7(921)295-42-72



Виталий Ерохин

Россия г. Москва

Data Science

Текущий стек: Python, Pandas, Keras, Pytorch, Catboost.

@warleagle
+7(928)008-93-57



Руслан Латипов

Россия, г. Зеленодольск

Капитан команды

Full stack developer

Текущий стек: ML, Python, SQL

@rus_lat116
+7(937)521-32-77

Юрий Дон

Россия г. Краснодар

Data Science

Текущий стек: Python, Sklearn, Pandas, Pytorch, Android, Kotlin

@Yuriy_Nikitich
+7(918)988-11-11

Татьяна Моисеева

Россия г. Москва

Full stack developer

Текущий стек: Python, SQL, PHP, ML

@Estochka
+7(910)453-10-79

Задача

Потери компании при увольнении квалифицированного работника могут быть значительны, и, имея инструмент, оценивающий вероятность его скорого увольнения, работодатель может среагировать на это, организовав личные беседы и оценку результатов работы сотрудника с последующей разработкой мер, снижающих риск его увольнения.

Решение задачи представляет собой сервис, рассчитывающий вероятность увольнения сотрудника по показателям использования корпоративной почты.



КОЛМАР

УГЛЕДОБЫВАЮЩАЯ КОМПАНИЯ





Используемые библиотеки

Базовый язык программирования - Python

База данных - PostgreSQL

Backend - Aiohttp

Работа с данными - Pandas, Numpy

ML model - Sklearn

Frontend - Streamlit, AG Grid

Тестирование API- Postman



POSTMAN

Yandex



DataLens

Мы используем агрегированные данные за определенный интервал времени, что позволило получать высокие метрики на легких и хорошо интерпретируемых моделях.

Данные получаем из нескольких баз данных:

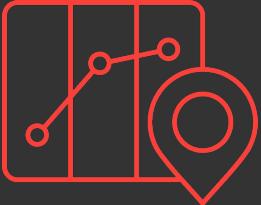
- База данных хранения статистики использования почтового сервиса (**кол-во входящих и исходящих сообщений, время отправки, объем и т.п.**).
- База данных хранения писем из почтовой переписки сотрудников (**письма**).
- База данных хранения HR статистики о персонале (**стаж, возраст, пол, статус уволен или не уволен и т.п.**).

Более подробная информации об агрегировании, конкатенации и обработке данных см. в документации к сервису.



Описание данных

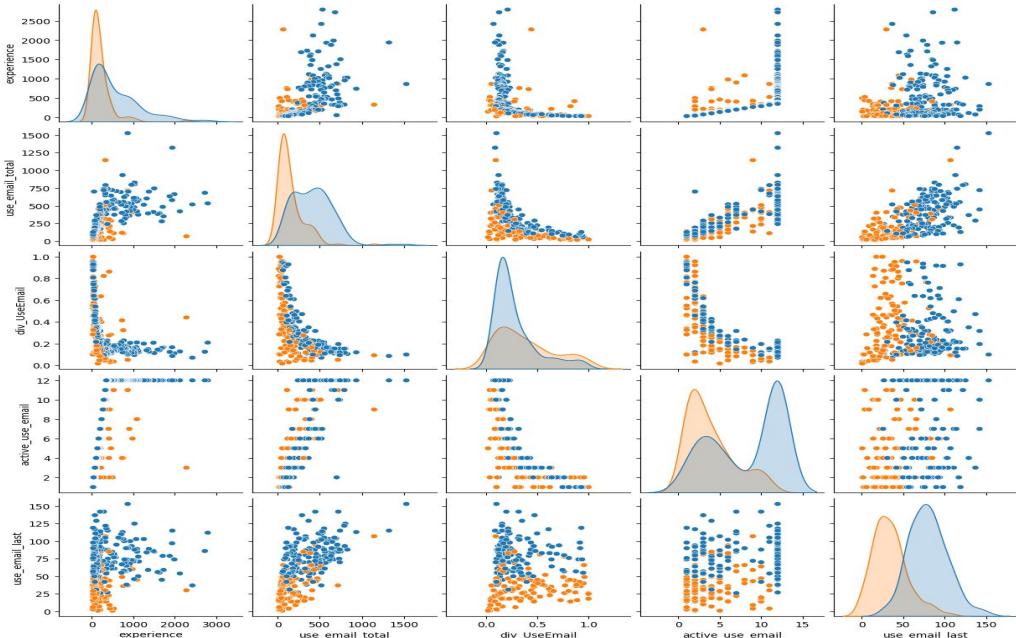
EDA



Phik корреляция признаков



Взаимосвязь признаков использования почты с целевым признаком

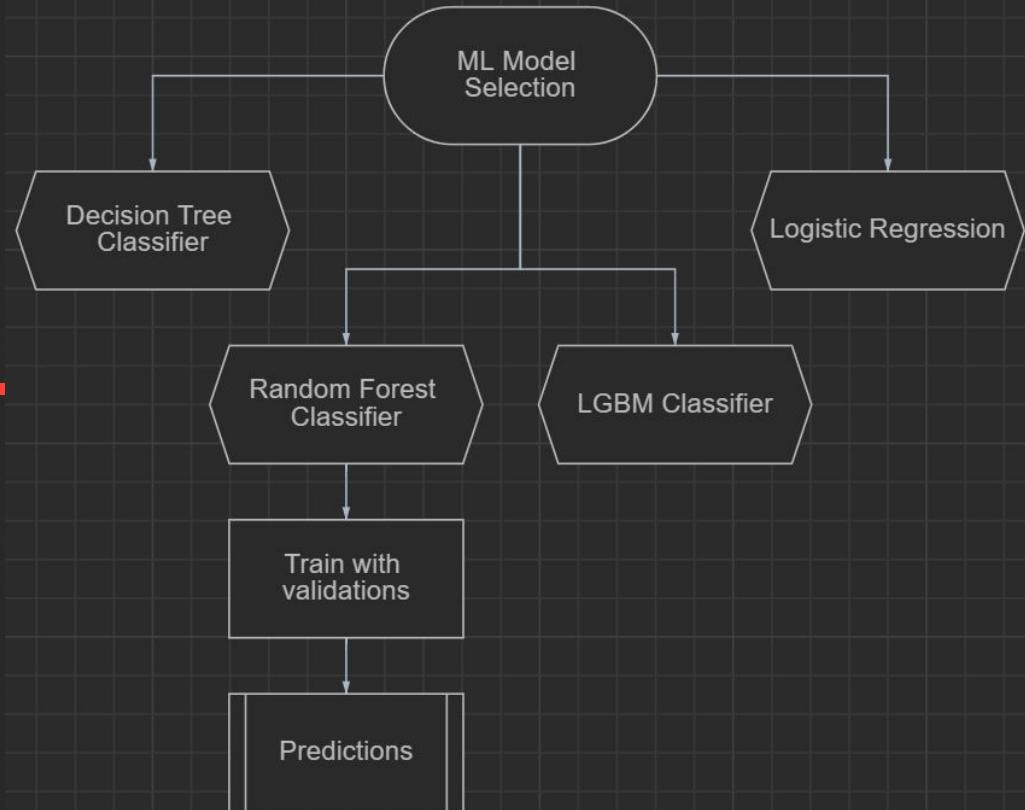


Обучение и выбор лучшей модели

Выбор лучшей модели и подбор гиперпараметров происходит с использованием GridSearchCV

В пайпайне обучения происходит нормализация данных с помощью StandardScaler

Обучение производилось с помощью кроссвалидации



Структура сервиса

Обработка данных соответствует поставленной задаче, структура отражена на схеме. На вход в сервис поступает предобработанный файл CSV и на основании него сервис выводит вероятности увольнения сотрудников, аналитику и делает рассылку.
Пошаговая логика на следующем слайде

[HR и email базы >>](#)

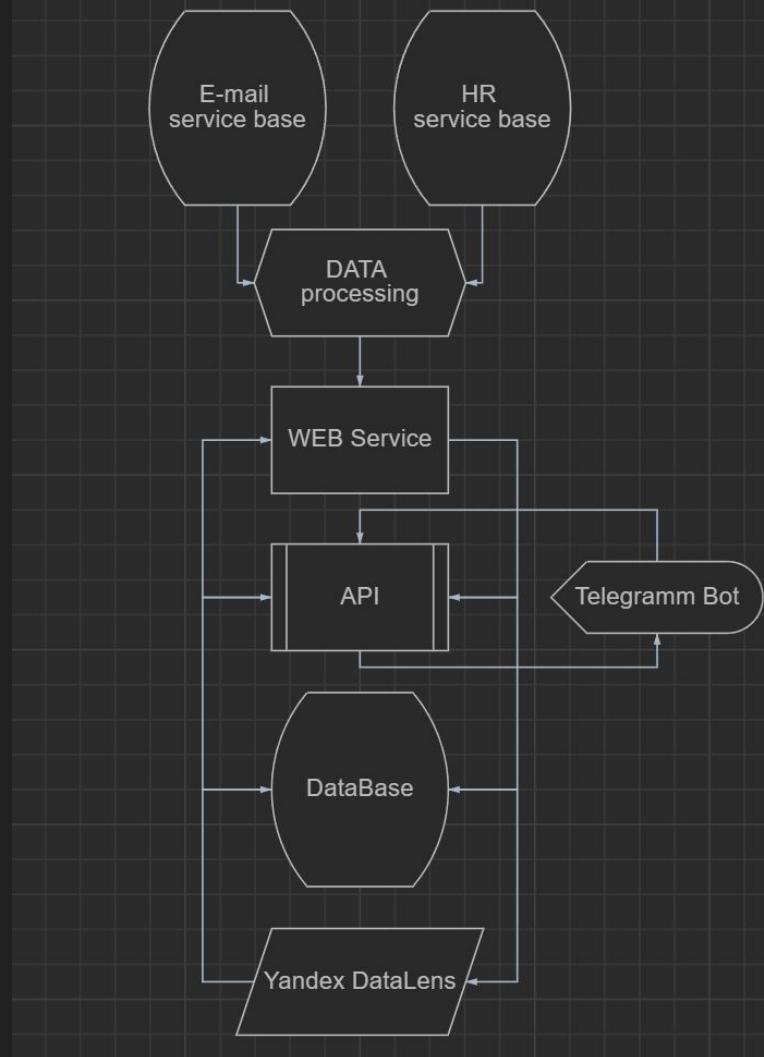
[Обработка данных >>](#)

[Front >>](#)

[Backend >>](#)

[Storage >>](#)

[Dashboard >>](#)





Сервис предсказания увольнения

01

первый шаг

02

второй шаг

03

третий шаг

04

четвертый шаг

05

пятый шаг

Загрузка данных в
сервис с
использованием
сводного файла в
формате csv.

Формирования
дополнительных
признаков путем
вычисления
соотношений
коррелирующих
признаков

Предсказание
вероятности
увольнение
сотрудников,
предварительно
обученной ML
модели (Random
Forest Classifier)

Визуализация
предсказаний в
виде
аналитических
дашбордов в
разрезе отделов и
департаментов

Рассылка таблиц,
файлов
ответственным
исполнителям по
работе с
текущестью на
почту и Телеграмм
Бот

КОЛМАР
УГЛЕДОБЫВАЮЩАЯ КОМПАНИЯ

**Сервис прогнозирования увольнения
на основе вовлеченности сотрудника**

О программе

- прогнозная оценка в % вероятности увольнения каждого сотрудника
- аналитика в разрезе отделов и департаментов
- постановка на контроль сотрудников с высокой вероятностью увольнения
- рассылка файлов с результатами прогнозации
- уведомления, контроль через телеграм-бот
- отображение ТОП-3 сотрудников с самой высокой вероятностью увольнения
- осуществление контроля за сотрудниками с высокой вероятностью увольнения
- сравнительная статистика
- чек лист бесед с сотрудниками

01

О программе

На данной странице выводится краткое описание сервиса

КОЛМАР
УГЛЕДОБЫВАЮЩАЯ КОМПАНИЯ

Персонал

Департаменты

Отделы

Персонал

Должности

Роль

Сотрудники

Шаг

Добавить/удалить роль

Название роли
Введите наименование роли...

Добавить Удалить

сохранить

ID роли	Роль
<input type="checkbox"/>	1 Руководитель
<input type="checkbox"/>	2 Начальник
<input type="checkbox"/>	3 Сотрудник

1 из 3 по 3 < > Страница 1 из 1

02

Персонал

На странице мы можем увидеть, добавить, удалить подразделение, отдел, сотрудников. Так же можем сортировать и фильтровать сотрудников.

Прогнозирование увольнения на основе вовлеченности сотрудника

Данные с почтовой перепиской

Текущий день
2023-12-16

December 2023

Su Mo Tu We Th Fr Sa

1 2
3 4 5 6 7 8 9
10 11 12 13 14 15 16
17 18 19 20 21 22 23
24 25 26 27 28 29 30
31

Browse files

Прогнозирование увольнения на основе вовлеченности сотрудника

Демонстрационный режим

Загрузка файла

Drag and drop file here
Limit 200MB per file + CSV

test_hr_pred_new.csv (26.1KB)

Browse files

Файл успешно загружен

name	employee_id	male	age	experience	use_email_total	active_use_email	use_email_last	out_work_intern
Чернова Оксана Петровна	612	0	53	1728	388	12	63	
Борисова Ирина Ильинич	537	0	44	1709	517	12	65	
Баранов Валерий Борисович	137	0	32	1687	301	12	42	
Захарова Екатерина Александровна	596	0	44	1629	432	12	63	
Наумова Надежда Игоревна	120	0	45	1588	415	12	45	
Иванова Ульяна Артемовна	422	0	42	1505	770	12	56	

03

Прогноз увольнения

Прогноз осуществляется на данных загружаемого CSV файла. Имеется возможность выбора даты в демонстрационном режиме, для наблюдения вероятности увольнения в заданный период.

04

Вывод предсказания

Предсказанные вероятности выводятся в виде таблицы с предсказаниями уволится или нет и вероятности увольнения либо нет.

Аналитика

Анализ прогнозирования увольнения на основе вовлеченности сотрудника					
НР - все сотрудники		Департамент №2		Отдел №1	
ФИО ответственного	Должность	Отдел	Ссылка на телеграмм		
Мария Попова	Начальник отдела	Отдел №1	https://t.me/+3uNhu7jW0IwNz		
Кирилл Прокопов	Начальник отдела	Отдел №2			
Полина Годзев	Руководитель департамента	Отдел №2			
Лука Шаров	Начальник отдела	Отдел №3			

Основные фильтры		Количество сотрудников		Шанс увольнения более 50	
Название департамента	Название отдела	Возраст более	Сотрудников отслеживается	Шанс увольнения более заданного	Пол
Нет выбранного значения	Нет выбранного значения	Пол	7	12	

Текущая вероятность увольнения							
id	Имя	Фамилия	Возраст	Пол	Опыт работы		
896	Вероника	Михайлова	52	Ж	36	Отдел №3	90%
102	Даниил	Артемьев	37	М	208	Отдел №4	81%
184	Борис	Глебов	29	М	307	Отдел №3	77%
871	Светлана	Коновалова	21	Ж	106	Отдел №3	72%
551	Федор	Попов	44	М	126	Отдел №3	72%
809	Алиса	Петрова	27	Ж	139	Отдел №4	68%
238	Денис	Щербаков	49	М	174	Отдел №3	59%
407	Владимир	Смирнов	40	М	375	Отдел №4	59%
297	Виктория	Жукова	47	Ж	544	Отдел №3	54%
678	Иван	Колзов	38	М	167	Отдел №2	54%
705	Альбина	Кирилева	49	Ж	67	Отдел №2	54%
441	Артемий	Петров	40	М	204	Отдел №2	54%

Не отслеживаются с высокой вероятностью увольнения					
id	Имя	Фамилия	Вероятность увольнения	Разница с прошлым периодом	Пол
102	Даниил	Артемьев	62%	+16%	М
426	Ольга	Михайлова	5%	-5%	Ж
551	Федор	Поганов	73%	+13%	М
612	Оксана	Чернова	9%	-1%	Ж
865	Арина	Кирилева	14%	0%	Ж

05

Подробная визуализация результатов предсказаний в сравнительной динамики прошлых предсказаний вероятности увольнения сотрудников, в табличном виде. Цветовое выделение категорий вероятности увольнения сотрудников

Аналитика

Аналитика

ФИО сотрудников для отображения (фильтры только по отдельно)																																															
Алия Котова, Денис Щербаков, Лилия Коновалова, Ольга Артемьева, София Наумова																																															
Динамика шансов увольнения																																															
19 Ноябрь 2023 Вс																																															
<table border="1"> <thead> <tr> <th>Дата</th> <th>Денис Щербаков</th> <th>Ольга Артемьева</th> <th>София Наумова</th> <th>Лилия Коновалова</th> <th>Алия Котова</th> </tr> </thead> <tbody> <tr> <td>06.11.23</td> <td>55%</td> <td>55%</td> <td>55%</td> <td>55%</td> <td>55%</td> </tr> <tr> <td>13.11.23</td> <td>56%</td> <td>56%</td> <td>56%</td> <td>56%</td> <td>56%</td> </tr> <tr> <td>20.11.23</td> <td>57%</td> <td>57%</td> <td>57%</td> <td>57%</td> <td>57%</td> </tr> <tr> <td>27.11.23</td> <td>58%</td> <td>58%</td> <td>58%</td> <td>58%</td> <td>58%</td> </tr> <tr> <td>04.12.23</td> <td>59%</td> <td>59%</td> <td>59%</td> <td>59%</td> <td>59%</td> </tr> <tr> <td>11.12.23</td> <td>60%</td> <td>60%</td> <td>60%</td> <td>60%</td> <td>60%</td> </tr> </tbody> </table>						Дата	Денис Щербаков	Ольга Артемьева	София Наумова	Лилия Коновалова	Алия Котова	06.11.23	55%	55%	55%	55%	55%	13.11.23	56%	56%	56%	56%	56%	20.11.23	57%	57%	57%	57%	57%	27.11.23	58%	58%	58%	58%	58%	04.12.23	59%	59%	59%	59%	59%	11.12.23	60%	60%	60%	60%	60%
Дата	Денис Щербаков	Ольга Артемьева	София Наумова	Лилия Коновалова	Алия Котова																																										
06.11.23	55%	55%	55%	55%	55%																																										
13.11.23	56%	56%	56%	56%	56%																																										
20.11.23	57%	57%	57%	57%	57%																																										
27.11.23	58%	58%	58%	58%	58%																																										
04.12.23	59%	59%	59%	59%	59%																																										
11.12.23	60%	60%	60%	60%	60%																																										

06

Графическая визуализация динамического изменения вероятности сотрудника или нескольких сотрудников на выбор пользователя.

Аналитика

Аналитика

Анализ прогнозирования увольнения на основе вовлеченности сотрудника

НР: все сотрудники | Департамент №2 | Отдел №1 | Письма

Отслеживаемые сотрудники: Нет выбранных элементов | Конкретный сотрудник: Нет выбранных элементов

История входа в почту по департаментам

Департаменты	05.11.2023	12.11.2023	19.11.2023	26.11.2023	03.12.2023	10.12.2023
Вход в почту (среднее)	83	43	36	45	90	48
Вход в почту за период (месяц)	83	43	36	45	90	48
Вход в почту за период (среднее)	83	43	36	45	90	48
Вход в почту за период (неделя)	83	43	36	45	90	48
Вход в почту за период (год)	83	43	36	45	90	48
Вход в почту за период (месяц)	83	43	36	45	90	48
Вход в почту за период (среднее)	83	43	36	45	90	48
Вход в почту за период (неделя)	83	43	36	45	90	48
Вход в почту за период (год)	83	43	36	45	90	48
Департамент №1	75	7	78	8	81	8
Департамент №2	75	7	78	8	81	8

Среднее количество дней между получением письма

Среднее количество сообщений, на которые направлен ответ

03 декабря 2023 Кс
Сообщение от отдела №1
Отдел №1-Департамент №1 5
Отдел №2-Департамент №2 4
Отдел №3-Департамент №3 4
Сумма 18

03 декабря 2023 Кс
Сообщение от отдела №1
Отдел №1-Департамент №1 5
Отдел №2-Департамент №2 4
Отдел №3-Департамент №3 4
Сумма 18

Контроль

Чек лист для работы с сотрудниками под риском увольнения

Выявлять причины недовольственности работой
ответ сотрудника

Выявлять потребности, которые не достаточно полно удовлетворяются у сотрудника
ответ сотрудника

При каких обстоятельствах сотрудник изменит свое отношение к компании
ответ сотрудника

Нужна ли сотруднику помощь в решении проблем, которые помогут ему остаться в компании
ответ сотрудника

07

Аналитика

Визуализация входящих признаков в виде таблиц и графиков

08

Контроль

В сервисе реализован функционал для оперативной проработки HR с сотрудниками с высокой вероятностью увольнения. Чек лист предназначен для выявления причин, нежелания работать и потребностей сотрудника м самоконтроля HR.

ТГ Бот

В телеграмм боте 3 типа доступа:

- HR
- Департамент
- Отдел

Бот умеет:

- Отправлять списки с сотрудниками и их вероятностями увольнения
- Отправлять файлы с той же информацией
- Выбирать топ 3 самых высоковероятностных к увольнению сотрудников
- Отслеживать выбранных сотрудников
- Выводить информацию по сотруднику

Якутия: HR
3 members

Yakutiya_bot /топ3 15:55 ✓

Yakutiya_bot 3 сотрудника с наибольшей вероятностью уволиться:
Сотрудник 896: Вероника Михайлова, вероятность: 90.91%
Сотрудник 102: Даниил Артемьев, вероятность: 77.27%
Сотрудник 871: Светлана Комарова, вероятность: 72.73% 15:55

Yakutiya_bot /команды 15:55 ✓

Yakutiya_bot топ3, отслеживать, неотслеживать, список, отчет, инфо 15:55

Yakutiya_bot /топ3 15:55 ✓

Yakutiya_bot 3 сотрудника с наибольшей вероятностью уволиться:
Сотрудник 896: Вероника Михайлова, вероятность: 90.91%
Сотрудник 102: Даниил Артемьев, вероятность: 77.27%
Сотрудник 871: Светлана Комарова, вероятность: 72.73% 15:56

Yakutiya_bot id: 2 ФИО Никита Наумов - Департамент №2
опыт работы: 33 15:56

Yakutiya_bot /инфо 5 15:56 ✓

Yakutiya_bot id: 5 ФИО Лука Шаров - Департамент №2
опыт работы: 33 15:56

Esta /список Отдел №2 15:58

Yakutiya_bot Сотрудники отдела: Отдел №2
Сотрудник 4: Кирилл Прохоров, Начальник отдела
Сотрудник 422: Ульяна Иванова, Сотрудник отдела
Сотрудник 290: Евгения Соколова, Сотрудник отдела
Сотрудник 515: Леонид Шилов, Сотрудник отдела
Сотрудник 815: Лев Петров, Сотрудник отдела

Якутия: Отдел №1
3 members

Yakutiya_bot /топ3 16:01 ✓

Yakutiya_bot 3 сотрудника с наибольшей вероятностью уволиться отдела:
Отдел №1
Сотрудник 778: Александра Князева, вероятность: 45.45%
Сотрудник 597: Полина Жукова, вероятность: 40.91%
Сотрудник 174: Александра Семенова, вероятность: 40.91% 16:01

Yakutiya_bot /отчет Отдел №1 16:01 ✓

Yakutiya_bot 20231216-Отдел 1.csv 5.9 KB 16:01

Yakutiya_bot /инфо 16:02 ✓

Yakutiya_bot В текущей версии доступно только HR 16:02

Yakutiya_bot /отчет 16:02 ✓

Yakutiya_bot 20231216-Отдел 1.csv 5.9 KB 16:02

Yakutiya_bot /отчет Отдел №2 16:02 ✓

Yakutiya_bot Нет доступа к данным 16:02

Yakutiya_bot /отчет Отдел №1 16:04 ✓

Yakutiya_bot 20231216-Отдел 1.csv 5.9 KB 16:04

Write a message...

Основные преимущества

01

#Первое

02

#Второе

03

#Третье

Высокая точность предсказания

За счет правильной агрегации и обработки данных, и качественно подобранный ML модели и ее гипер параметров, сервис предлагает высокую скорость и точность предсказания вероятности увольнения сотрудника.

Подробная аналитика

Визуализация, обобщения информации по вероятности увольнения сотрудников в разрезе отделов и департаментов в виде таблиц и графиков

Оповещения ответственных сотрудников (телеграмм, почта)

Отправка отчетов на почту ответственным сотрудникам
Предоставление всей статистики по предсказаниям в ТГ Бот по запросу

04
#Четвертое

05
#Пятое

06
#Шестое

07
#Седьмое

Реализация desktop и web решения

для удобства использования и дополнительных возможностях масштабирования реализованы оба варианта и обеспечения работы алгоритма при наличии интернета и без интернета (при условии развертывания базы данных и API на локальном хосте)

User Friendly интерфейс

удобный и логически понятный интерфейс сервиса, за счет простоты реализации с привычным для пользователей стандартных веб приложений

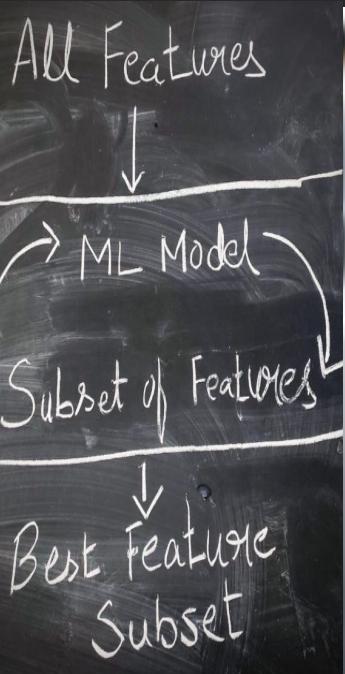
Ролевая модель пользователей

в сервисе реализована адресная рассылка результатов предсказания, а также возможно разграничение доступа на соответствующие чарты дашборда

Асинхронное API и реляционная БД

высокая скорость запросов и обработки информации, надежное хранение данных

Перспектива развития



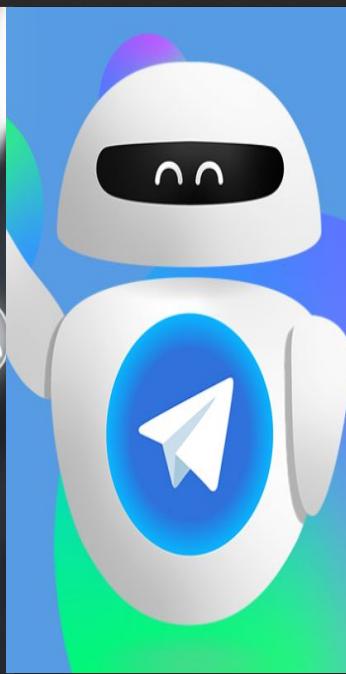
Признаки

Поиск новых дополнительных взаимосвязей, в отчетности для усовершенствования результатов предсказания



Контроль

Реализовать автоматическую систему контроля за работой HR специалиста с сотрудниками с высоким уровнем риска увольнения.



ТГ Бот

Реализовать автоматическое добавления данных через ТГ, поиск в БД по фамилии, изменение информации в базе., добавить еженедельный отчет.



NLP

Использовать нейросети для работы с текстом, в целях выявления по тексту настроение сотрудников и их вовлеченность.



Ссылка на сервис

[Web сервис](<http://46.243.227.129:5981/>)



Спасибо



за внимание