

Команда 14

# GlowByte Хакатон

2025

Прогноз самовозгорания угля при  
открытом хранении



# Состав команды

## Патрушева Виолетта

- ML специалист
- @hakloi

## Куликова Алёна

- Тимлид,  
фронтенд,  
бэкенд, ML
- @kulikova\_alyona

## Берсенев Андрей

- Аналитик  
данных, DS
- @BerAndrey

## Эшмуродов Азизбек

- FullStack  
разработчик
- @it\_is\_worth\_it

# Описание решения

На сайте пользователь может:

## Просматривать графики по четырем категориям:

- Выгрузка/отгрузка угля
- Самовозгорания (пожары)
- Температура в штабелях
- Погодные условия

## Создавать новые графики через диалоговые окна

- выбрать файл, колонку с датой, параметры (ось Y), период и тип графика (линейный, гистограмма и др.)

## Загружать новые погодные данные

в формате CSV под названием `weather_data_YYYY.csv`

## Формировать прогноз риска самовозгорания угля

- загрузить файл с данными штабелей под названием `schedule_for_prediction.csv`,
- запустить ML-модель,
- получить результаты с вероятностями и бинарными предсказаниями,
- скачать отчёт в формате TXT и данные в CSV.

## Переобучать модель на актуальных данных

(кнопка «Пересоздать модель»).

## Автоматически сохранять все настройки графиков

между сессиями (в файле `schedule.json`).

## Возвращаться на главную страницу

в любой момент.

# Главная страница

Deploy

## Прогноз самовозгорания угля

24.11.2025 14:32

СЭУ-файл № 11 • Создано графиков: 4

Добавить погоду

Главная

Добавить погоду

### Прогноз

Добавить предикторы Запустить прогноз Пересекать модели

### Метеоданные

Режим отображения Выбрать годы Все годы Годы 2021 Даты Выбрать 1 год День для отображения 0 365

Параметры для отображения Температура Влажность Насыщ.

Метеоданные (3 лет) – среднее по дням

Даты 01.11.2020 01.01.2021 01.04.2021 01.07.2021 01.10.2021 01.01.2022 01.04.2022 01.07.2022 01.10.2022

Графики (1/4) температура (2021) влажность (2021) насыщ. (2021) предиктор (2022)

### Быстрое создание графиков

Выгрузка/Отгрузка Самовозгорания Температура Погода

Выгрузка/Отгрузка Самовозгорание Температура Погода

### Выгрузка и отгрузка на склад

supplies\_На склад, гг\_На судно, тн

\* Часть страницы

# Главная страница (секция “Метеоданные”)

## Метеоданные

Режим отображения

Выбрать годы  Все годы

Годы

2021 x 2016 x 2019 x

Выбрано: 3 лет

Дней для отображения

26

1

365

Параметры для отображения

weather\_code x

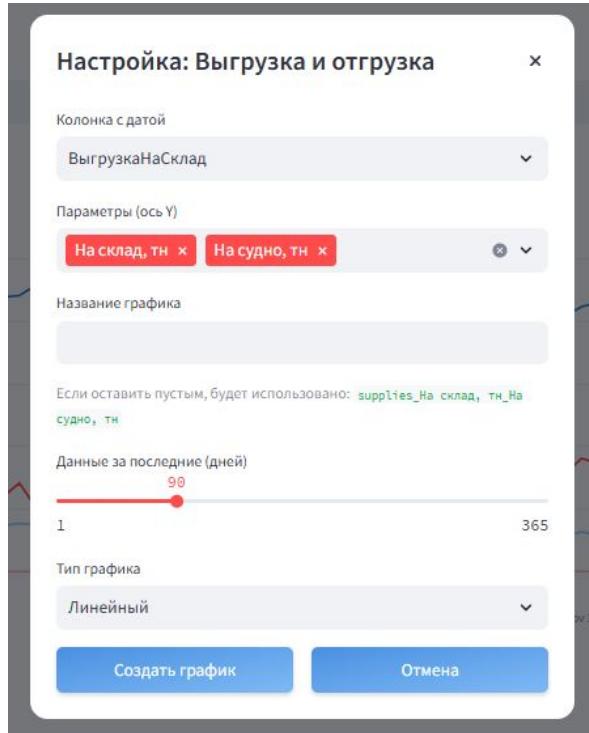
Метеоданные (3 лет) — среднее по дням



\* Метеоданные на 2021, 2016, 2019

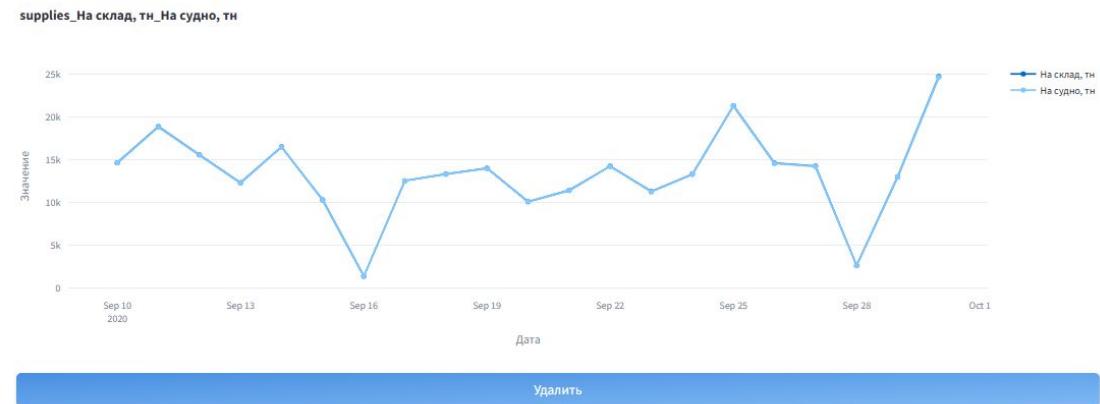
# Главная страница

(секция “Выгрузка и отгрузка”)



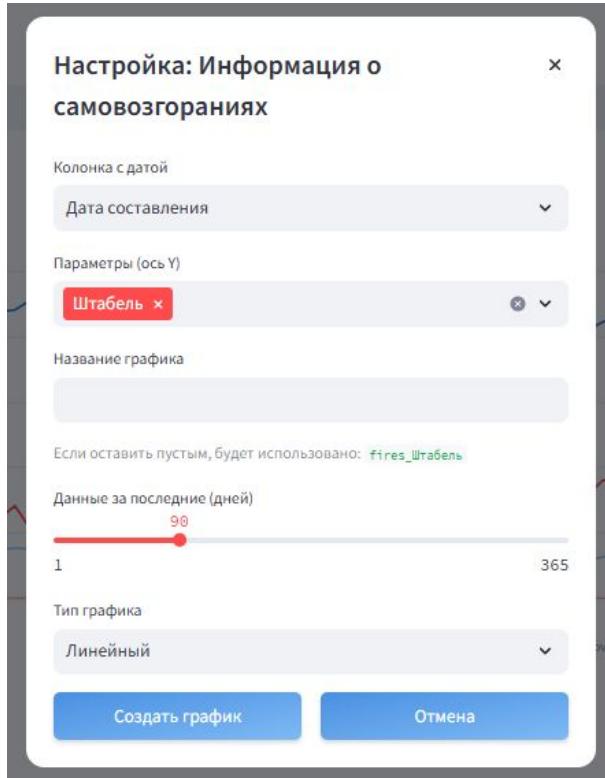
Окно для создания графика “Выгрузка и отгрузка”

## Выгрузка и отгрузка на склад

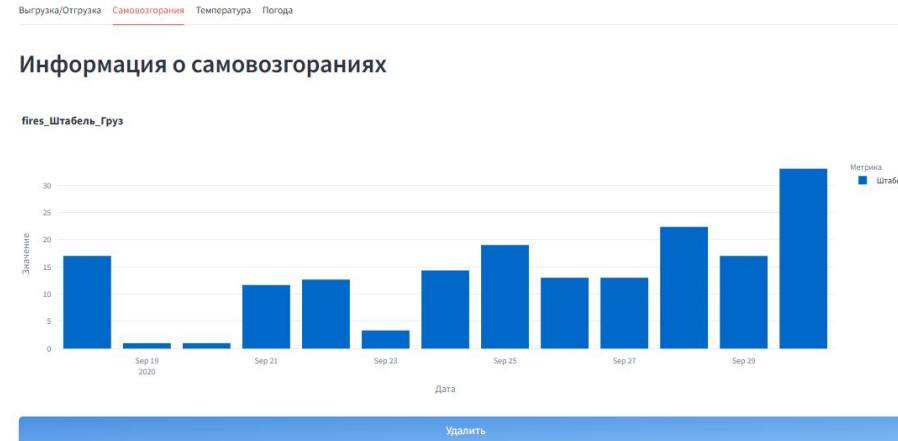


Созданный график во вкладке “Выгрузка и отгрузка”

# Главная страница (секция “Самовозгорание”)



Окно для создания графика “Самовозгорание”



Созданный график во вкладке “Самовозгорание”

# Главная страница (секция “Температура”)

Настройка: Показатели температуры в штабелях

Колонка с датой  
Дата акта |

Параметры (ось Y)  
Максимальная ...

Название графика

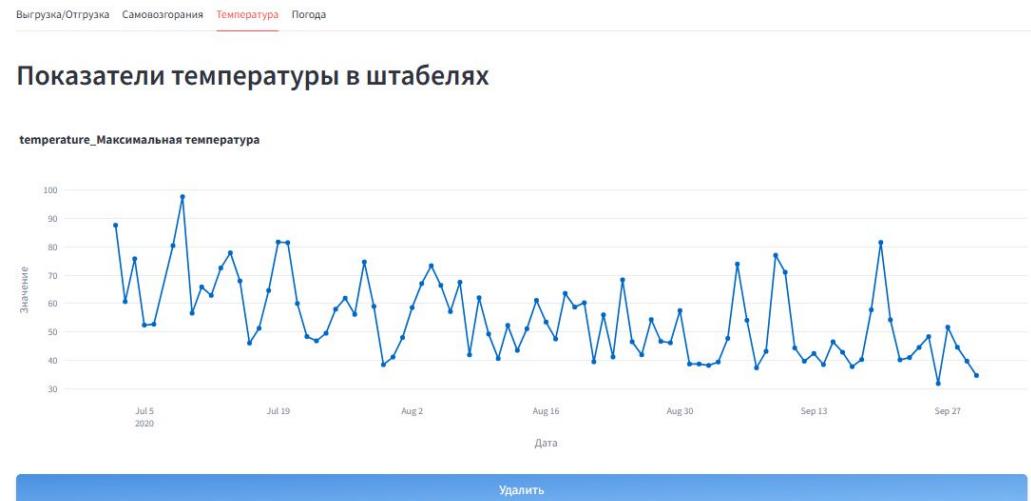
Если оставить пустым, будет использовано: temperature\_Максимальная температура

Данные за последние (дней)  
90

Тип графика  
Линейный

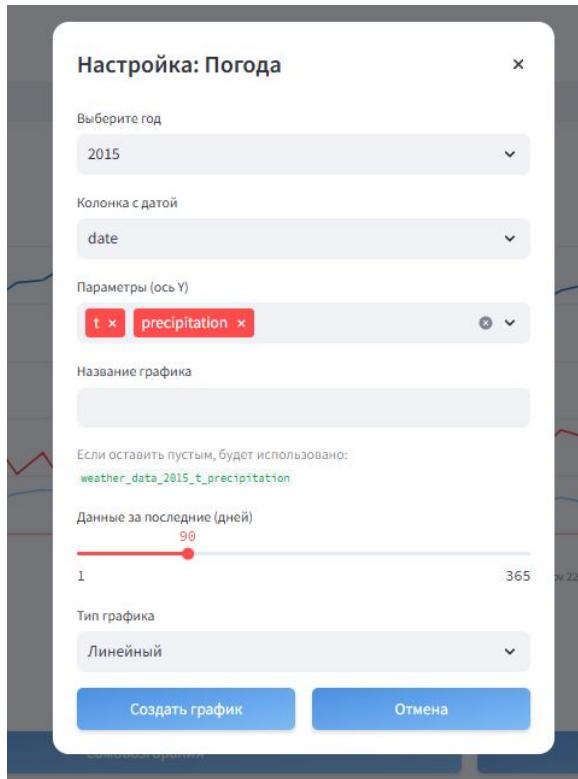
Создать график      Отмена

Окно для создания графика “Температура”

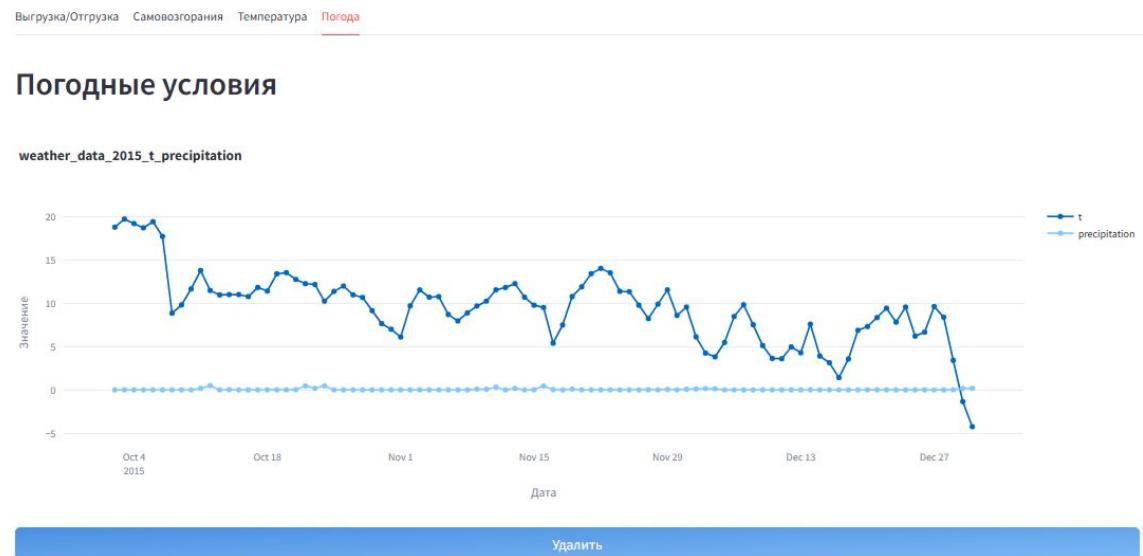


Созданный график во вкладке “Температура”

# Главная страница (секция “Погода”)



Окно для создания графика “Погода”



Созданный график во вкладке “Погода”

# Главная страница (Получение прогноза)

## Прогноз самовозгорания угля

**24.11.2025 14:50**

CSV-файлов: 11 • Создано графиков: 4

Главная

Добавить погоду

Добавить погоду

Прогноз

Добавить predict файл

Запустить прогноз

Пересоздать модель

### Результаты прогноза риска самовозгорания

Модель загружена: <class 'sklearn.ensemble.\_forest.RandomForestClassifier'>

LabelEncoder загружен.

Прогноз завершен: 3 из 5 записей — высокий риск

Макс. вероятность

**0.478**

Ср. вероятность

**0.195**

Порог классификации

**0.10**

[Скачать результаты \(CSV\)](#)

[Скачать аналитический отчет \(TXT\)](#)

Первые 10 записей:

	Марка	Возраст_дн	mass	Максимальная температура	Temp_изменение	weekday	month	t	p	humidity	fire_proba	fire_pred
0	A1	30	1,000	25	1	1	6	20	1,013	70	0.016	0
1	E5	150	3,000	60	5	3	7	30	1,008	45	0.3504	1
2	A1	250	6,000	85	12	5	8	35	1,005	30	0.4779	1
3	A1	45	1,500	35	-2	2	9	25	1,015	60	0.0035	0
4	E5	180	4,500	75	8	4	10	15	1,020	80	0.1291	1

# Техническая проработка решения

## Backend

Python

## Frontend

Python/streamlit

## ML

Python, pandas, numpy, scikit-learn,  
matplotlib

## Модель\*

RandomForestClassifier

\* Задача подразумевает, что таргет  $y_{3d}$  имеет два варианта *True/False*, что подталкивает нас на *binary classification* решение

# Аналитика для прогнозирования самовозгорания угля

## Цель анализа

- Оценить объемы и динамику грузов по складам и штабелям
- Изучить температурный режим штабелей и риски перегрева
- Проверить качество данных перед дальнейшим моделированием

## Используемые источники

- fires.csv – акты по штабелям (формирование / события, ~500 строк)
- supplies.csv – логистика склад-судно (несколько тысяч записей)
- temperature.csv – измерения максимальной температуры (~4 тыс. измерений)

## Анализ EDA в Colab



<https://colab.research.google.com/drive/1f0MnJOH25du1Ez8OdfwCaZl2LLVvoxQf#scrollTo=dfda01d8>

# Аналитика для прогнозирования самовозгорания угля

- Структура и объемы

- supplies.csv – основной и самый полный слой; склады 3 и 4 дают основную массу грузов
- fires.csv – небольшой реестр «особых» штабелей с несколькими крайне крупными актами
- temperature.csv – фон 30–40°C, есть единичные экстремальные значения до сотен °C

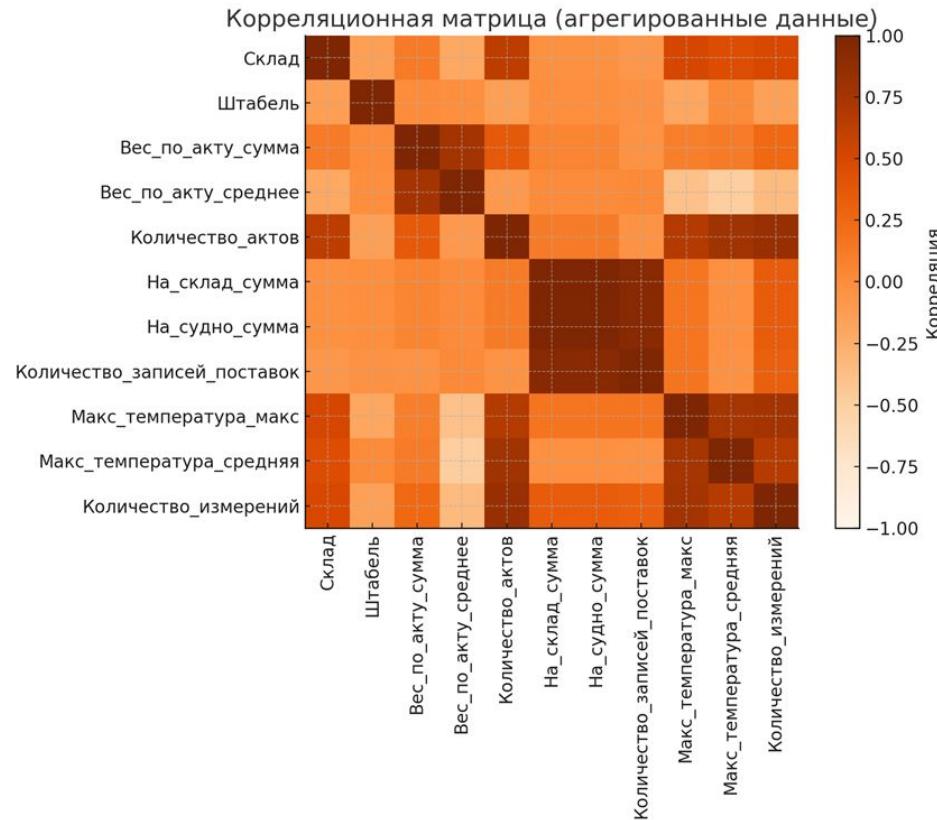
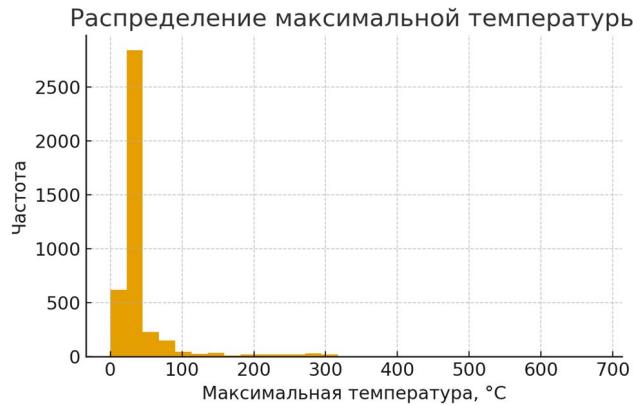
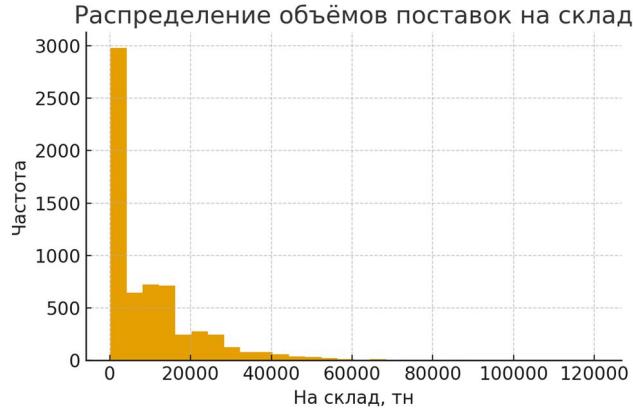
- Качество данных

- Хорошая заполняемость, особенно по логистике
- Есть временные несогласованности (отрицательные интервалы между началом формирования и поставками/измерениями)
- Выбросы по температуре, требующие проверки или фильтрации

- Связи и риски

- Явной зависимости «чем больше штабель, тем выше температура» не выявлено
- Один из складов выглядит «горячее» остальных по распределению температур

# Аналитика для прогнозирования самовозгорания угля



# ML решение подробнее

## Цель

прогнозирование вероятности пожара на 3 дня вперед ( $y_{3d}$  1/0) для каждого штабеля

## Решение

Были взяты несколько основных **features**:

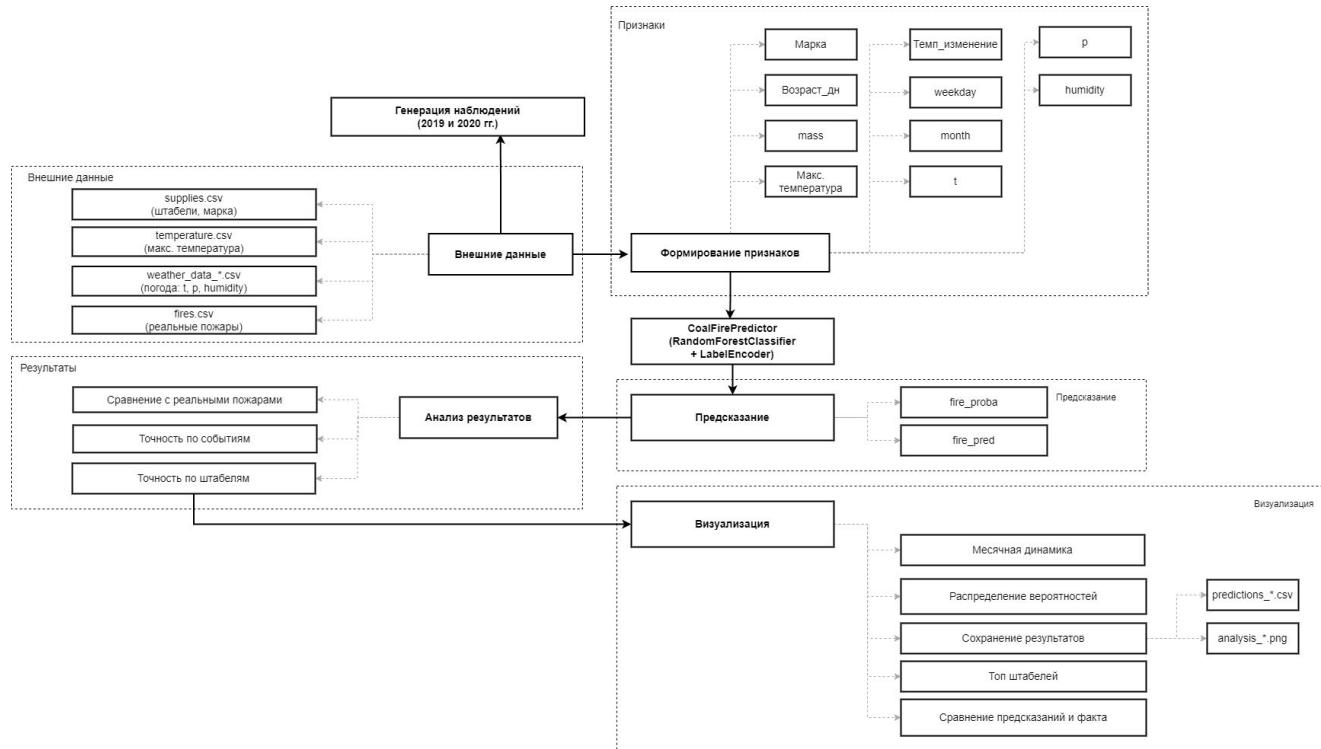
- Марка угля (категор. данные - LabelEncoder)
- Возраст штабеля (текущая дата – дата формирования)
- Масса штабеля в тоннах
- Максимальная температура штабеля
- Скорость изменения температуры
- Сезон/месяц
- Температура воздуха
- Атмосферное давление
- Влажность воздуха

**Обучение модели** в Colab



[https://colab.research.google.com/drive/1m\\_ay4lnoQD4758SvMLgGFeRq5rHr3EqJ?usp=sharing#scrollTo=-CTLsMiRRods](https://colab.research.google.com/drive/1m_ay4lnoQD4758SvMLgGFeRq5rHr3EqJ?usp=sharing#scrollTo=-CTLsMiRRods)

# Архитектура ML-системы для прогнозирования самовозгорания угля

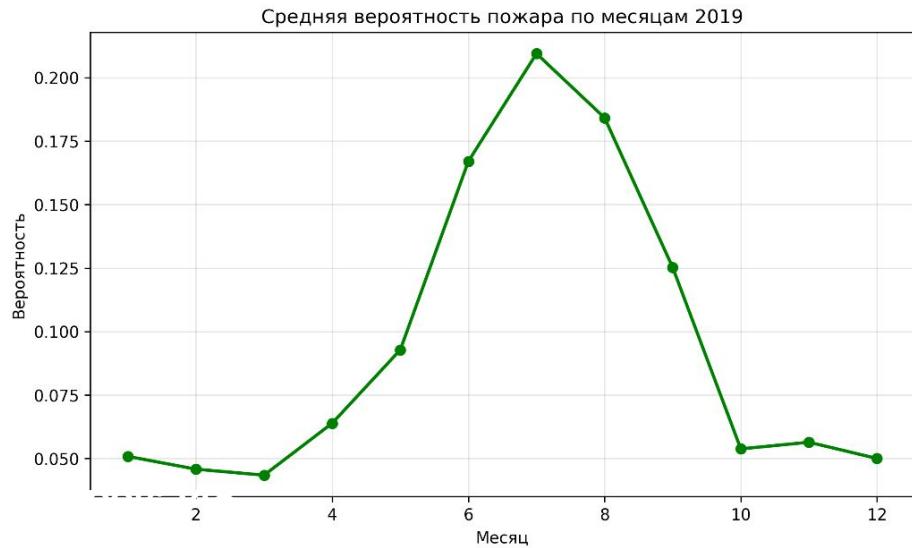


Описание представлено  
в виде отчета

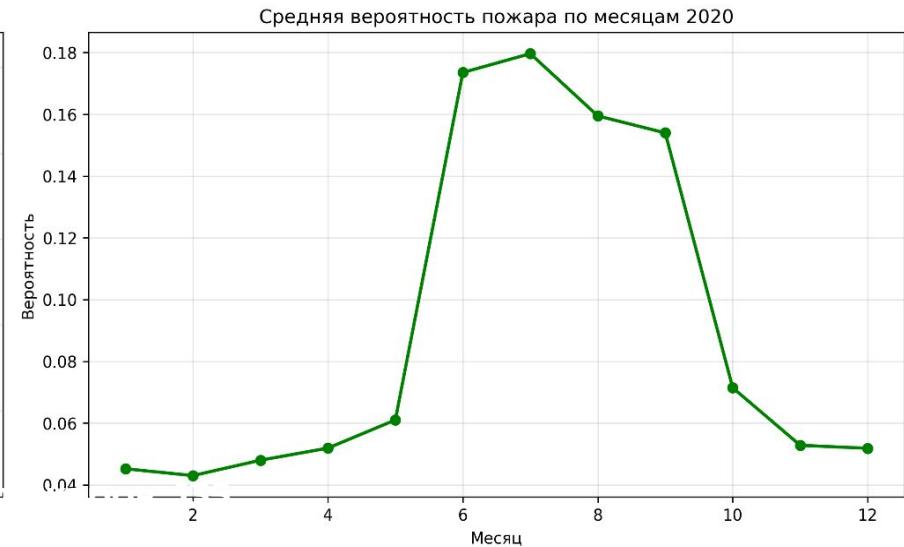


<https://disk.yandex.ru/i/wo2lct65rr5IAA>

# ML решение - анализ (тест 2019)



# ML решение - анализ (тест 2020)



# Метрики качества

## Precision

из всех случаев ложноположительных было 24%

## Recall

правильно обнаруженных пожаров в среднем 93%

## F1-score

дисбаланс precision x recall, НО важнее не пропустить потенциальный пожар

## ROC AUC

почти идеальное разделение классов, показывает способность модели выделять пожар среди множества случаев

Таблица - Метрики качества модели

### Метрики качества модели

Метрика	Значение
Precision	0.241
Recall	0.932
F1-score	0.383
ROC AUC	0.987

# Наша работа

Работа с инструкцией по запуску представлена в **GitHub**



[https://github.com/Kulikova-A18/glowbyte\\_hac\\_2025\\_command\\_14](https://github.com/Kulikova-A18/glowbyte_hac_2025_command_14)

**Спасибо за внимание!**

Команда 14

# GlowByte Хакатон

2025

Прогноз самовозгорания угля при  
открытом хранении

