МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«Московский государственный технический университет имени Н.Э. Баумана

(национальный исследовательский университет)»

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

**по курсу**

«Data Science»

**Тема:** «**Прогнозирование конечных свойств новых материалов**

**(композиционных материалов)»**

Слушатель Глорин Артём Васильевич

# 

Москва, 2025

# **Содержание**

[Содержание 2](#_Toc106232838)

[Введение 3](#_Toc106232839)

[1. Аналитическая часть 7](#_Toc106232840)

[1.1. Постановка задачи 7](#_Toc106232841)

[1.2. Описание используемых методов 10](#_Toc106232842)

[1.3. Разведочный анализ данных 16](#_Toc106232843)

[2. Практическая часть 21](#_Toc106232844)

[2.1. Предобработка данных 21](#_Toc106232845)

[2.2. Разработка и обучение модели 23](#_Toc106232846)

[2.3. Тестирование модели 25](#_Toc106232847)

[2.4. Нейронная сеть 26](#_Toc106232848)

[2.5. Разработка приложения 29](#_Toc106232849)

[2.6. Создание удалённого репозитория и загрузка 32](#_Toc106232850)

[2.7. Заключение 33](#_Toc106232851)

[2.8. Список используемой литературы и веб ресурсы. 33](#_Toc106232852)

# **Введение**

Композитные материалы представляют собой гетерогенные системы, объединяющие два или более компонента с различными физико-химическими свойствами, что позволяет достигать синергетического эффекта и превосходить характеристики традиционных материалов. Благодаря сочетанию высокой прочности, жесткости, коррозионной стойкости и малого веса композиты находят применение в авиакосмической, автомобильной, энергетической и медицинской отраслях. В данной статье рассматриваются ключевые типы композитных материалов, их классификация, преимущества и области применения, а также перспективные направления исследований, включая разработку «умных» и самовосстанавливающихся композитов.

История композитных материалов насчитывает тысячелетия: первые примеры их использования относятся к древним цивилизациям, где солома и глина применялись для укрепления строительных конструкций. Однако революционное развитие композитов началось в середине XX века с появлением синтетических полимеров и высокопрочных волокон. Сегодня композиционные материалы являются неотъемлемой частью современных технологий, обеспечивая прорывные решения в авиации, машиностроении, возобновляемой энергетике и биомедицине. Композитные материалы состоят из двух основных компонентов:

1. Матрицы – связующей основы, определяющей форму изделия и передающей нагрузку на армирующие элементы.

- Полимерная (эпоксидные смолы, термопласты),

- Металлическая (алюминиевые, магниевые, титановые сплавы),

- Керамическая (карбид кремния, оксид алюминия),

2. Армирующего элемента – усиливающего компонента, отвечающего за механические свойства.

- Волокна (углеродные, стеклянные, базальтовые, арамидные),

- Частицы (нанодисперсные порошки, керамические включения),

- Структурные наполнители (сотовые конструкции, многослойные композиты),

В зависимости от сочетания матрицы и армирования выделяют:

- Полимерные композиты (наиболее распространенные, например, углепластик),

- Металломатричные композиты (применяются в высокотемпературных условиях),

- Керамоматричные композиты (используются в экстремальных средах),

Ключевые преимущества композитов перед традиционными материалами:

- Высокая удельная прочность (прочность при минимальном весе),

- Коррозионная и химическая стойкость,

- Управляемые механические характеристики (возможность проектирования под конкретные задачи),

- Усталостная долговечность (устойчивость к циклическим нагрузкам),

- Диэлектрические и теплозащитные свойства.

Эти свойства делают композиты незаменимыми в конструкциях, где критичны масса, надежность и долговечность.

Применение:

1. Авиация и космонавтика

- До 50% массы современных самолетов (МС-21, Boeing 787, Airbus A350) составляют композиты.

- Лопатки турбин, обшивка корпусов, элементы шасси.

2. Автомобилестроение

- Карбоновые кузова спортивных автомобилей.

- Полимерные композиты в электромобилях для снижения веса.

3. Энергетика

- Лопасти ветрогенераторов из стекло- и углепластика.

- Теплоизоляционные композиты для солнечных панелей.

4. Медицина

- Биосовместимые имплантаты (углеродные и полимерные).

- Композитные протезы и ортопедические конструкции.

5. Строительство

- Армированные бетоны с фиброволокном.

- Композитные панели для фасадов и мостов.

6. Военное применение

- Лопасти вертолета КА-52,

- Средства снижения радиолокационной заметности СУ-57,

- Корпуса БПЛА,

- Элементы корпуса и динамической защиты Т-14 «Армата»,

- Морские дроны,

- Средства индивидуальной бронезащиты,

- Всевозможные «обвесы» стрелкового оружия.

Современные исследования направлены на:

- Самовосстанавливающиеся композиты (с микрокапсулами полимера или биомиметическими структурами) для автоматического «залечивания» пробоин.

- Наноармированные материалы (графен, углеродные нанотрубки) для сверхпрочной брони.

- Адаптивная маскировка – материалы, меняющие цвет или свойства в зависимости от среды.

- Экологичные и перерабатываемые композиты (биополимеры, разлагаемые матрицы).

- «Умные» композиты (с интегрированными датчиками для мониторинга состояния).

Композитные материалы продолжают трансформировать инженерные решения, предлагая уникальное сочетание легкости, прочности и функциональности. Дальнейшее развитие технологий их производства и переработки откроет новые возможности в создании высокоэффективных и экологически устойчивых материалов для технологий будущего.

1. **Аналитическая часть**
   1. **Постановка задачи**

Для выполнения данной задачи были предоставлены 2 файла: X\_bp.csv (с данными о параметрах, состоящими из 1024 строки и 11 столбцов) и X\_nup.csv (c данными нашивок, состоящими из 1041 строки и 4 столбцов).

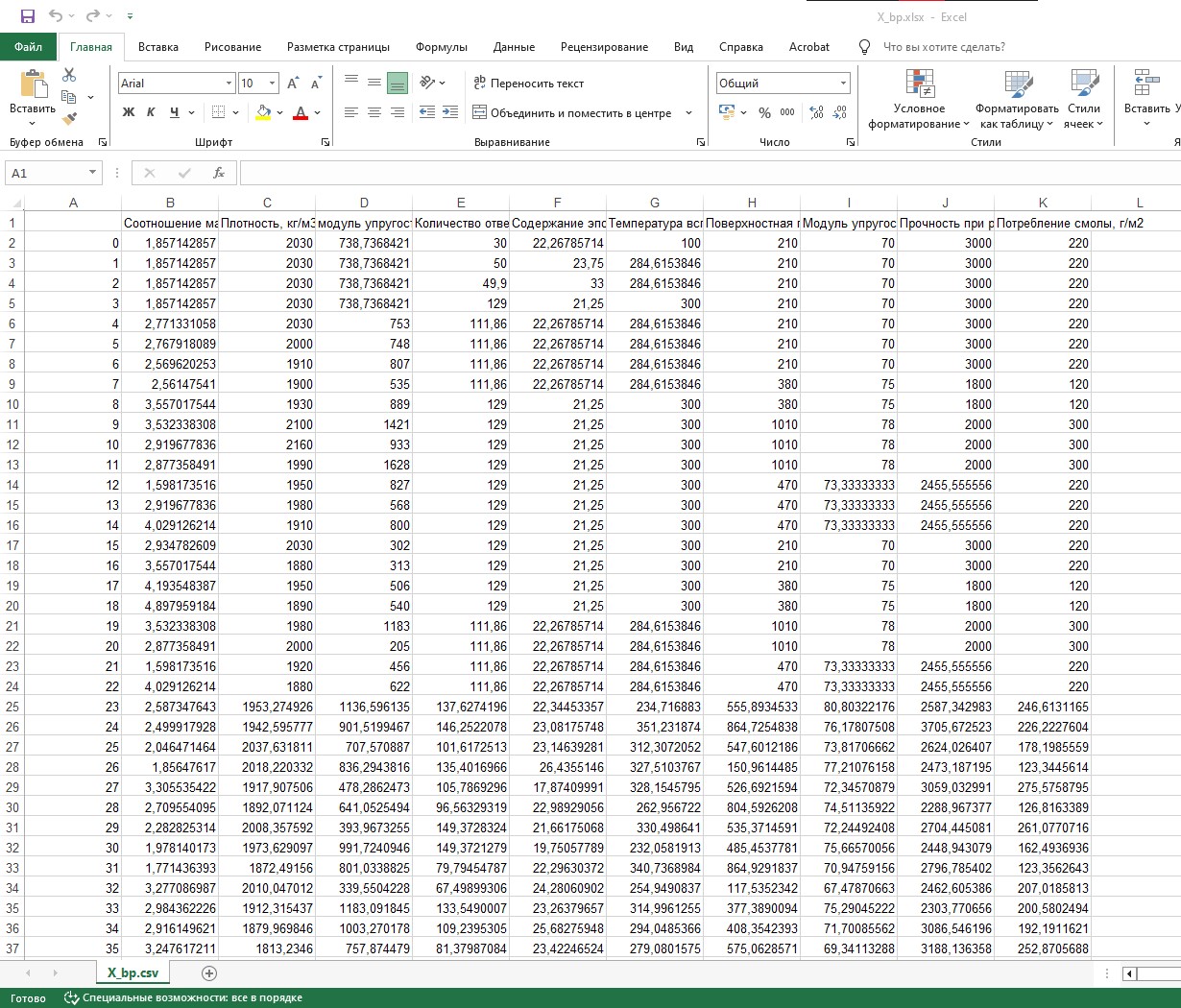


Рисунок 1 – файл X\_bp.csv (фрагмент)

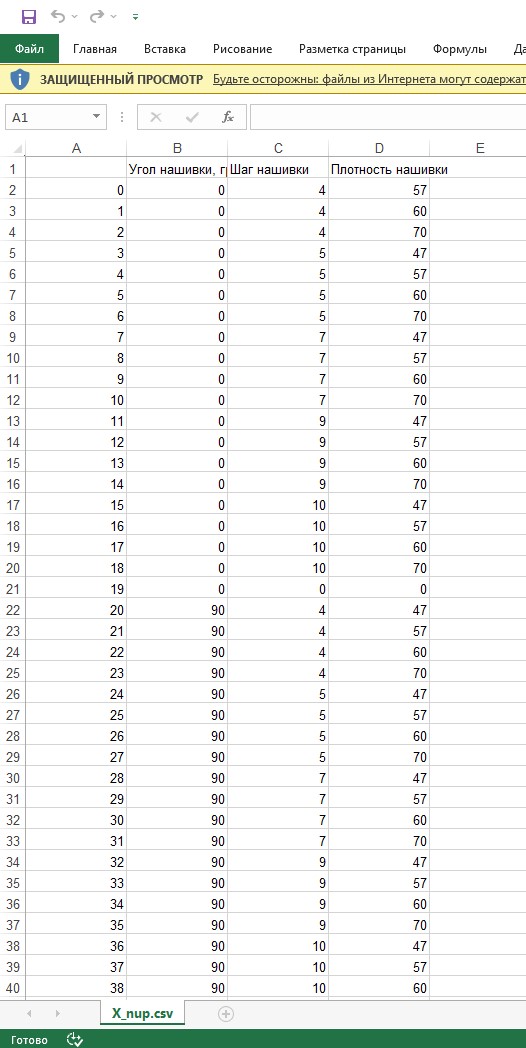


Рисунок 2 – файл X\_nup.csv (фрагмент)

Проект направлен на создание моделей прогнозирования модуля упругости и прочности при растяжении, а также на разработку нейросетевого алгоритма для рекомендации состава «матрица-наполнитель».

На подготовительном этапе данные из файлов X\_bp и X\_nup были объединены в единый датасет для дальнейшего комплексного анализа.

Первым шагом работы стало создание датафрейма (df) из импортированных данных. Датафрейм включает 1023 строки и 13 столбцов.

В процессе загрузки данных в качестве индекса был взят столбец «ID».

Метод df.isnull().sum() вывел отсутствие пропусков по каждому столбцу, а метод df.info() выявил, что параметр «угол нашивки, град» принадлежит к типу данных int, который позже был приведен к типу float.

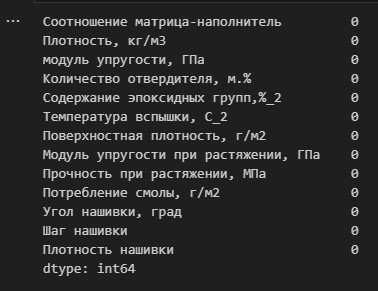


Рис. 3 Пропуски.

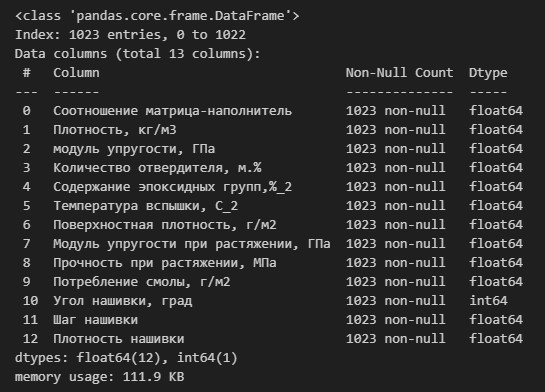


Рис 4. Типы данных

* 1. **Описание используемых методов**

Анализ исходных данных выявил потребность в углублённом разведочном анализе, который предполагает исследование распределений с использованием гистограмм и выявление аномалий посредством диаграмм Boxplot.

В данном проекте были применены следующие методы:

1. метод межквартильного размаха IQR для обработки выбросов;

2. метод K-means++ для кластеризации;

3. Gradient Boosting Regressor для обучения первой модели;

4. AdaBoost Regressor для обучения второй модели.

1. Метод межквартильного размаха IQR – устойчивая к выбросам половина выборки относительно медианы. Это удобный показатель изменчивости признаков для асимметричных распределений и данных с аномальными значениями.  
 Преимущества:

а) Робастность к выбросам — в отличие от стандартного отклонения, IQR не зависит от экстремальных значений.

б) Простота интерпретации — чётко показывает, где сосредоточено 50% центральных данных.

в) Работает с ненормальными распределениями — не требует симметричности данных.

г) Универсальность — применим для любых непрерывных данных (финансы, медицина, инженерия).

д) Используется в box-plot — визуализация разброса и выбросов становится наглядной.

Недостатки:

а) Игнорирует форму распределения — не различает симметричные и асимметричные данные.

б) Может пропускать выбросы — если аномалии находятся близко к Q1/Q3, но не за границами ±1.5×IQR.

в) Не учитывает все точки данных — использует только 25-й и 75-й процентили, "теряя" остальную информацию.

г) Эмпирический коэффициент (1.5) — границы выбросов выбраны условно и могут не подходить для некоторых задач.

2. K-means++ — это улучшенная версия алгоритма K-means для кластеризации данных, которая предлагает более эффективный способ выбора начальных центроидов (начальных точек кластеров). Основная цель — избежать плохой инициализации, которая может привести к неоптимальным результатам в стандартном K-means.

Ключевая идея:

-Первый центроид выбирается случайно из всех точек данных.

-Каждый следующий центроид выбирается с вероятностью, пропорциональной квадрату расстояния до ближайшего уже выбранного центроида.

-Процесс повторяется, пока не будут выбраны все k центроидов.

Преимущества:

а) Более быстрая сходимость - уменьшает количество итераций, необходимых для стабилизации кластеров.

б) Устойчивость к начальным условиям – позволяет лучше распределить центроиды на начальном этапе и уменьшает вероятность попадания алгоритма в локальные минимумы.

Недостатки:

а) Не решает все проблемы K-means:

-чувствительность к форме кластеров (лучше работает с выпуклыми кластерами одинакового размера),

-не подходит для данных с шумом или выбросами (можно комбинировать с методами вроде DBSCAN).

б) Зависимость от метрики расстояния — как и K-means, использует евклидово расстояние, что может быть неудачным выбором для данных сложной структуры.

в) Требует заданного числа кластеров (k) — как и оригинальный K-means, нуждается в предварительной оценке k (например, через "метод локтя").

г) Вычислительная нагрузка при инициализации — выбор центроидов сложнее, чем случайный, но всё равно быстрее, чем, например, иерархическая кластеризация.

3. Gradient Boosting Regressor (GBR) — это ансамблевый алгоритм машинного обучения, который последовательно строит деревья решений, каждое из которых корректирует ошибки предыдущих, минимизируя заданную функцию потерь с помощью градиентного спуска.

Преимущества:

а) Высокая точность – Один из самых мощных алгоритмов машинного обучения, часто даёт лучшие результаты по сравнению с линейными моделями и случайным лесом.

б) Гибкость – Может работать с разными функциями потерь (MSE, MAE, Huber) и поддерживает кастомные.

в) Автоматическая обработка нелинейных зависимостей – Хорошо улавливает сложные взаимосвязи в данных без явной feature engineering.

г) Устойчивость к выбросам (при использовании MAE или Huber loss) – Менее чувствителен к аномалиям, чем, например, линейная регрессия.

д) Встроенный отбор признаков – Учитывает важность фичей, что помогает в интерпретации модели.

е)Работает с разными типами данных – Может обрабатывать как числовые, так и категориальные признаки (после кодирования).

Недостатки:

а) Склонность к переобучению – Требует тщательной настройки гиперпараметров (глубина деревьев, learning rate, число итераций).

б) Вычислительная сложность – Обучение может быть медленным на больших датасетах по сравнению с Random Forest.

в) Чувствительность к шуму и выбросам (при использовании MSE) – Может подстраиваться под аномалии.

г) Требует больше памяти – По сравнению с линейными моделями из-за хранения ансамбля деревьев.

д) Менее интерпретируем, чем линейные модели – Хотя важность признаков можно оценить, сам механизм предсказания сложнее для объяснения.

е) Зависимость от начальных параметров – Результаты могут сильно меняться при разных random\_state и learning rate.

5. AdaBoost Regressor – ансамблевый метод, в основе которого лежит процедура бустинга, предполагающая итеративное построение ансамбля, где на каждом шаге новая слабая модель оптимизируется относительно остаточных ошибок, полученных от композиции предыдущих. Это обеспечивает адаптивное улучшение совокупного прогноза.

Преимущества:

а) Улучшение точности - бустинг позволяет значительно улучшить точность модели путем комбинирования множества слабых прогнозов. В задачах регрессии это достигается через усреднение их прогнозов, а в задачах классификации - через механизм голосования, что в совокупности дает более точный финальный результат.

б) Устойчивость к переобучению - бустинг может снизить риск переобучения за счет присвоения весов неправильно классифицированным входным данным.

в) Улучшенная обработка несбалансированных данных — бустинг позволяет справиться с несбалансированными данными, уделяя больше внимания точкам данных, которые неправильно классифицированы.

г) Лучшая интерпретируемость — бустинг может повысить интерпретируемость модели за счет разбиения процесса принятия решения на несколько процессов.

Недостатки:

а) Чрезмерная чувствительность к выбросам.

б) AdaBoost создаёт сложные модели, которые работают как 'чёрный ящик' - даже если он использует простые деревья, их комбинация становится настолько запутанной, что невозможно понять, как именно модель принимает решения.

в) Методология AdaBoost не позволяет эффективно конструировать ансамбли малого размера (при небольшом T) из алгоритмов высокой сложности, к которым относятся SVM и нейронные сети.

г) Необходима большая обучающая выборка (бэггинг может обходиться более маленькой выборкой).

Сводная таблица предпосылок работоспособности каждого метода.

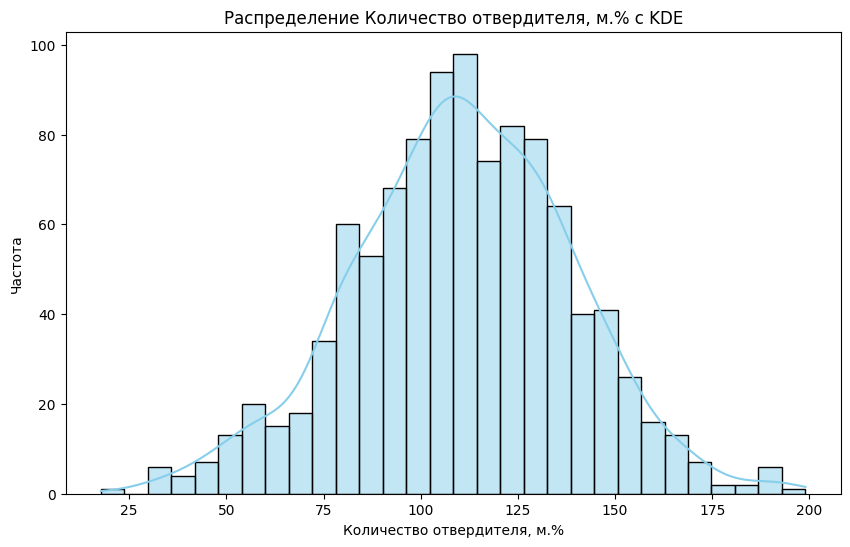
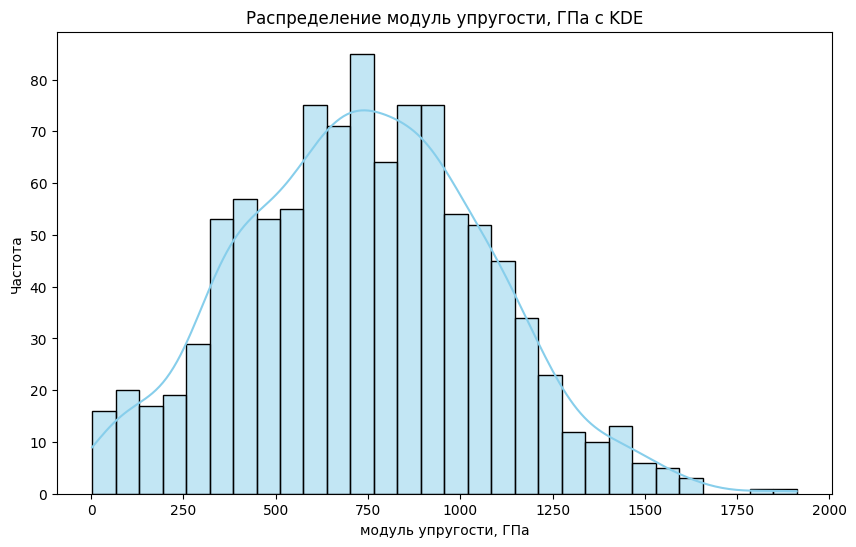
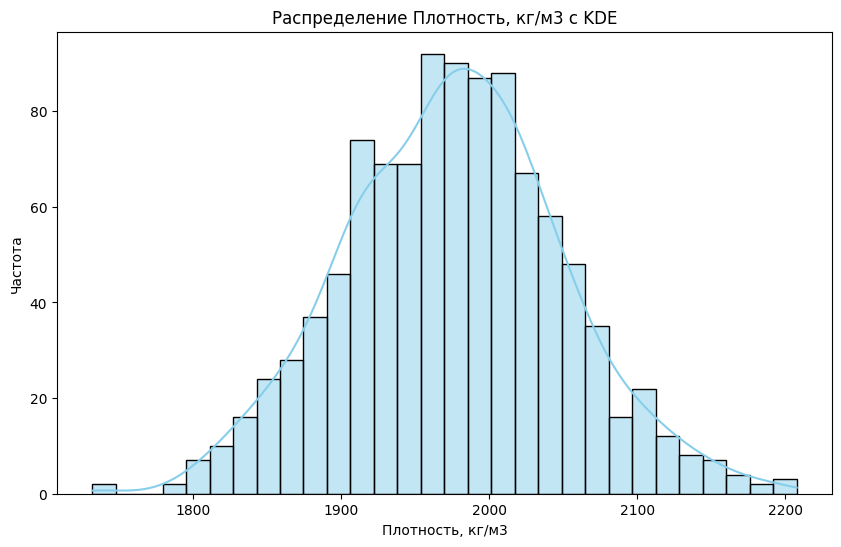
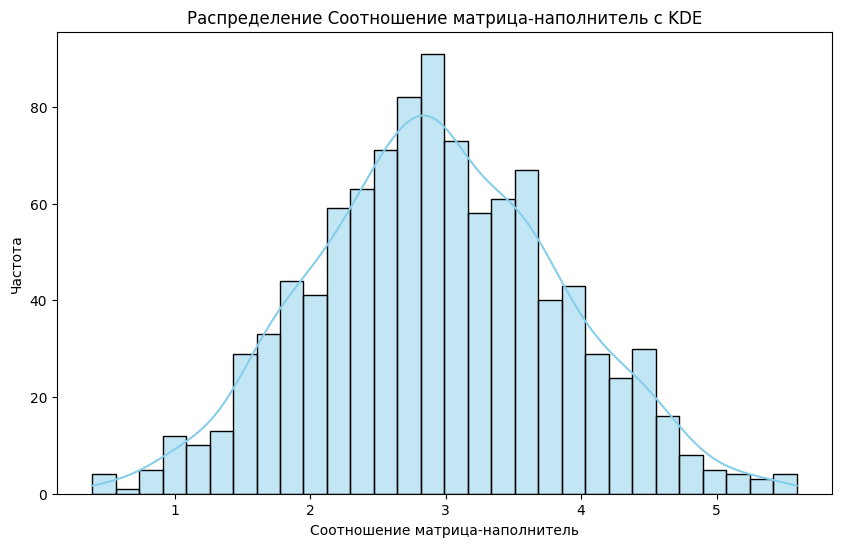
|  |  |  |
| --- | --- | --- |
| Метод | Тип задачи | Предпосылки применения |
| Межквартильный размах | Обнаружение и обработка выбросов | Ненормальное распределение некоторых признаков |
| K-means++ | Кластеризация | Числовые данные, известно k (методом локтя) |
| Gradient Boosting Regressor (GBR) | Регрессия | Для задач с небольшими/средними объёмами данных, где важна точность |
| AdaBoost | Регрессия | Нужна высокая точность при несбалансированных данных и слабых моделях |

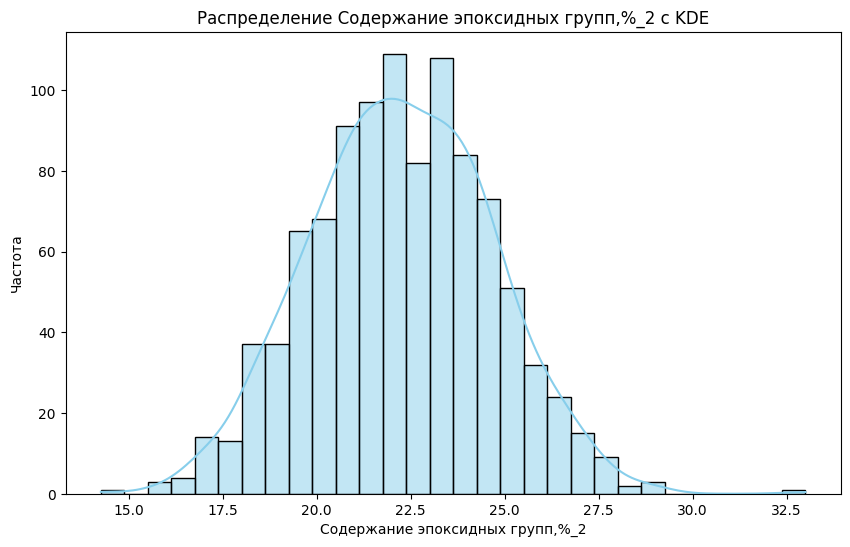
* 1. **Разведочный анализ данных**

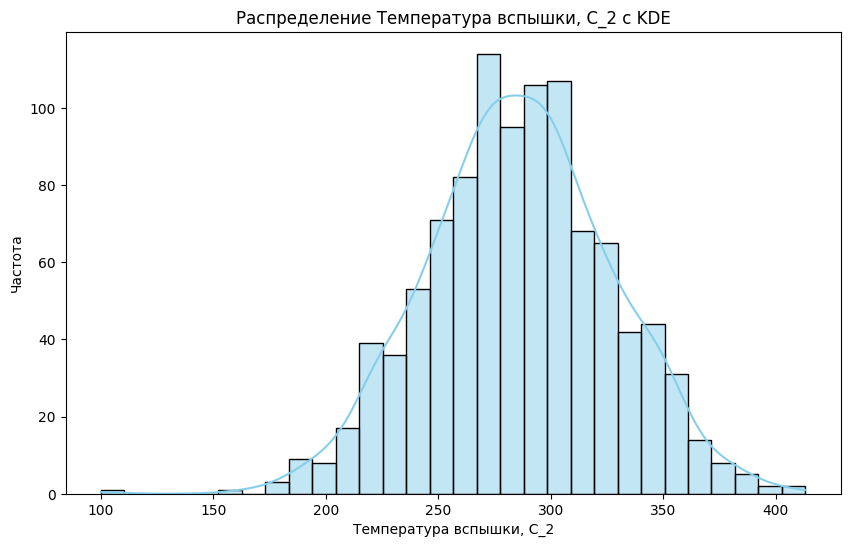
Первоочередная цель текущего этапа - визуализировать распределение значений всех признаков с помощью гистограмм. Такой подход даст возможность:

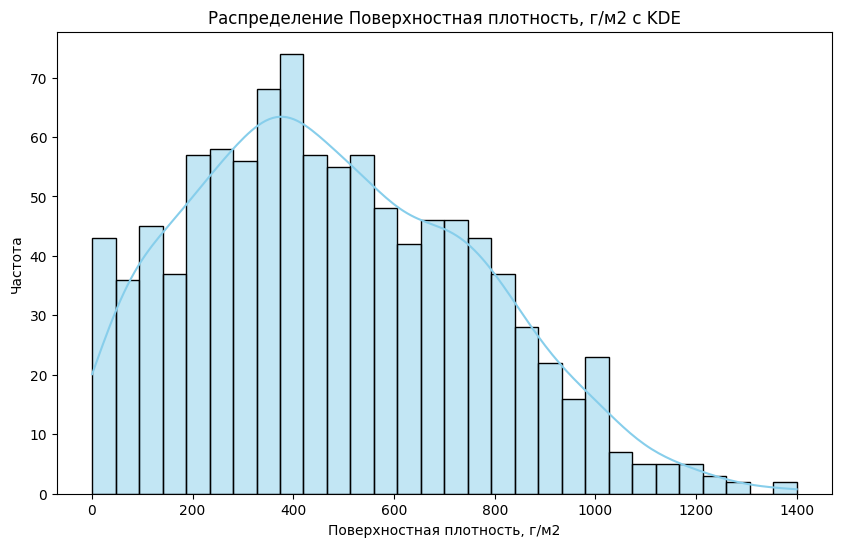
* выявить основные закономерности и тенденции в данных
* обнаружить скрытые взаимосвязи между переменными
* проанализировать частотность попадания значений в различные интервалы
* определить наличие аномалий в виде резких всплесков или спадов распределения

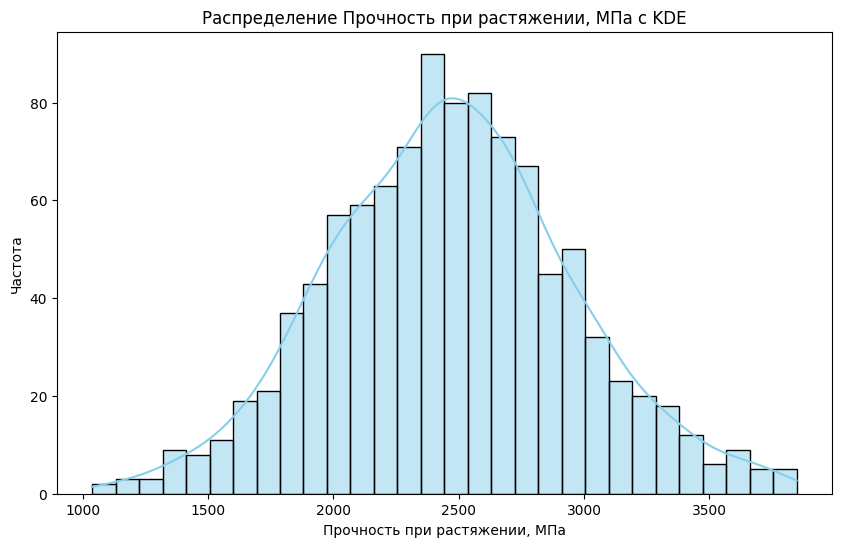
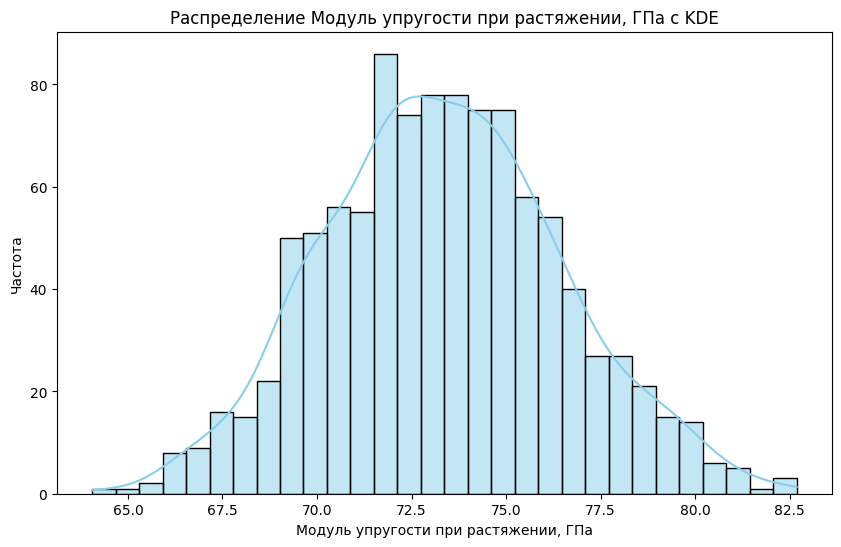
Гистограммное представление данных является эффективным инструментом для первичного анализа структуры информации и выявления ключевых характеристик каждого признака.

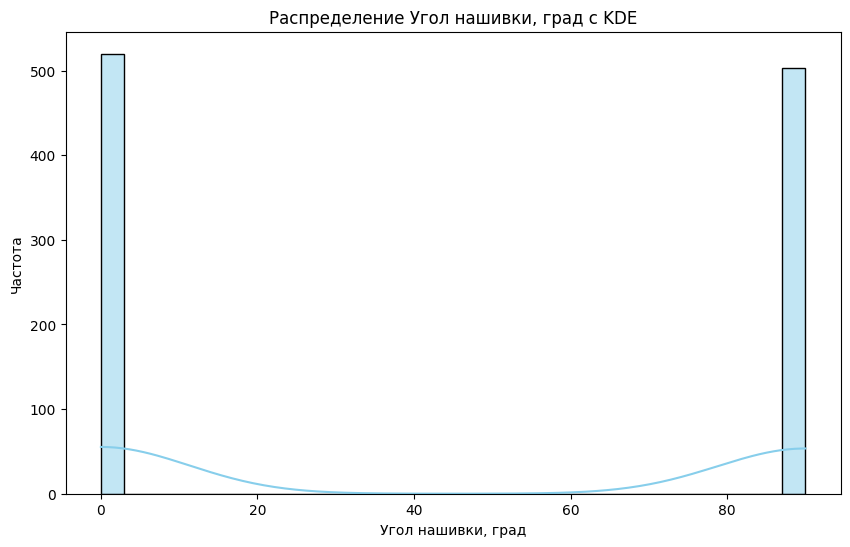
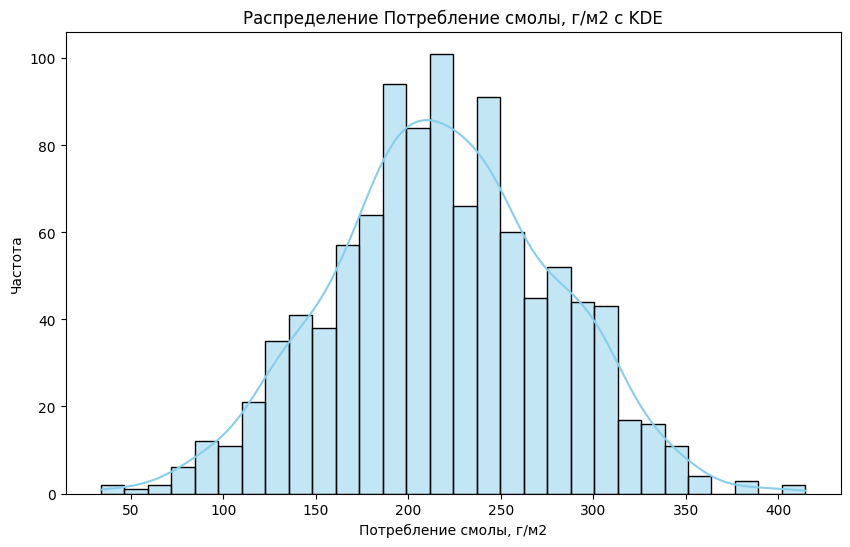


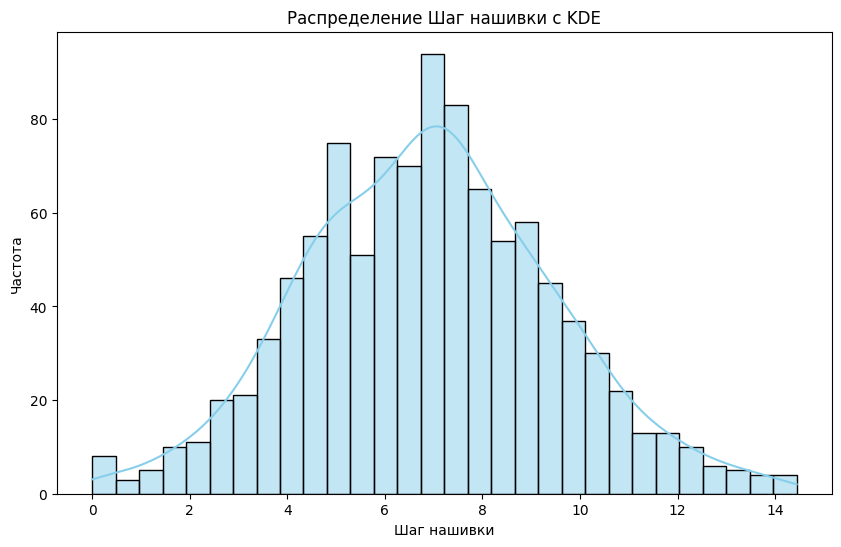












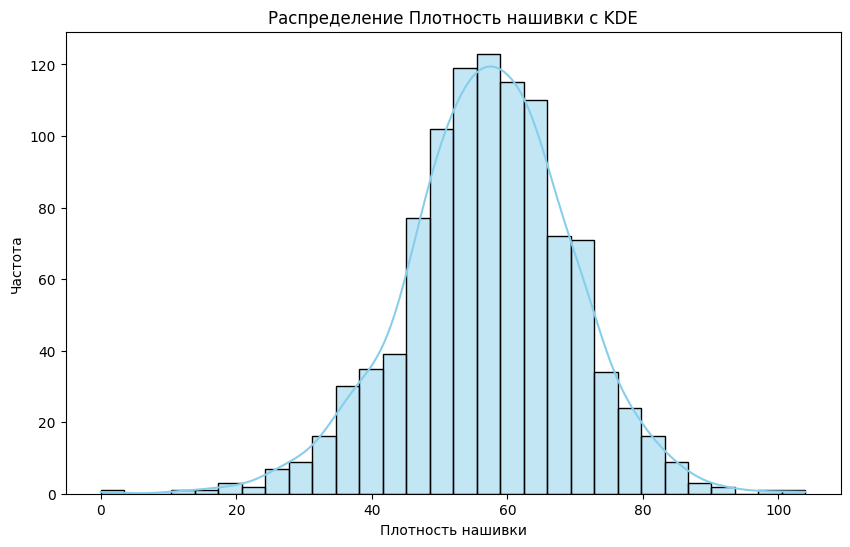


Рис. 5-17 – гистограммы признаков по столбцам.

Тепловая карта – это визуализация данных в виде матрицы, где цвет ячеек соответствует величине коэффициентов корреляции между признаками.

Она позволяет оценить линейные взаимосвязи между переменными: чем интенсивнее цвет, тем сильнее зависимость. В данном случае анализ показывает очень слабую корреляционную связь между исследуемыми параметрами.

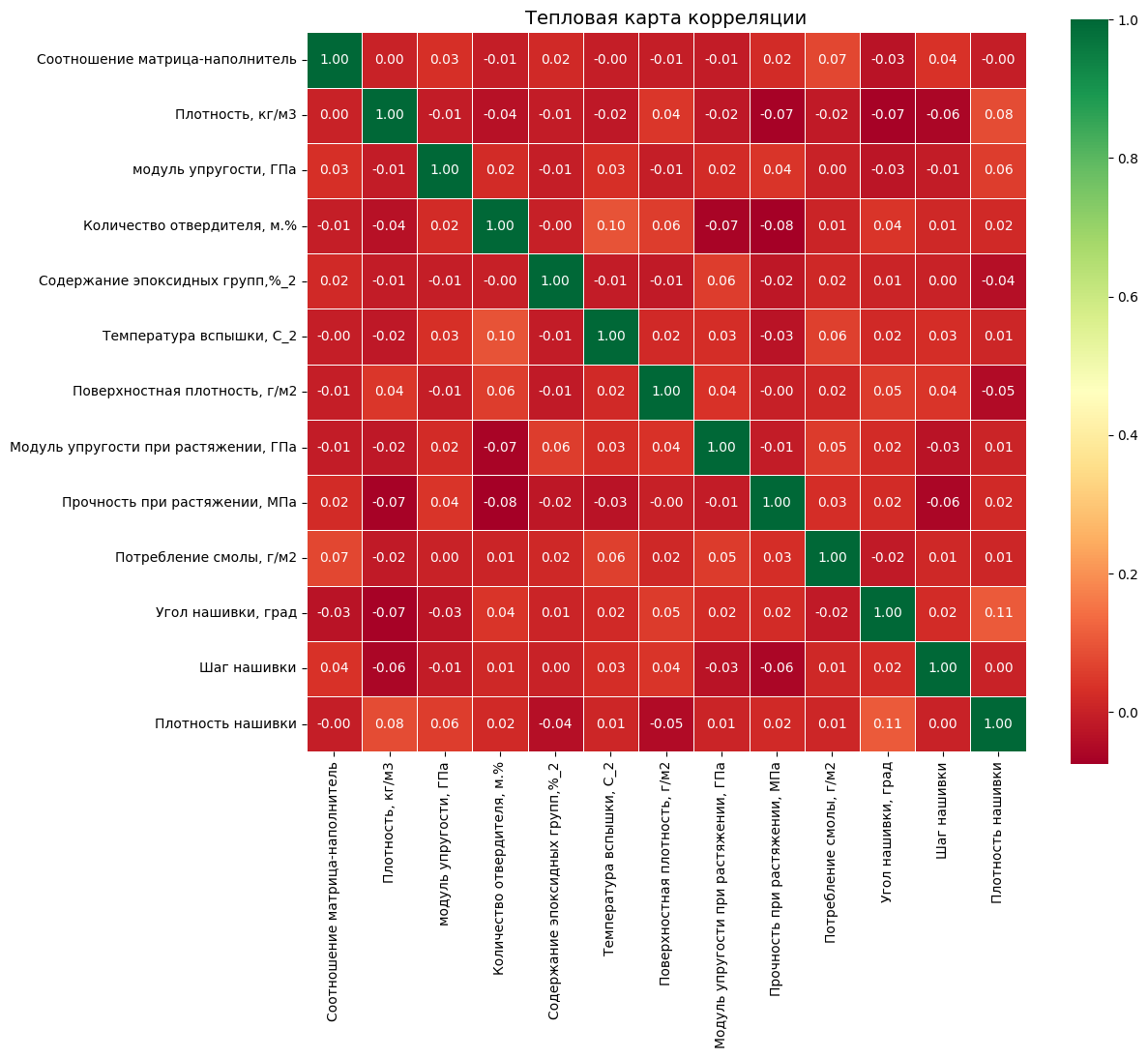


Рис. 18 – тепловая карта корреляции.

Также расчет корреляции Пирсона и Кендалла не выявил очевидных взаимосвязей.

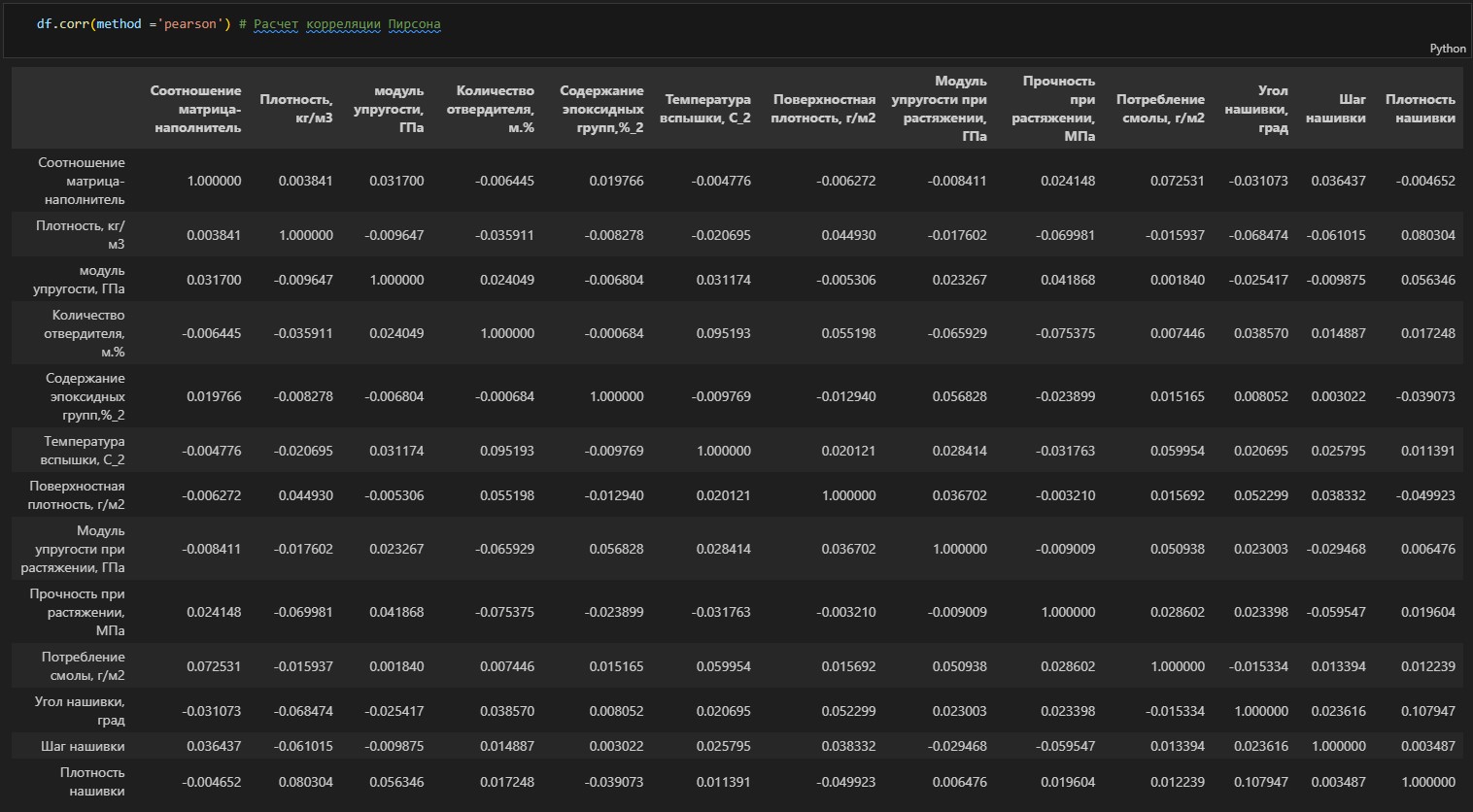


Рис. 19 – расчет корреляции Пирсона

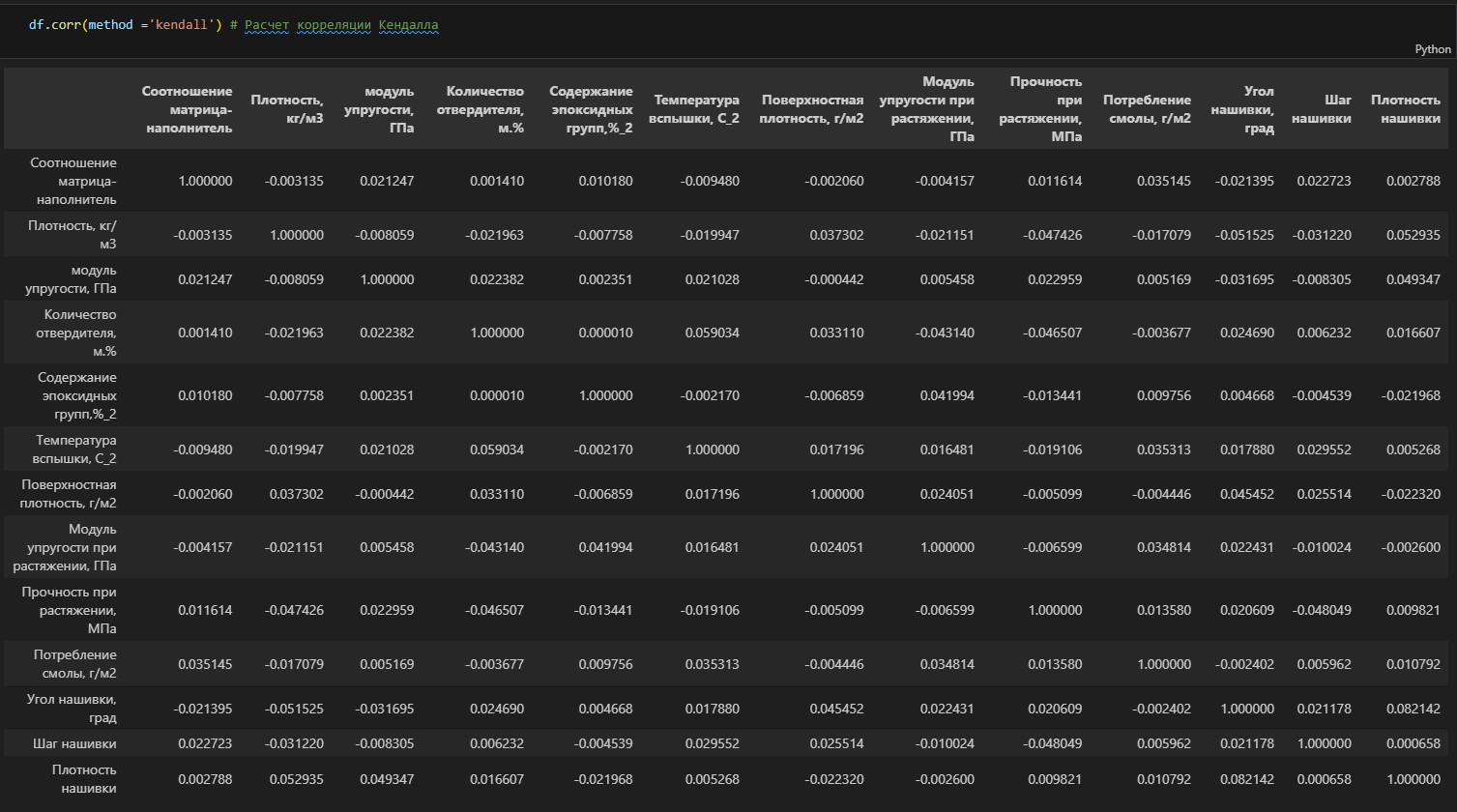


Рис. 20 – расчет корреляции Кендалла.

Выбросы были обработаны методом межквартильного размаха, так как он показал лучший результат.

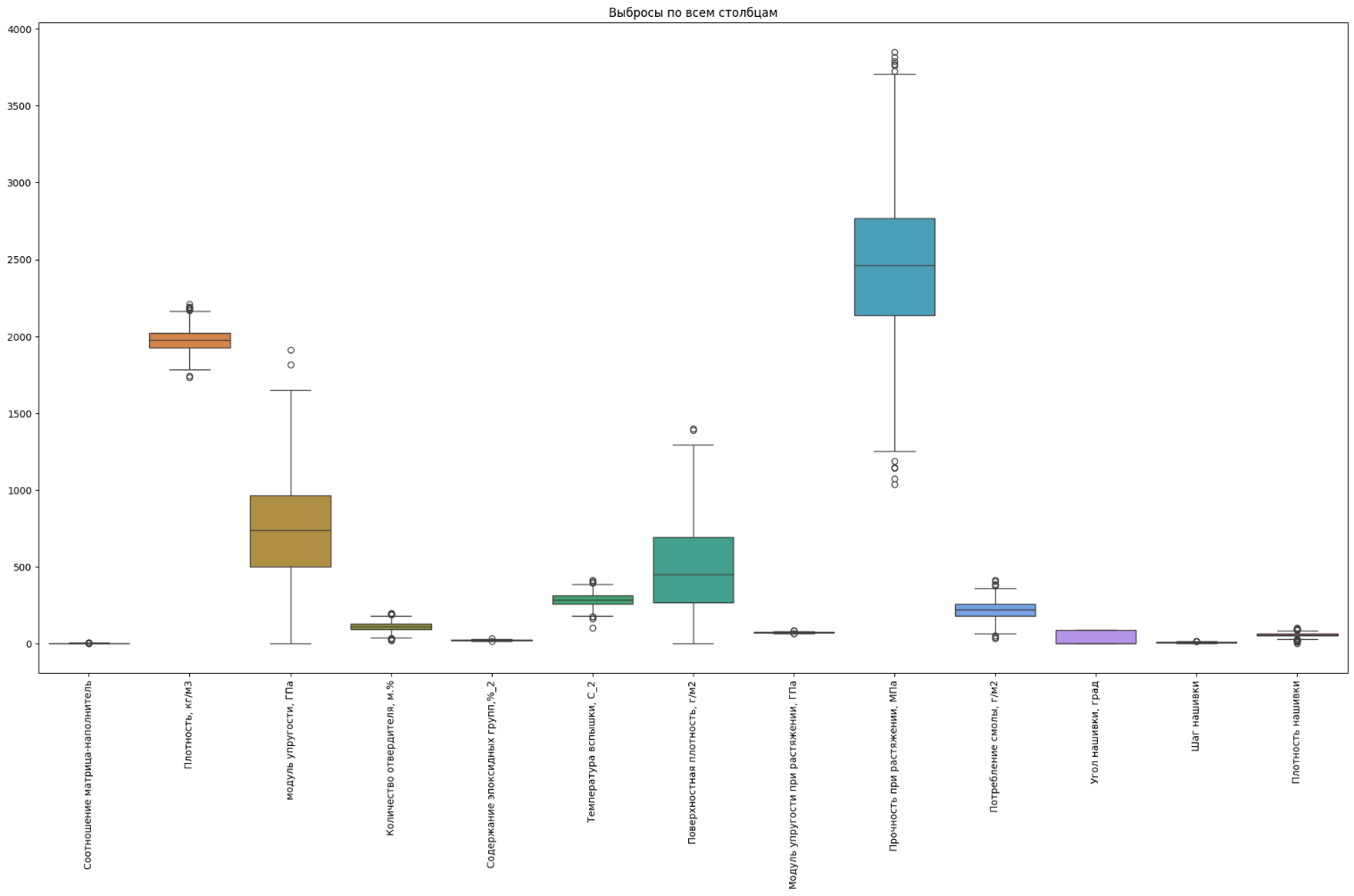


Рис. 21 – обработка выбросов методом IQR.

Ограниченный объем данных обусловил выбор метода обработки выбросов - их замену средними значениями, что минимизирует риск недообучения моделей по сравнению с полным удалением аномалий.

1. **Практическая часть**
   1. **Предобработка данных**

На этапе предобработки данных была произведена нормализация при помощи Standard Scaler и кластеризация при помощи KMeans++.

Методом локтя было выявлено оптимальное количество кластеров и их данные были добавлены в датафрейм.

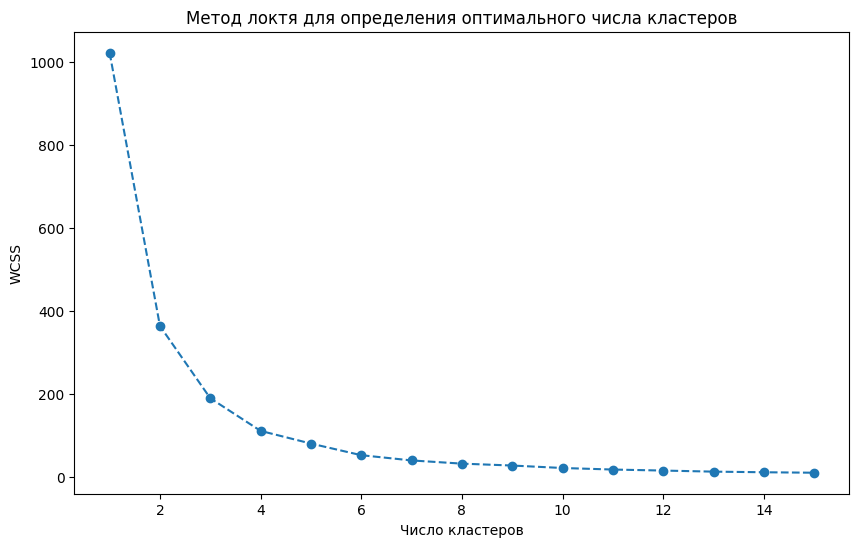
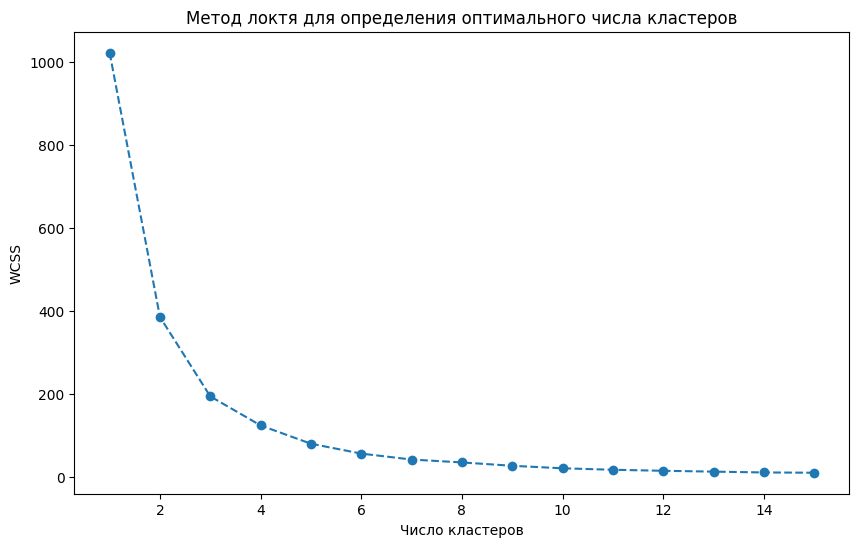


Рис. 22, 23 – определение оптимального количества кластеров методом локтя.

**2.2 Разработка и обучение модели**

В процессе создания и обучения моделей применялись два алгоритма: Gradient Boosting Regressor и AdaBoost Regressor. После тщательного ручного подбора и перебора гиперпараметров было установлено, что Gradient Boosting Regressor демонстрирует более высокую эффективность в первом случае, тогда как AdaBoost Regressor оказался предпочтительнее во втором. Для оценки качества моделей исходные данные были разделены на обучающий (70%) и тестовый (30%) наборы. Была проведена предварительная обработка данных для последующего обучения алгоритмов: из набора переменных были определены предикторы и целевой показатель.

На основе имеющегося набора данных создан обучающий алгоритм Gradient Boosting Regressor из 100 последовательных деревьев решений со скоростью обучения 0,1 и глубиной деревьев 7.

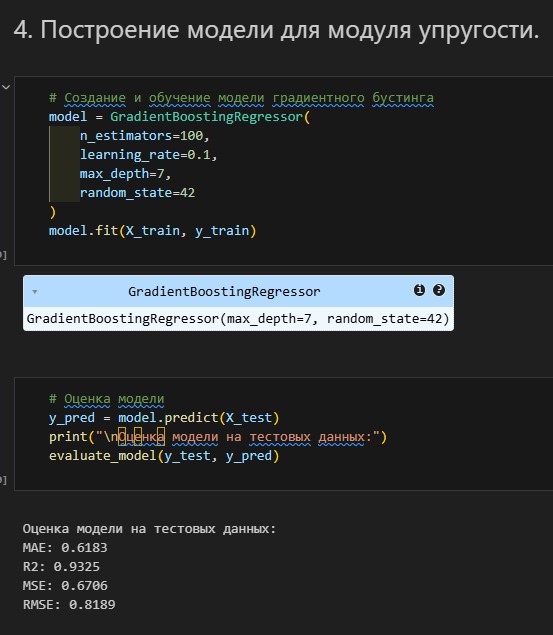


Рис. 24 – создание модели 1 и метрики

Для второй модели действия, начиная от кластеризации, были аналогичными.

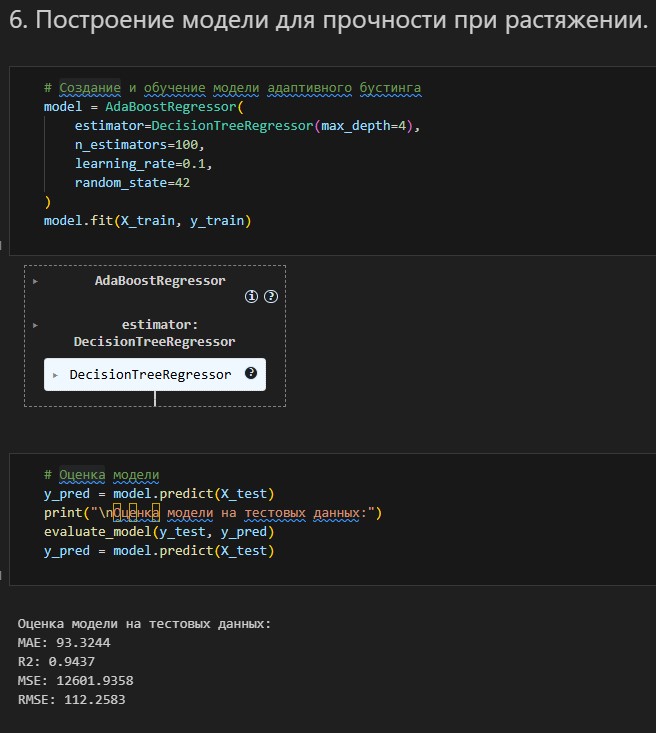


Рис. 25 - создание модели 2 и метрики

**2.3 Тестирование модели**

В связи с регрессионной природой задачи, метрики будут представлять собой среднюю оценку разности между действительными и спрогнозированными значениями.

MAE (Mean Absolute Error) – оценивает насколько в среднем прогнозы модели отклоняются от средних значений по модулю. MAE измеряется в тех же единицах, что и исходные данные. Чем выше значение MAE, тем больше ошибок в прогнозах модели. Модель 1 в среднем ошибается на ±0.6183 единицы измерения целевой переменной. Модель 2 в среднем ошибается на ±93.3244 единицы измерения целевой переменной.

R2 – коэффициент детерминации. Показывает, какую долю дисперсии целевой переменной объясняет модель. Аналог точности в процентах для задач регрессии. Модель 1 объясняет 93.25% изменчивости данных, модель 2 объясняет 94.37% изменчивости данных. Эти значения можно интерпретировать как то, что модель 1 имеет очень высокое значение R2 и объясняет почти всю вариацию зависимой переменной. Модель 2 имеет высокое значение R2 и объясняет большую часть вариации зависимой переменной. Данные значения согласуются с предметной областью, т.к. в физических экспериментах ожидается высокая точность.

MSE (Mean Squared Error) – показывает насколько в среднем прогнозы модели отклоняются от реальных значений в квадрате. То есть среднее значение квадратов разностей между фактическими и предсказанными значениями. MSE используется в регрессионных задачах основной целью которых является предсказание числовых значений. Чем меньше значение MSE, тем лучше. В связи с тем, что MSE в отличие от MAE не сохраняет единицы измерения исходных данных - корректная интерпретация значений MSE сложна. Для того, чтобы показатель эффективности MSE имел размерность исходных данных, из него извлекают квадратный корень и получают показатель эффективности RMSE.

Таким образом для анализа вариации ошибок лучше использовать MAE и RMSE.

На этапе тестирования модели показали следующие результаты:

Модель 1 (Gradient Boosting Regressor) –

MAE: 0.6183

RMSE: 0.8189

R2: 0.9325

MSE: 0.6706

Модель 2 (AdaBoost Regressor) –

MAE: 93.3244

RMSE: 112.2583

R2: 0.9437

MSE: 12601.9358.

Отношение MAE/RMSE для модели 1 составляет 0.755. Модель демонстрирует умеренно согласованные MAE и RMSE, без аномально крупных ошибок.

Отношение MAE/RMSE для модели 2 – 0.831. Значение говорит о наличии некоторых аномалий в данных, ошибки распределены не строго нормально, но и не катастрофично.

**2.4 Нейронная сеть, рекомендующая соотношение матрица-наполнитель.**

Представленная нейронная сеть представляет собой полносвязную (dense) архитектуру (многослойный перцептрон), предназначенную для решения задачи регрессии. Основная цель - предсказание соотношения матрица-наполнитель на основе 12 физико-химических характеристик композитных материалов.

Архитектура сети

1. Входной слой:
   * Принимает 12 признаков (размерность входных данных)
   * Используется стандартизация данных (StandardScaler)
2. Скрытые слои:
   * Два идентичных скрытых слоя по 16 нейронов каждый
   * Функция активации: ReLU (Rectified Linear Unit)
   * Dropout-слои с вероятностью 0.2 для регуляризации
3. Выходной слой:
   * Один нейрон с линейной активацией (для регрессии)

Ключевые особенности

1. Регуляризация:
   * Dropout (20%) для предотвращения переобучения
   * Ранняя остановка (EarlyStopping) с отслеживанием валидационных потерь
2. Оптимизация:
   * Оптимизатор: Adam с learning\_rate=0.001
   * Функция потерь: MSE (Mean Squared Error)
   * Дополнительные метрики: MAE и MAPE
3. Обучение:
   * Размер батча: 32
   * Максимальное число эпох: 100
   * 5-кратная кросс-валидация
   * Разделение данных: 70/30 (train/test)

Метрики оценки

1. Основные метрики:
   * MAE (Mean Absolute Error) - 0.1465
   * MSE (Mean Squared Error) - 0.0236
   * RMSE (Root Mean Squared Error) - 0.1535
   * R² (Коэффициент детерминации) - 0.9717
   * MAPE (Mean Absolute Percentage Error) – 5.0

(Отношение MAE/RMSE для модели составляет 0.95. Ошибки модели распределены относительно симметрично)

1. Визуализация:
   * Графики обучения/валидации
   * Диаграмма рассеяния предсказаний
   * Важность признаков

Особенности реализации

1. Предобработка данных:
   * Автоматическое выделение признаков и целевой переменной
   * Масштабирование признаков для каждого фолда отдельно
2. Воспроизводимость:
   * Фиксированный random\_state (42)
   * Сохранение и загрузка модели
3. Интерпретируемость:
   * Визуализация важности признаков
   * Подробный вывод метрик.

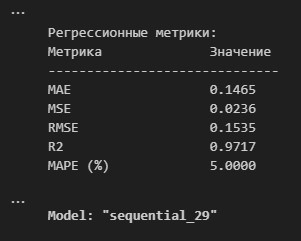


Рис.26 – метрики модели

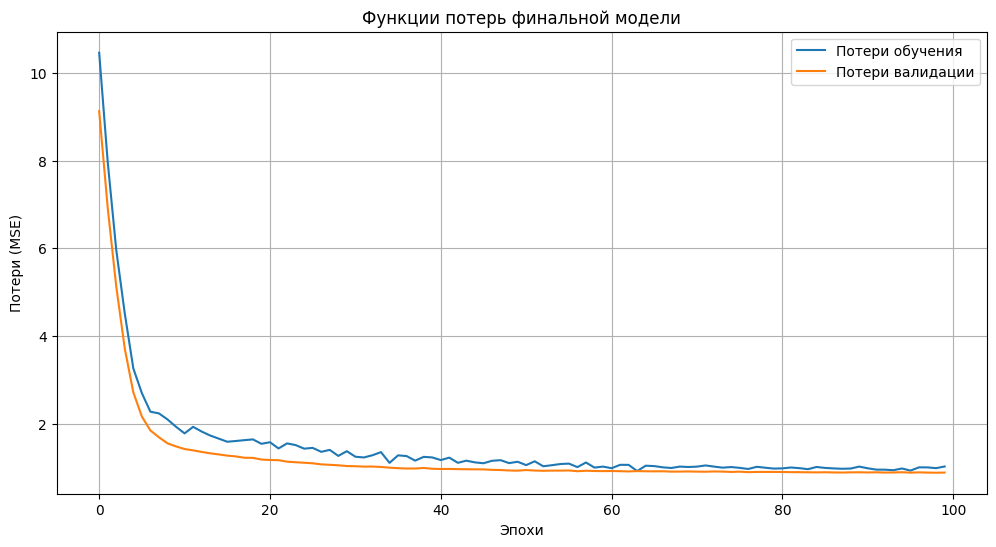


Рис.27 – график функций потерь

**2.5 Разработка приложения**

Это веб-приложение на Flask, которое использует предварительно обученную нейронную сеть (Keras/TensorFlow) для выполнения прогнозов на основе введенных пользователем параметров. Приложение предоставляет простой интерфейс для ввода данных и отображения результатов предсказания.

Основной функционал

1. Загрузка модели нейронной сети

- При запуске приложения загружается модель из файла 'dnn\_model.keras'

- Проверяется существование файла модели перед загрузкой

2. Веб-интерфейс

- Главная страница с формой для ввода параметров

- Обработка GET и POST запросов

3. Обработка входных данных

- Получение 12 числовых параметров от пользователя через форму:

1) Плотность, кг/м3

2) Модуль упругости, ГПа

3) Количество отвердителя, м.%

4) Содержание эпоксидных групп, %\_2

5) Температура вспышки, С\_2

6) Поверхностная плотность, г/м2

7) Модуль упругости при растяжении, ГПа

8) Прочность при растяжении, МПа

9) Потребление смолы, г/м2

10) Угол нашивки, град

11) Шаг нашивки

12) Плотность нашивки

4. Прогнозирование

- Преобразование входных данных в формат, подходящий для нейронной сети

- Выполнение предсказания с использованием загруженной модели

- Возвращение результата пользователю

5. Обработка ошибок

- Перехват и обработка возможных ошибок при вводе данных или предсказании

- Отображение сообщений об ошибках пользователю

Особенности приложения

1. Простота развертывания

- Использует Flask - легковесный фреймворк для веб-приложений

- Минимальные зависимости

2. Защита от ошибок

- Проверка существования файла модели при запуске

- Обработка исключений при вводе данных и предсказании

3. Масштабируемость

- Легко можно добавить новые модели или функционал

- Простая интеграция с другими сервисами

4. Интерфейс

- Использует шаблонизатор Flask (render\_template) для генерации HTML

- Передает в шаблон результат предсказания или сообщение об ошибке

Использование

1. Пользователь открывает главную страницу приложения

2. Вводит необходимые параметры в форму

3. Отправляет данные (POST-запрос)

4. Получает результат предсказания или сообщение об ошибке

Приложение предназначено для задач, где требуется прогнозирование на основе множества числовых параметров, однако, в силу отсутствии данного материала по теме и освоения оного самостоятельно, может работать некорректно.

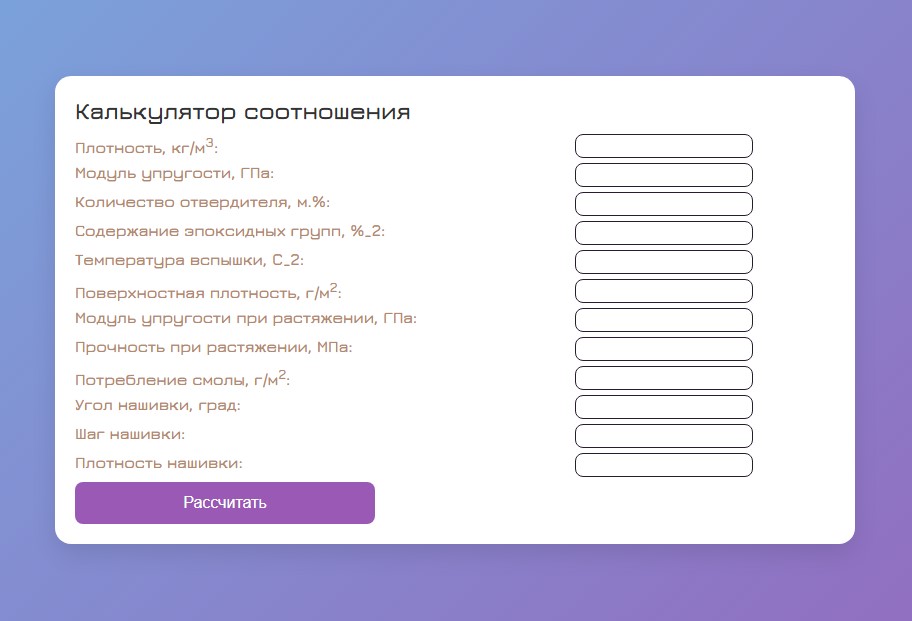


Рис. 28 – приложение.

**2.5 Создание удаленного репозитория и загрузка**

Репозиторий был создан на github.com по адресу: https://github.com/Karbofos-13/diploma.git

**Заключение**

В ходе исследования были выявлены определенные ограничения, в связи с чем был реализован ряд мер по повышению качества данных, включая обработку выбросов и создание дополнительных производных признаков. Однако отсутствие возможности привлечения экспертов в предметной области не позволило добиться значительного улучшения результатов. При этом важно отметить, что достигнутые метрики качества как нейронной сети, так и других моделей подтверждают принципиальную возможность построения точных прогнозных систем для решения поставленной задачи. Таким образом, для дальнейшего развития исследования необходим экспертный анализ полученных результатов с привлечением специалистов в области материаловедения.

**2.7 Список используемой литературы**

1. Араки Масахиро: \*Занимательное машинное обучение\*. ДМК Пресс, 2020. 2. В. В. Кудинов, И. К. Крылов: \*Влияние компонентов на свойства полимерных материалов\*. Монография-справочник. Изд. Наука, 2021.

3. В. И. Костиков: \*Технология композиционных материалов\*. Учебное пособие. Инфра-Инженерия, 2021.

4. И. Ф. Кобылкин, В. В. Селиванов: \*Материалы и структуры легкой бронезащиты\*. Изд. МГТУ им. Н. Э. Баумана, 2014.

5. Луис Серрано: \*Грокаем машинное обучение\*. Изд. Питер, 2025.

6. Пратхап Дангети: Статистика для машинного обучения (Pratap Dangeti: Statistics for Machine Learning). Packt Publishing, 2017.

7. Питер Брюс, Эндрю Брюс: Практическая статистика для специалистов по данным (Peter Bruce, Andrew Bruce: Practical Statistics for Data Scientists). O’Reilly Media, 2017.

8. Себастьян Рашка, Вахид Мирджалили: Машинное обучение на Python (Sebastian Raschka, Vahid Mirjalili: \*Python Machine Learning\*). Packt Publishing, 2019.

9. Стефани Молин: Анализ данных с Pandas на практике (Stefanie Molin: Hands-on Data Analysis with Pandas). Packt Publishing, 2019.

10. Т. А. Гузева, Г. Е. Нехороших, А. И. Долгих: Определение физико-механических характеристик полимерных композиционных материалов. Учебно-методическое пособие. Изд. МГТУ им. Н. Э. Баумана, 2020.

11. Томас Нильд: Основы математики для Data Science (Thomas Nield: Essential Math for Data Science). O’Reilly Media, 2022.

12. Уэс МакКинни: Python для анализа данных (Wes McKinney: Python for Data Analysis). O’Reilly Media, 2022.

13. Франсуа Шолле: Глубокое обучение на Python. Изд. Питер, 2025.

14. Чарльз Куфс: Статистика с котиками (Charles Kufs: Stats with Cats). Wheatmark, 2011.

15. Шарапова В. А.: Композиционные материалы специального назначения. Учебное пособие. Уральский федеральный университет, 2020.

16. Эндрю Траск: Грокаем глубокое обучение. Изд. Питер, 2025.

17. Ли Вон: Инструменты Python для ученых (Lee Vaughan: Python Tools for Scientists). No Starch Press, 2023.

18. 25+ Most Used Keras Snippets in 2025: – Режим доступа: [https://www.geeksforgeeks.org/25-most-used-keras-snippets-in-2025/](https://www.geeksforgeeks.org/25-most-used-keras-snippets-in-2025/) (дата обращения 01.07.2025).

19. Библиотека Matplotlib: – Режим доступа: [https://devpractice.ru/files/books/python/Matplotlib.book.pdf](https://devpractice.ru/files/books/python/Matplotlib.book.pdf) (дата обращения 15.06.2025).

20. Enhancing Multi-Layer Perceptron Performance: Demystifying Optimizers: – Режим доступа: [https://towardsai.net/p/l/enhancing-multi-layer-perceptron-performance-demystifying-optimizers](https://towardsai.net/p/l/enhancing-multi-layer-perceptron-performance-demystifying-optimizers) (дата обращения 02.07.2025).

21. Introduction to the Keras Tuner: – Режим доступа: [https://www.tensorflow.org/tutorials/keras/keras\_tuner](https://www.tensorflow.org/tutorials/keras/keras\_tuner) (дата обращения 01.07.2025).

22. Multi-Layer Perceptrons: – Режим доступа: [https://towardsdatascience.com/multi-layer-perceptrons-8d76972afa2b/](https://towardsdatascience.com/multi-layer-perceptrons-8d76972afa2b/) (дата обращения 04.07.2025).

23. Simple NN with Python: Multi-Layer Perceptron: – Режим доступа: [https://www.kaggle.com/code/androbomb/simple-nn-with-python-multi-layer-perceptron](https://www.kaggle.com/code/androbomb/simple-nn-with-python-multi-layer-perceptron) (дата обращения 05.07.2025).

24. Машинное обучение и Data Science: Совершенствуем AI-модели с помощью AdaBoost: – Режим доступа: [https://www.mql5.com/ru/articles/14034](https://www.mql5.com/ru/articles/14034) (дата обращения 21.06.2025).