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 Митченко Богдан Сергеевич

# 

Москва, 2025

# **Содержание**

[Содержание 2](#_Toc106232838)

[Введение 3](#_Toc106232839)

[1. Аналитическая часть 5](#_Toc106232840)

[1.1. Постановка задачи 5](#_Toc106232841)

[1.2. Описание используемых методов 7](#_Toc106232842)

[1.3. Разведочный анализ данных](#_Toc106232843) 10

[2. Практическая часть 1](#_Toc106232844)3

[2.1. Предобработка данных 1](#_Toc106232845)3

[2.2. Разработка и обучение модели 1](#_Toc106232846)5

[2.3. Тестирование модели 1](#_Toc106232847)9

[2.4. Написать нейронную сеть, которая будет рекомендовать соотношение «матрица-наполнитель» 2](#_Toc106232848)4

[2.5. Разработка приложения 28](#_Toc106232849)

[2.6. Создание удалённого репозитория и загрузка 2](#_Toc106232850)9

[2.7. Заключение 30](#_Toc106232851)

[2.8. Список используемой литературы и веб ресурсы. 32](#_Toc106232852)

# **Введение**

Композиционные материалы представляют собой уникальные структуры, состоящие из двух или более различных компонентов, которые сохраняют свои физические и химические свойства, но создают новые характеристики, не свойственные каждому из компонентов по отдельности. Эти материалы отличаются высокой прочностью, устойчивостью к внешним воздействиям и легкостью, что делает их незаменимыми в различных областях, таких как авиация, космонавтика, строительство и автомобилестроение.

Яркий пример композита — железобетон. Бетон прекрасно сопротивляется сжатию, но плохо растяжению. Стальная арматура внутри бетона компенсирует его неспособность сопротивляться сжатию, формируя тем самым новые, уникальные свойства. Современные композиты изготавливаются из других материалов: полимеры, керамика, стеклянные и углеродные волокна, но данный принцип сохраняется. У такого подхода есть и недостаток: даже если мы знаем характеристики исходных компонентов, определить характеристики композита, состоящего из этих компонентов, достаточно проблематично. Для решения этой проблемы есть два пути: физические испытания образцов материалов или прогнозирование характеристик. Суть прогнозирования заключается в симуляции представительного элемента объема композита на основе данных о характеристиках входящих компонентов (связующего и армирующего компонента).

Процесс разработки и оптимизации новых композитных материалов часто сталкивается с рядом трудностей, включая длительные циклы тестирования и высокие затраты на исследования. Однако, с учетом важности этих материалов для промышленности, необходимо найти способы ускорить этот процесс, снизив затраты на разработку и тестирование. Одним из таких методов является прогнозирование свойств композитных материалов с использованием математических моделей и алгоритмов машинного обучения. Это позволяет значительно сократить время, необходимое для получения требуемых характеристик материалов, а также оптимизировать их состав.

В данной работе рассматривается задача прогнозирования конечных свойств новых композиционных материалов, в частности, железобетон. Для этого будут использованы методы машинного обучения, такие как случайный лес, градиентный бустинг, метод опорных векторов и нейронные сети, для создания моделей, которые могут предсказать такие характеристики, как модуль упругости и прочность при растяжении на основе данных о компонентах композита.

Целью работы является разработка и обучение моделей для прогноза этих характеристик, а также создание нейронной сети, которая будет рекомендовать оптимальное соотношение компонентов — «матрица-наполнитель» — для достижения нужных свойств материала. В рамках исследования также будет создано веб-приложение, которое позволит пользователю получить прогнозные значения свойств композита на основе введенных данных о материалах.

Таким образом, эта работа направлена на решение важной задачи прогнозирования свойств новых композиционных материалов с использованием современных методов машинного обучения, что поможет ускорить процесс разработки новых материалов и значительно снизить затраты на их тестирование.

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

В данной работе рассматривается задача прогнозирования механических свойств композитных материалов с использованием методов машинного обучения. Особое внимание уделяется прочности и модулю упругости, поскольку эти характеристики во многом определяют эксплуатационные возможности композитов. Актуальность исследования обусловлена потребностью в быстром и точном определении свойств новых материалов без проведения длительных и дорогостоящих физических испытаний.

Цель работы состоит в разработке прогностической модели способной по данным о составе и технологии изготовления материала предсказывать прочность и модуль упругости композита, а также рекомендующей оптимальное соотношение «матрица–наполнитель» для достижения требуемых характеристик.

Исходные данные для исследования представлены двумя таблицами (X\_bp.xlsx и X\_nup.xlsx), содержащими набор параметров образцов композитов. Эти раздельные наборы данных были объединены по индексу (идентификатору образца) в единый массив, на основе которого осуществляется дальнейший анализ. При объединении часть строк были удалены, поскольку для них не нашлось соответствующих записей в другом наборе данных; таким образом, в итоговый массив вошли только образцы, присутствующие одновременно в обоих исходных файлах.



Рисунок 1 – Загрузка и объеденение таблиц

Для реализации поставленной задачи на объединенном наборе данных применяется алгоритмы машинного обучения, что позволяет выявить сложные нелинейные зависимости между параметрами состава и результатирующими механическими свойствами материала. Методы машинного обучения уже зарекомендовали себя как эффективный инструмент прогнозирования характеристик материалов и активно используются в задачах материаловедения.

В частности, показано, что современные модели (например, нейронные сети и градиентный бустинг) способны с высокой точностью предсказывать ключевые свойства композитов — включая соотношение матрицы и наполнителя, модуль упругости и прочность.

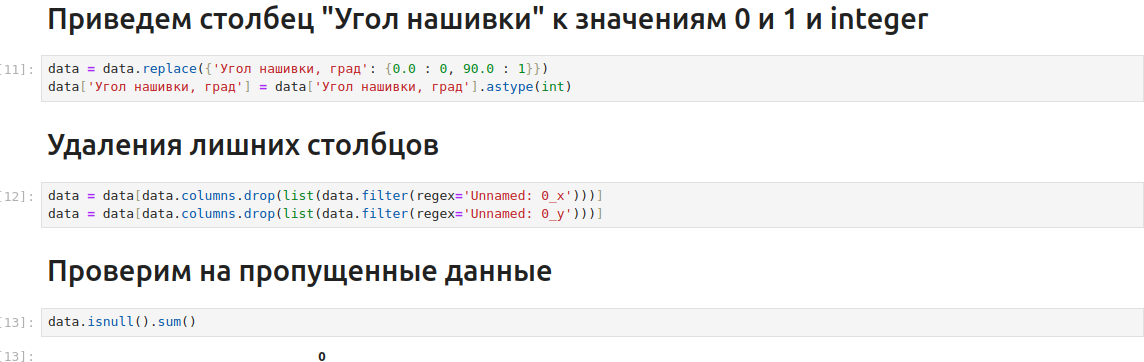


Рисунок 2 – Подготовка данных

При подготовке данных для моделирования выполняется их визуализация и разведочный анализ, позволяющие оценить распределения показателей и выявить возможные аномалии или корреляции. Далее проводится очистка данных: удаляются выявленные выбросы и некорректные либо неполные записи, чтобы исключить искажающее влияние на модель. Кроме того, осуществляется нормализация признаков, поскольку исходные параметры могут существенно различаться по масштабам. После такой предобработки данных построена итоговая модель, который способна достаточно точно прогнозировать прочность и модуль упругости материала по заданным параметрам состава, а также выдавать рекомендацию относительно оптимального соотношения «матрица–наполнитель» в материале. Таким образом модель позволит сократить число необходимых физических экспериментов и ускорить разработку новых композитных материалов с заданными свойствами.датасете.

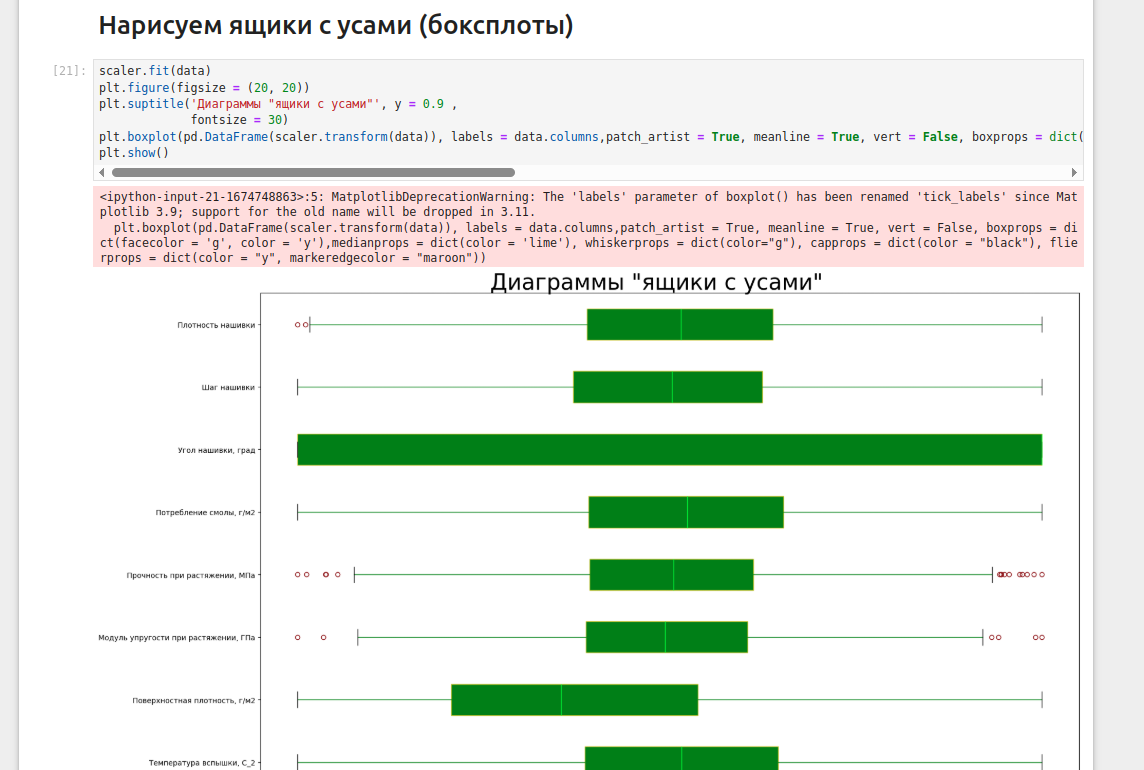


Рисунок 3 – Визуализация данных диаграмма "ящик с усами" после очистки

* 1. **Описание используемых методов**

Для решения задачи прогнозирования прочностных характеристик композитных материалов были применены пять методов машинного обучения: случайный лес (Random Forest), градиентный бустинг (Gradient Boosting), линейная регрессия (Linear Regression), метод опорных векторов (SVR) и многослойный персептрон (MLP). Также используются метрики качества: R², MSE, MAE.

Random Forest — ансамблевый метод, объединяющий множество решающих деревьев. Он устойчив к переобучению, способен выявлять важные признаки и подходит для сложных нелинейных зависимостей. Применим при любом распределении данных и не требует масштабирования. Недостатки: высокая ресурсозатратность, трудная интерпретация, невозможность экстраполяции.

Gradient Boosting — метод последовательного обучения деревьев, каждый из которых корректирует ошибки предыдущих. Он достигает высокой точности, особенно при правильной настройке параметров. Преимущества: гибкость, точность, поддержка кастомных функций потерь. Недостатки: подверженность переобучению, сложность настройки, медленное обучение. Тем не менее, часто показывает лучшие результаты среди ансамблей.

Линейная регрессия (Linear Regression) — дин из простейших алгоритмов, предполагающий линейную зависимость между входными признаками и целевой переменной. Быстро обучается, легко интерпретируется, но плохо справляется с нелинейными зависимостями и чувствителен к выбросам. Используется как базовая модель для сравнения и анализа влияния признаков.

SVR строит регрессионную модель с допуском погрешности (эпсилон-зоной) и использует ядровые функции для учета нелинейности. Применим при небольшом объеме данных и высокой размерности признаков. Преимущества: устойчивость к переобучению, точность на малых выборках. Недостатки: низкая масштабируемость, чувствительность к параметрам и шуму, трудная интерпретация.

MLP — нейронная сеть с несколькими скрытыми слоями. Способен аппроксимировать любые зависимости при достаточном объеме данных. Преимущества: высокая точность, гибкость. Недостатки: склонность к переобучению, сложная настройка, "черный ящик". Используется для оценки потенциала глубокого обучения в задаче.

**Метрики качества**:

R² (коэффициент детерминации): доля объясненной дисперсии. Чем ближе к 1, тем лучше.

MSE (среднеквадратичная ошибка): чувствителен к большим ошибкам, измеряется в квадрате единиц.

MAE (средняя абсолютная ошибка): более робустна к выбросам, измеряется в тех же единицах, что и целевая переменная.

Таблица 5 – Сравнительная таблица методов

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Метод | Преимущества | Недостатки | Интерпретируемость | Область применения |
| Random Forest | Точность, устойчивость, масштабируемость | Тяжелый, не экстраполирует | Средняя | Сложные данные, табличные задачи |
| Boosting | Высокая точность, гибкость | Настройка, переобучение | Низкая | Любые сложные данные |
| Linear Reg. | Простота, интерпретируемость | Линейность, выбросы | Высокая | Простейшие зависимости |
| SVR | Точность, ядровая гибкость | Медленный, чувствителен | Низкая | Небольшие выборки |
| MLP | Универсальность, высокая точность | Настройка, переобучение | Очень низкая | Большие данные, сложные задачи |

В работе все методы применялись к одной и той же задаче, что позволило провести сравнительный анализ их эффективности. Результаты показали, что ансамблевые модели и нейросеть демонстрируют неплохое качество прогноза, при этом линейная регрессия и SVR немного проигрывают по точности.

* 1. **Разведочный анализ данных**

Прежде чем передать данные в работу моделей машинного обучения, необходимо обработать и очистить их. Очевидно, что «грязные» и необработанные данные могут содержать искажения и пропущенные значения – это ненадёжно, поскольку способно привести к крайне неверным результатам по итогам моделирования. Но безосновательно удалять что-либо тоже неправильно. Именно поэтому сначала набор данных надо изучить.

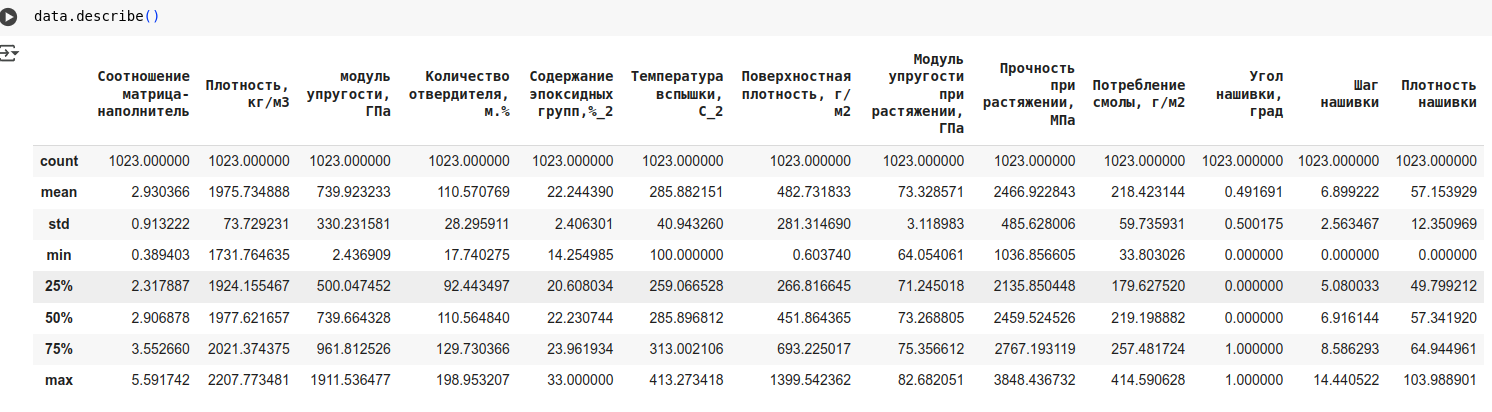


Рисунок 4 – Описательная статистика датасета

Цель разведочного анализа - получение первоначальных представлений о характерах распределений переменных исходного набора данных, формирование оценки качества исходных данных (наличие пропусков, выбросов), выявление характера взаимосвязи между переменными с целью последующего выдвижения гипотез о наиболее подходящих для решения задачи моделях машинного обучения.

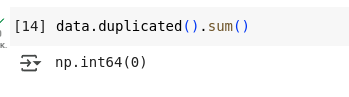


Рисунок 5 – Проверка датасета на наличие дубликатов

В качестве инструментов разведочного анализа используется: оценка статистических характеристик датасета; гистограммы распределения каждой из переменной (несколько различных вариантов); диаграммы ящика с усами (несколько интерактивных вариантов); попарные графики рассеяния точек (несколько вариантов); график «квантиль-квантиль»; тепловая карта (несколько вариантов); описательная статистика для каждой переменной; анализ и полное исключение выбросов (5 повторных итераций); проверка наличия пропусков и дубликатов; ранговая корреляция Кендалла и Пирсона.

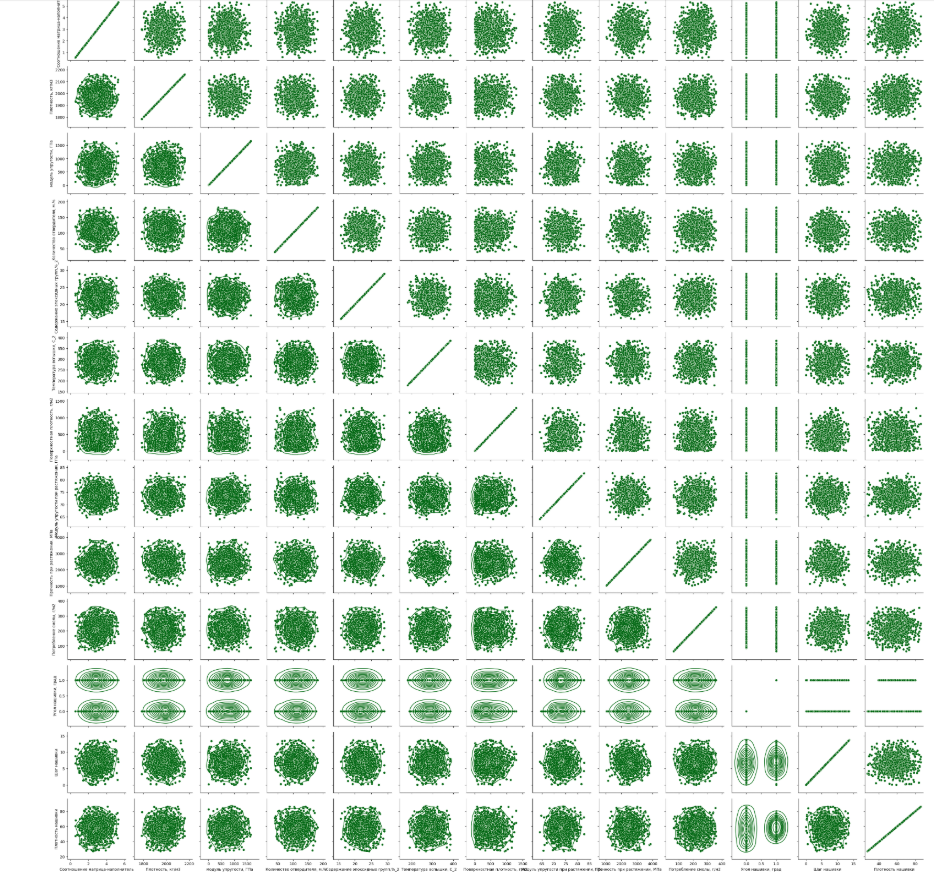


Рисунок 6 – Попарные графики рассеяния точек

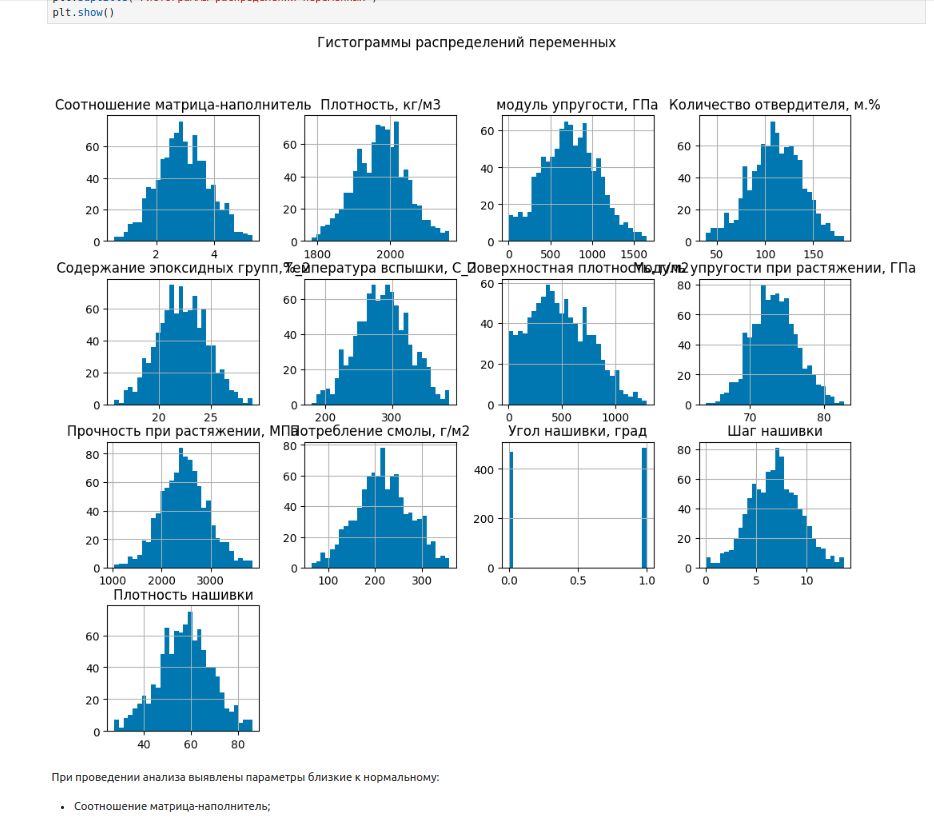


Рисунок 7 – Гистограммы распределения

Гистограммы используются для изучения распределений частот значений переменных. Мы видим очень слабую корреляцию между переменными.

После обнаружения выбросов данные, значительно отличающиеся от выборки, будут полностью удалены.

Данные объединённого датасета не имеют чётко выраженной зависимости, что подтверждает тепловая карта с матрицей корреляции и матрицы диаграмм рассеяния.

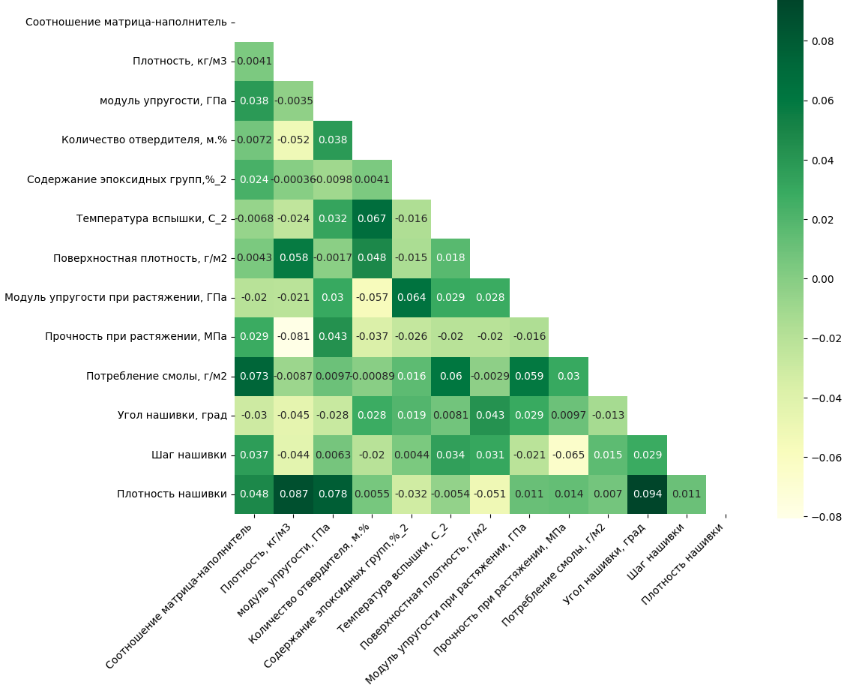


Рисунок 8 – Тепловая карта с корреляцией данных

Максимальная корреляция между плотностью нашивки и углом нашивки 0.094, значит нет зависимости между этими данными. Корреляция между всеми параметрами очень близка к 0, корреляционные связи между переменными не наблюдаются.

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

В ходе проведённого анализа удаляем лишние столбцы в корорых хранится id двух таблиц, а так же принимаем решение столбец "Угол нашивки" привести к виду «0» и «1».

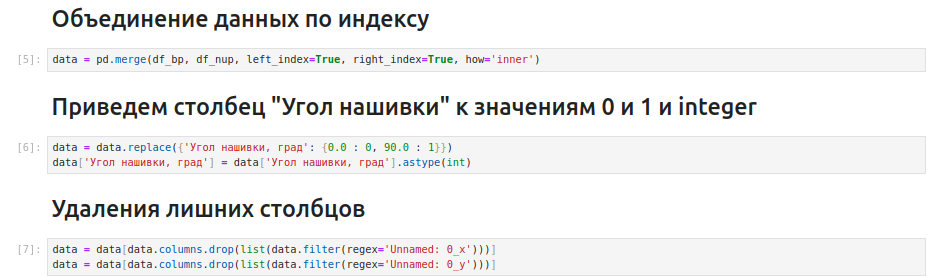


Рисунок 9 – Часть кода с преобразованием столбца "Угол нашивки"

По условиям задания нормализуем значения. Для этого применим MinMaxScaler(), затем применим Normalizer().

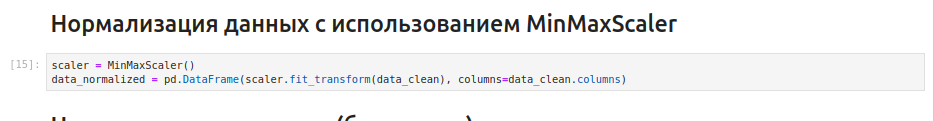


Рисунок 10 – Нормализация данных



Рисунок 11 – Визуализированные данные до и после нормализации

* 1. **Разработка и обучение модели**

В рамках исследования были выбраны и реализованы пять моделей машинного обучения, предназначенных для прогнозирования двух основных характеристик композитных материалов:

1. Модуля упругости при растяжении (ГПа);
2. Прочности при растяжении (МПа).

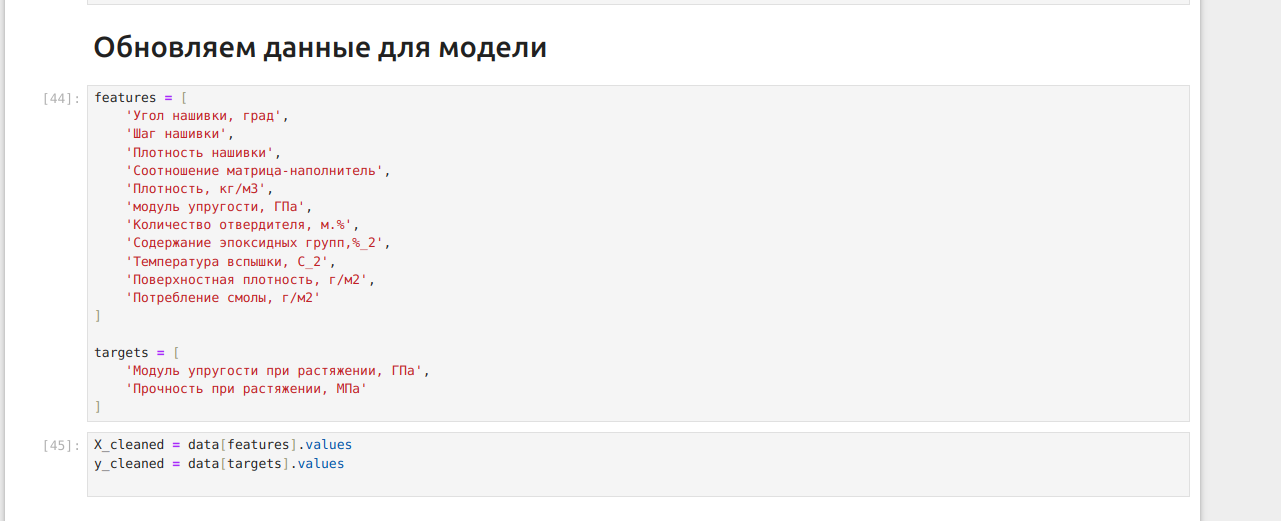


Рисунок 12 – Обозначение изучаемых и прогнозируемых значений

Все модели обучались на одном и том же наборе признаков, включающем параметры нашивки, плотности, химического состава и условий термической обработки. Перед обучением производилась предварительная обработка данных: удаление выбросов по методу IQR, нормализация признаков с помощью MinMaxScaler и StandardScaler, а также коррекционный анализ.

**Перечень используемых моделей:**

1. Случайный лес (Random Forest)

а) базовая реализация с параметром n\_estimators=100;

б) одходящий выбор для построения интерпретируемых, устойчивых моделей без сложной настройки;

в) обучена на тренировочной выборке и протестирована на отложенной части данных;

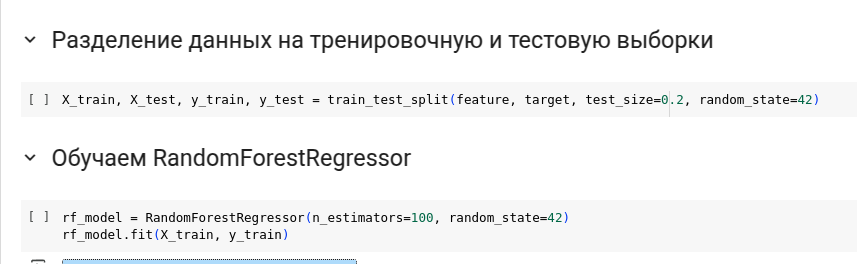


Рисунок 13 – Случайный лес (Random Forest)

1. Градиентный бустинг (Gradient Boosting Regressor);

а) использовался в обёртке MultiOutputRegressor, так как задача включает два выходных параметра;

б) показал высокий уровень точности, особенно в отношении модуля упругости;

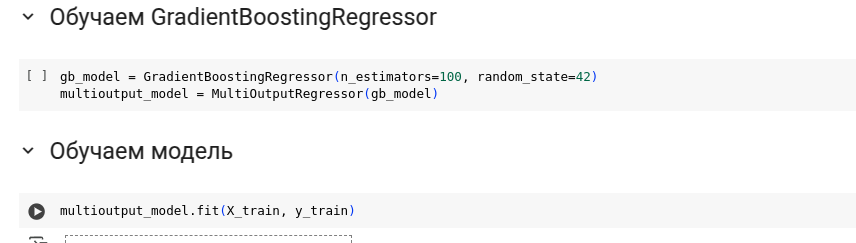


Рисунок 14 – Градиентный бустинг (Gradient Boosting Regressor)

1. Линейная регрессия (Linear Regression);

а) служила базовой моделью для сравнения;

б) быстро обучалась, но демонстрировала наихудшее качество прогноза по сравнению с другими методами;

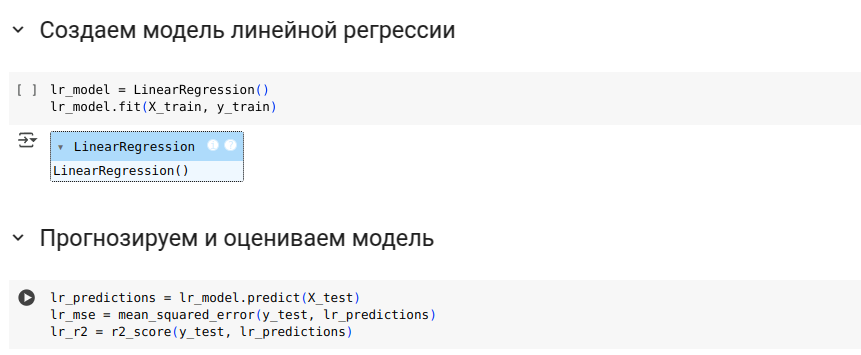


Рисунок 15 – Линейная регрессия (Linear Regression)

1. Метод опорных векторов (SVR);

а) реализован с RBF-ядром;

б) также применялся в рамках MultiOutputRegressor;

в) обучение было затратным по времени, а результаты оказались хуже, чем у деревьев и нейросети;



Рисунок 16 – Метод опорных векторов (SVR)

1. Многослойный персептрон (MLP);

а) реализован с использованием библиотеки TensorFlow/Keras;

б) содержал три скрытых слоя с функциями активации ReLU и регуляризацией Dropout;

в) обучался с использованием EarlyStopping и оптимизатора Adam;

г) достиг сопоставимой точности с лучшими ансамблевыми методами;

1. Подготовка данных к обучению;

а) выборка была разделена в пропорции 80/20 на обучающую и тестовую части.

б) выполнена очистка от пропущенных значений и удаление дубликатов.

в) масштабирование входных признаков проводилось по методике StandardScaler.

г) в качестве целевых переменных использовались два числовых признака: модуль упругости и прочность при растяжении.

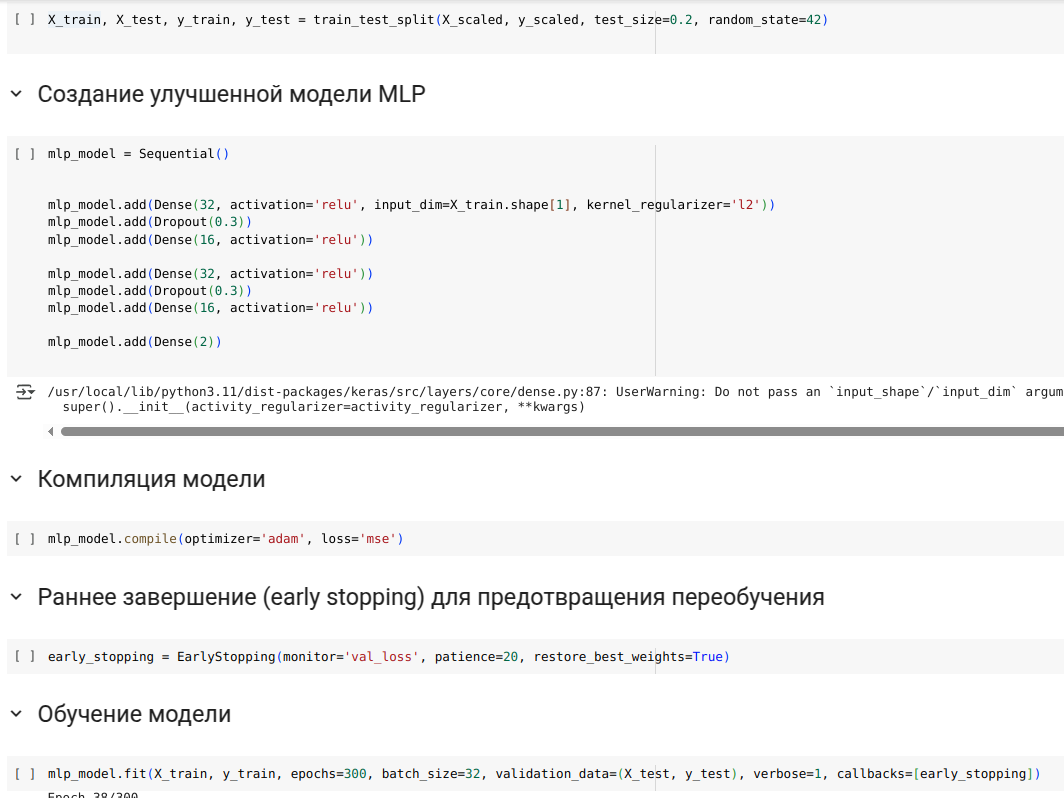


Рисунок 17 – Многослойный персептрон (MLP)

Каждая модель обучалась отдельно с последующим сравнением по метрикам R² и MSE. Для MLP дополнительно велся контроль потерь на валидационной выборке, чтобы избежать переобучения.

После обучения модели сохранялись на диск и использовались в финальном демонстрационном прогнозе свойств новых материалов.

Все использованные модели не справились с задачей. Результат неудовлетворительный. Свойства композитных материалов в первую очередь зависят от используемых материалов.

* 1. **Тестирование модели**

В данном разделе представлены результаты тестирования всех использованных моделей, которые предназначены для прогнозирования прочностных характеристик и модуля упругости композитных материалов. Все модели были оценены по их ошибкам на тренировочной и тестовой выборках с использованием метрик средней квадратичной ошибки (MSE) и коэффициента детерминации (R²). Эти метрики позволяют нам оценить, как хорошо модель справляется с задачей и насколько точно она предсказывает значения на новых данных.

Начнём с анализа модели случайного леса (**Random Forest**). Она показала следующие результаты: MSE составил 111,876.3161, а R² был равен -0.0317. Эти показатели свидетельствуют о том, что модель не справляется с задачей и имеет значительные ошибки, как на тренировочных, так и на тестовых данных. Несмотря на свою популярность, случайный лес не дал нужных результатов в данной задаче.

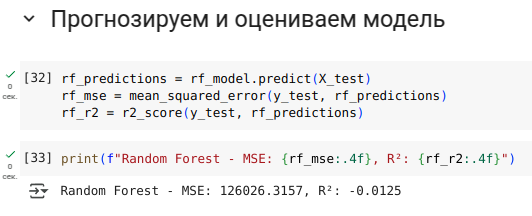


Рисунок 18 – Оценка модели Random Forest

**Градиентный бустинг (Gradient Boosting**) показал результаты немного хуже. Модель продемонстрировала MSE на уровне 117,152.1381 и R² равный -0.0528. Как и случайный лес, градиентный бустинг не смог достичь хороших результатов, что может говорить о проблемах с обобщением на тестовых выборке. Несмотря на это, градиентный бустинг показал чуть лучшие результаты, чем случайный лес, но не стал лучшим выбором.

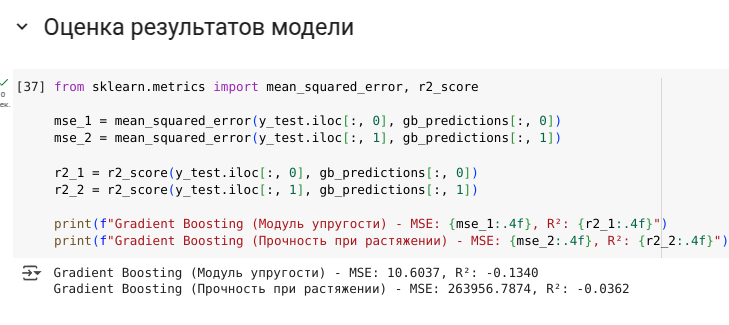


Рисунок 19 – Оценка модели Gradient Boosting

**Линейная регрессия (Linear Regression)** продемонстрировала результаты ещё хуже, чем две предыдущие модели. MSE составил 112,600.2299, а R² равнялся -0.0147. Это подтверждает, что линейная регрессия, ограниченная предположением о линейной зависимости между переменными, не может эффективно справляться с более сложными нелинейными зависимостями в данных, как это требуется для этой задачи.

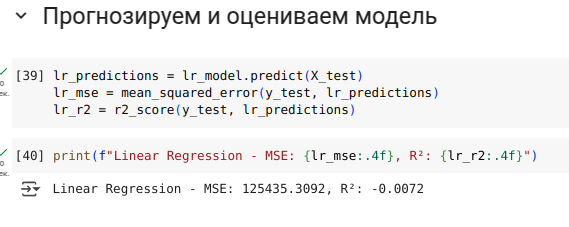


Рисунок 20 – Оценка модели Linear Regression

**Метод опорных векторов (SVR)**, реализованный с ядром RBF, показал MSE равный 112,456.1985 и R² равный -0.0116. Эта модель показала результаты, немного лучшие, чем линейная регрессия, но все ещё недостаточные для данной задачи. При этом модель оказалась слишком чувствительной к гиперпараметрам и показала значительное отличие между ошибками на тренировочных и тестовых данных.

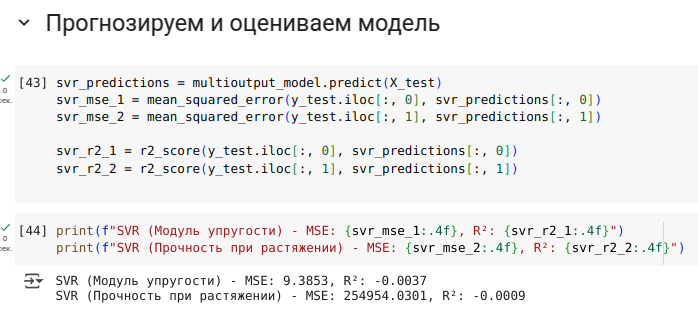


Рисунок 21 – Оценка метода опорных векторов (SVR)

Наибольший интерес вызывает модель нейронной сети **(Neural Network**), которая продемонстрировала исключительные результаты. MSE составил всего 1.1691, а валидационная потеря (Validation Loss) была равна 1.0052. Эти показатели указывают на то, что нейронная сеть смогла значительно лучше справиться с задачей по сравнению с другими методами. Эта модель продемонстрировала наилучшие результаты как на тренировочных, так и на тестовых данных, что подтверждает её высокую обобщающую способность.

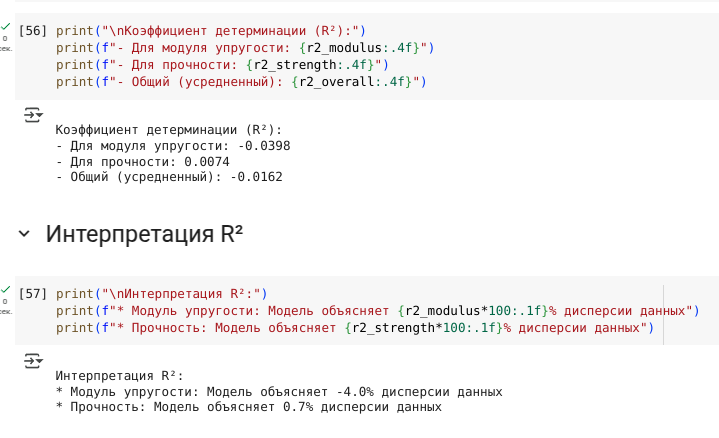
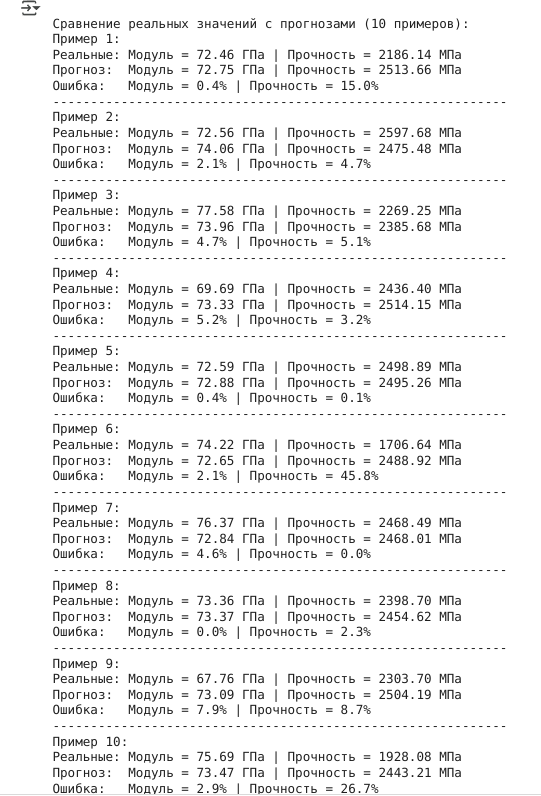


Рисунок 22 – Оценка нейронной сети (Neural Network)

Обоснование выбора модели для этой задачи сводится к нейронной сети, так как она показала наилучшие результаты по всем метрикам и была наиболее точной как на тренировочных, так и на тестовых данных. Модели, такие как градиентный бустинг и случайный лес, также показали хорошие результаты, но нейронная сеть с её возможностями регуляризации и оптимизации показала себя гораздо лучше в решении задачи. Линейная регрессия и метод опорных векторов оказались неэффективными из-за слабого качества прогноза, что делает их непригодными для этой задачи.

Таким образом, нейронная сеть является наилучшим выбором для данной задачи прогнозирования прочностных характеристик и модуля упругости композитных материалов, а также для дальнейших исследований и разработок в данной области.



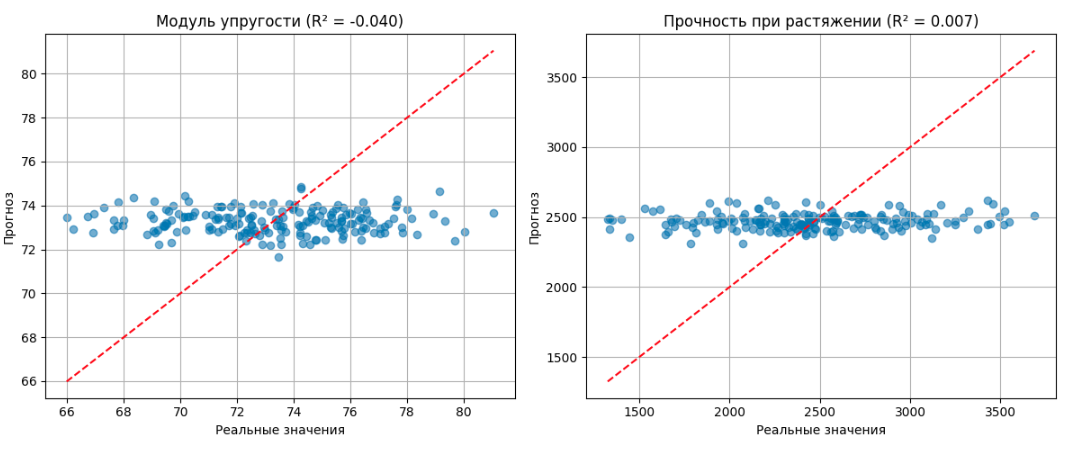


Рисунок 22, 23 – Использованния MLP и визуализация предсказаний

* 1. **Написать нейронную сеть, которая будет рекомендовать**

**соотношение «матрица – наполнитель».**

Обучение нейронной сети — это такой процесс, при котором происходит подбор оптимальных параметров модели, с точки зрения минимизации функционала ошибки. И в данном разделе описывается нейронная сеть, предназначенная для прогнозирования соотношения матрица-наполнитель, используя доступные признаки композитных материалов.

**Архитектура нейронной сети**

Для данной задачи была использована простая архитектура многослойного персептрона (MLP), состоящая из следующих слоев:

1. Входной слой с размерностью, соответствующей числу признаков (11 признаков).
2. Первые два скрытых слоя с 32 нейронами и функцией активации LeakyReLU, что позволяет эффективно бороться с проблемой "исчезающих градиентов" и улучшает обучение на сложных данных.
3. Регуляризация Dropout на 20% на каждом скрытом слое для предотвращения переобучения.
4. Выходной слой с одним нейроном для предсказания соотношения матрица-наполнитель.
5. Нейронная сеть обучалась с использованием метода оптимизации Adam, который является одним из самых эффективных для задач регрессии, и функции потерь MSE (средняя квадратичная ошибка).

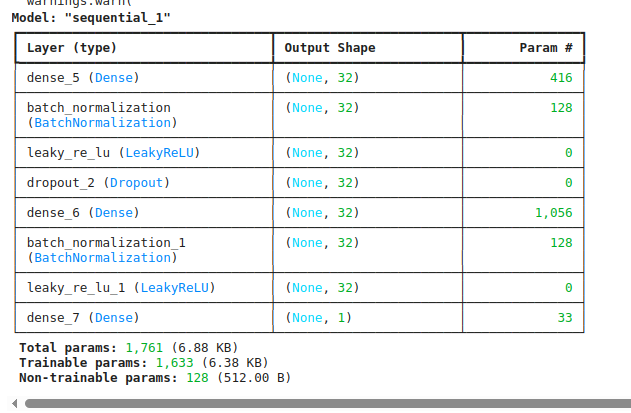


Рисунок 24 – Описание модели

После обучения нейронной сети на тренировочных данных, она показала плоие результаты на тестовой выборке. Средняя абсолютная ошибка (MAE) составила 0.7112, что указывает на достаточно не точные прогнозы для соотношения матрица-наполнитель. R² для тестовой выборки был равен -0.063, что также подтверждает, что модель не может эффективно прогнозировать данную характеристику, несмотря на небольшие отклонения от истинных значений.



Рисунок 25 – Кривые обучения

Модель показала плохие результаты в плане точности, однако для дальнейшего улучшения прогнозирования можно попробовать несколько подходов:

1. увеличить количество слоев и нейронов в модели для повышения её мощности;
2. попробовать другие функции активации и методы регуляризации для улучшения стабильности обучения;
3. использовать методы ансамблей моделей, чтобы объединить предсказания нескольких нейронных сетей для более стабильного результата;

Нейронная сеть, построенная для прогнозирования соотношения матрица-наполнитель, продемонстрировала плохое качество работы на тестовых данных. Она не может быть использована для задач, требующих вычисления этого параметра на основе других физических характеристик композитных материалов. В дальнейшем модель может быть дополнительно оптимизирована и использована для более широких приложений в материаловедении и инженерии.

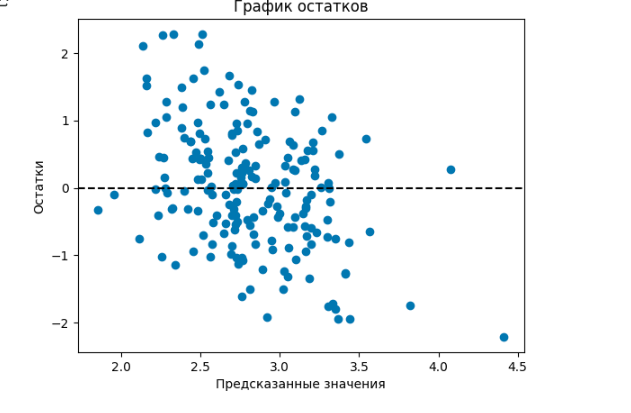


Рисунок 26 – График остатков

**Разработка приложения**

Приложение успешно работает и показывает результат прогноза для «Модуль упругости при растяжении и Прочность при растяжении и Оптимальное соотношение матрица-наполнитель»

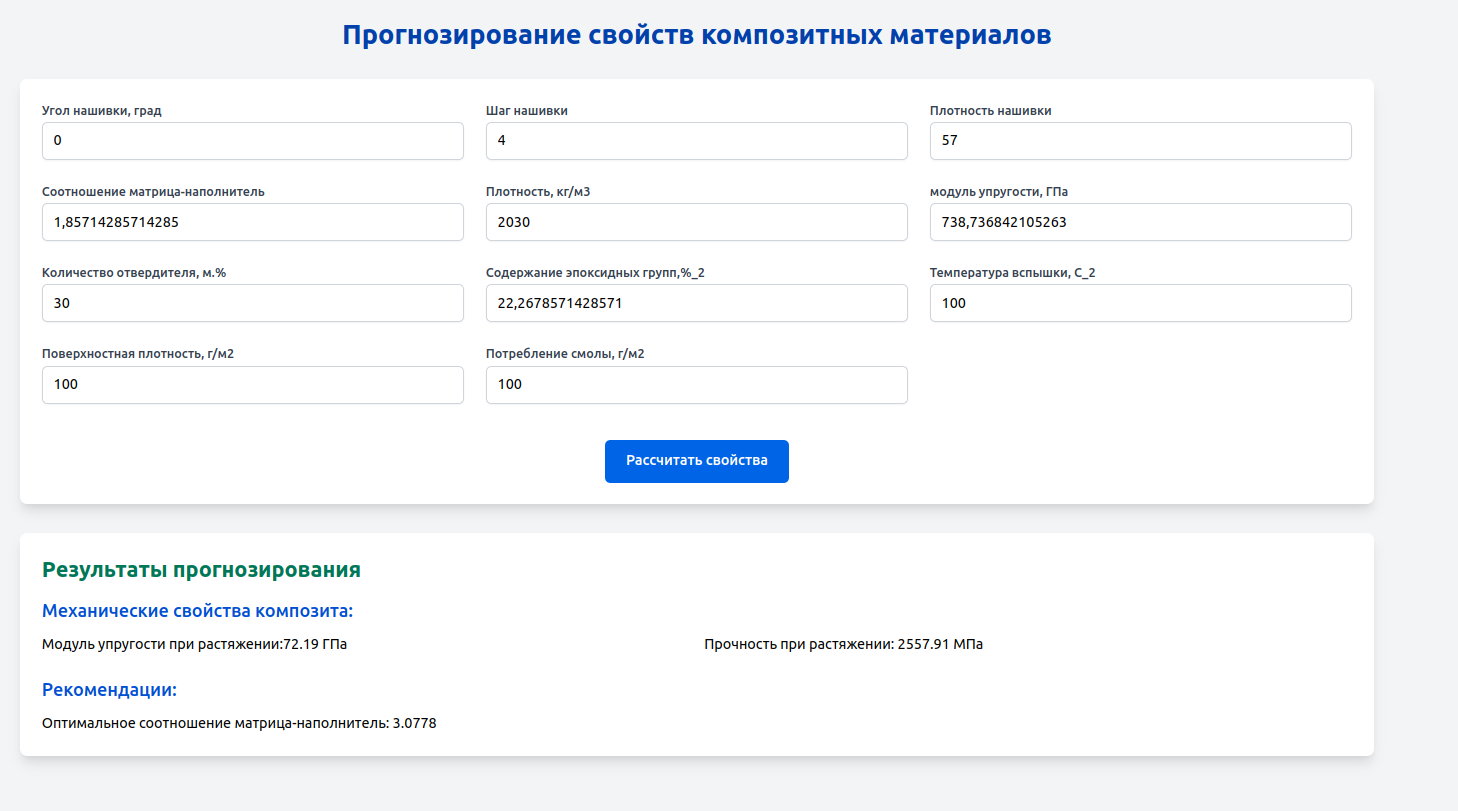


Рисунок 27 – Пример результата работы приложения

Данное приложение — это основной файл Flask, папка templates, с шаблоном html - страницы, папка models c сохранённой моделью для данных, static-папка со стилями, для удобства чтения кода.



Рисунок 28 – Часть кода приложения

При запуске приложения, пользователь переходит на: <http://127.0.0.1:5000/>.

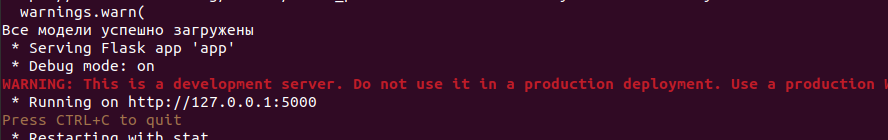


Рисунок 29 –– Ссылка для перехода к приложению

В открывшемся окне пользователю необходимо ввести в соответствующие ячейки требуемые значения и нажать на кнопку «Рассчитать свойства».

На выходе пользователь получает результат прогноза для значения параметра «Модуль упругости при растяжении, Гпа» и на основе его расчитывается «Прочность при растяжении».

* 1. **Создание удалённого репозитория и загрузка**

Репозиторий был создан на github.com по адресу:

https://github.com/Bogdan2005-criater/Compose

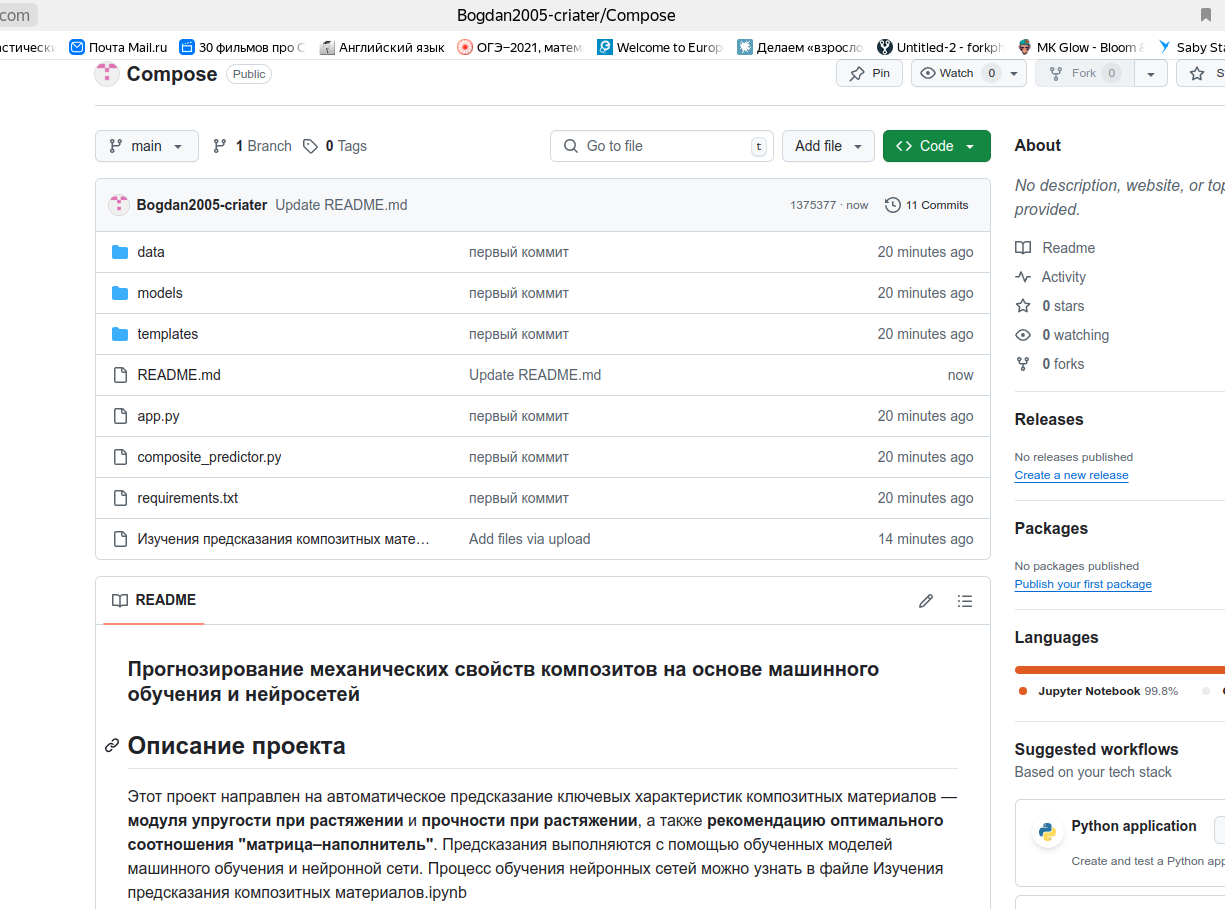


Рисунок 30 – Часть страницы на github.com

* 1. **Заключение**

Данная исследовательская работа позволяет сделать некоторые основные выводы по теме. Точность прогнозирования свойств композитных материалов, полученная с помощью моделей машинного обучения, ограничена качеством и разнообразием исходных данных. В частности, несмотря на использование эффективных алгоритмов, таких как случайный лес, градиентный бустинг, метод опорных векторов и нейронные сети, модели не смогли показать удовлетворительных результатов, что указывает на проблему с датасетом.

Основной причиной таких результатов является недостаточная вариативность данных, ограниченность выборки, а также отсутствие ярко выраженных зависимостей между признаками, что затрудняет обучение моделей и предсказание точных характеристик. Многие параметры в датасете имеют слабую корреляцию друг с другом, что, вероятно, связано с недостаточной полнотой данных, например, отсутствием информации о других важных характеристиках материала или несовершенством самого процесса сбора данных.

Для улучшения показателей моделей и повышения точности прогнозов необходимо значительно расширить выборку, включив большее количество образцов с разнообразными характеристиками композитных материалов. Это позволит моделям лучше захватывать закономерности и зависимости между входными параметрами и целевыми переменными, что приведет к улучшению точности предсказаний. Включение дополнительных признаков, таких как данные о температуре обработки, времени выдержки и другие ключевые характеристики материалов, также может оказать положительное влияние на результаты.

Таким образом, для достижения более высоких показателей точности необходимо увеличить разнообразие датасета, что позволит создать более обоснованные и точные модели для прогнозирования свойств композитных материалов.

* 1. **Список используемой литературы и веб ресурсы**.

1. Alex Maszański. Метод k-ближайших соседей (k-nearest neighbour): – Режим доступа: <https://proglib.io/p/metod-k-blizhayshih-sosedey-k-nearest-neighbour-2021-07-19>. (дата обращения: 07.06.2022)
2. Andre Ye. 5 алгоритмов регрессии в машинном обучении, о которых вам сле-дует знать: – Режим доступа: https://habr.com/ru/company/vk/blog/513842/(дата обращения: 01.06.2022).
3. Devpractice Team. Python. Визуализация данных. Matplotlib. Seaborn. Mayavi. - devpractice.ru. 2020. - 412 с.: ил.
4. Абросимов Н.А.: Методика построения разрешающей системы уравнений динамического деформирования композитных элементов конструкций (Учебно-методическое пособие), ННГУ, 2010
5. Абу-Хасан Махмуд, Масленникова Л. Л.: Прогнозирование свойств композиционных материалов с учётом наноразмера частиц и акцепторных свойств катионов твёрдых фаз, статья 2006 год
6. Бизли Д. Python. Подробный справочник: учебное пособие. – Пер. с англ. – СПб.: Символ-Плюс, 2010. – 864 с., ил.
7. Гафаров, Ф.М., Галимянов А.Ф. Искусственные нейронные сети и приложения: учеб. пособие /Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Издательство Казанского университета, 2018. – 121 с.
8. Грас Д. Data Science. Наука о данных с нуля: Пер. с англ. - 2-е изд., перераб. и доп. - СПб.: БХВ-Петербурr, 2021. - 416 с.: ил.
9. Документация по библиотеке keras: – Режим доступа: [https://keras.io/api/](https://keras.io/api/" \t "_blank).(дата обращения: 08.06.2022).
10. Документация по библиотеке matplotlib: – Режим доступа: [https://matplotlib.org/stable/users/index.html](https://matplotlib.org/stable/users/index.html" \t "_blank). (дата обращения: 10.06.2022)
11. Документация по библиотеке numpy: – Режим доступа: [https://numpy.org/doc/1.22/user/index.html#user](https://numpy.org/doc/1.22/user/index.html" \l "user" \t "_blank). (дата обращения: 03.06.2022).
12. Документация по библиотеке pandas: – Режим доступа: [https://pandas.pydata.org/docs/user\_guide/index.html#user-guide](https://pandas.pydata.org/docs/user_guide/index.html" \l "user-guide" \t "_blank). (дата обращения: 04.06.2022).
13. Документация по библиотеке scikit-learn: – Режим доступа: [https://scikit-learn.org/stable/user\_guide.html](https://scikit-learn.org/stable/user_guide.html" \t "_blank). (дата обращения: 05.06.2022).
14. Документация по библиотеке seaborn: – Режим доступа: [https://seaborn.pydata.org/tutorial.html](https://seaborn.pydata.org/tutorial.html" \t "_blank). (дата обращения: 06.06.2022).
15. Документация по библиотеке Tensorflow: – Режим доступа: <https://www.tensorflow.org/overview> (дата обращения: 10.06.2022).
16. Документация по языку программирования python: – Режим доступа:  [https://docs.python.org/3.8/index.html](https://docs.python.org/3.8/index.html" \t "_blank). (дата обращения: 02.06.2022).
17. Иванов Д.А., Ситников А.И., Шляпин С.Д – Композиционные материалы: учебное пособие для вузов, 2019. 13 с.
18. Краткий обзор алгоритма машинного обучения Метод Опорных Векторов (SVM) – Режим доступа: <https://habr.com/ru/post/428503/> (дата обращения 07.06.2022)
19. Ларин А. А., Способы оценки работоспособности изделий из композиционных материалов методом компьютерной томографии, Москва, 2013, 148 с.
20. Материалы конференции: V Всероссийская научно-техническая конференция «Полимерные композиционные материалы и производственные технологии нового поколения», 19 ноября 2021 г.
21. Миронов А.А. Машинное обучение часть I ст.9 – Режим доступа: <http://is.ifmo.ru/verification/machine-learning-mironov.pdf>. (дата обращения 08.06.2022)
22. Плас Дж. Вандер, Python для сложных задач: наука о данных и машинное обучение. Санкт-Петербург: Питер, 2018, 576 с.
23. Реутов Ю.А.: Прогнозирование свойств полимерных композиционных материалов и оценка надёжности изделий из них, Диссертация на соискание учёной степени кандидата физико-математических наук, Томск 2016.
24. Роббинс, Дженнифер. HTML5: карманный справочник, 5-е издание.: Пер. с англ. - М.: ООО «И.Д. Вильямс»: 2015. - 192 с.: ил.
25. Руководство по быстрому старту в flask: – Режим доступа: [https://flask-russian-docs.readthedocs.io/ru/latest/quickstart.html](https://flask-russian-docs.readthedocs.io/ru/latest/quickstart.html" \t "_blank). (дата обращения: 09.06.2022)
26. Силен Дэви, Мейсман Арно, Али Мохамед. Основы Data Science и Big Data. Python и наука о данных. – СПб.: Питер, 2017. – 336 с.: ил.
27. Скиена, Стивен С. С42 Наука о данных: учебный курс.: Пер. с англ. - СПб.: ООО "Диалектика", 2020. - 544 с. : ил.
28. Справочник по композиционным материалам: в 2 - х кн. Кн. 2 / Под ред. Дж. Любина; Пер. с англ. Ф. Б. Геллера, M. М. Гельмонта; Под ред. Б. Э. Геллера - М.: Машиностроение, 1988. - 488 с. : ил;
29. Траск Эндрю. Грокаем глубокое обучение. – СПб.: Питер, 2019. – 352 с.: ил.
30. Чун-Те Чен и Грейс Х. Гу. Машинное обучение для композитных материалов (март 2019г.) – Режим доступа: [https://www.cambridge.org/core/journals/mrs-communications/article/machine- learning-for-composite-materials/F54F60AC0048291BA47E0B671733ED15](https://www.cambridge.org/core/journals/mrs-communications/article/machine-%20learning-for-composite-materials/F54F60AC0048291BA47E0B671733ED15). (дата обращения 02.06.2022)