**УДК 519.6**

**Применение методов машинного обучения и полносвязных нейронных сетей для прогнозирования свойств композиционных материалов**

**Казаков М.С.**1**,** заместитель начальница цеха – начальник лаборатории, к.т.н

*1 - АО «РКЦ «Прогресс», г. Самара, Российская Федерация*

*E-mail:* [*kazakov.mishel@gmail.com*](mailto:kazakov.mishel@gmail.com)

*Рассмотрены современные методы и алгоритмы машинного обучения, выполнена оценка их эффективности в моделировании зависимостей между исходными параметрами композиционного материала и его механическими свойствами. Описаны этапы подготовки данных включая сбор, очистку и выбор признаков. Показано, что использование машинного обучения имеет существенный потенциал для ускорения процесса разработки новых материалов, оптимизации существующих, а также снижения затрат на экспериментальные исследования.*

*Ключевые слова: машинное обучение, анализ данных, нейронные сети, композиционные материалы*

**APPLICATION OF MACHINE LEARNING METHODS AND FULLY CONNECTED NEURAL NETWORKS TO PREDICT PROPERTIES OF COMPOSITE MATERIALS**

**Kazakov M.S.**1**,** Deputy Head of Workshop - Head of Laboratory, candidate of technical sciences

1 - JSC «RSC «Progress», Samara , Russian Federation

E-mail: [kazakov.mishel@gmail.com](mailto:kazakov.mishel@gmail.com)

*Modern methods and algorithms of machine learning are considered, their efficiency in modeling dependencies between the initial parameters of a composite material and its mechanical properties is assessed. The stages of data preparation, including collection, cleaning and selection of features, are described. It is shown that the use of machine learning has significant potential for accelerating the process of developing new materials, optimizing existing ones, and reducing the costs of experimental research.*

*Key words: machine learning, data analysis, neural networks, composite materials*

Технология композиционных материалов (КМ) позволила решить насущный в авиации и двигателестроении вопрос уменьшения веса и повышения эксплуатационных характеристик. Несмотря на это, внедрение технологии в производство осложняется некоторыми причинами, такими как:

– дороговизна изготовления деталей из КМ, поскольку без промежуточного численного моделирования физический прототип является единственным способом проверить качество и свойства конечного продукта, что возможно уже в самом конце процесса разработки;

– проблема выявления свойств композиционного материала и их дальнейшего использования в моделировании и прогнозировании поведения, исследования прочностных свойств и надежности изделия в целом;

- сложность задач прогнозирования свойств композиционных материалов также связана и с их существенной анизотропией [1].

В связи с этим возникает необходимость в разработке методов и технологий, которые позволят сократить время и средства, необходимые для создания новых материалов. Одним из основных методов является прогнозирование конечных свойств новых материалов [2,3,4]. Это позволяет ускорить процесс исследований, уменьшить затраты на эксперименты и сделать процесс создания новых материалов более эффективным и безопасным.

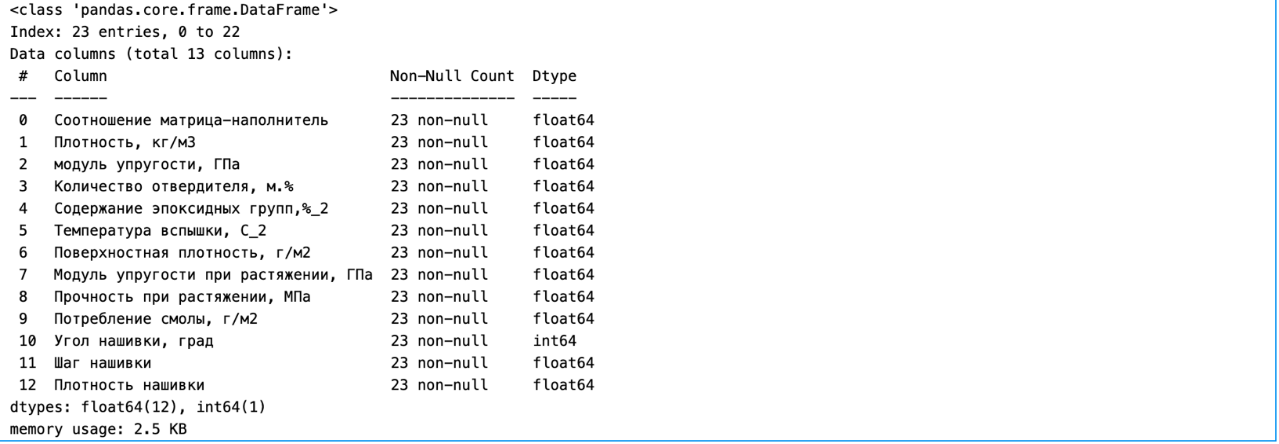
В этом контексте машинное обучение представляет собой мощный инструмент, способный значительно улучшить процессы анализа и прогнозирования. Машинное обучение – это область искусственного интеллекта, занимающаяся разработкой алгоритмов и моделей, которые позволяют компьютерам обучаться на основе данных и делать предсказания или принимать решения без явного программирования. Используя методы машинного обучения можно выявлять сложные зависимости между различными параметрами композитных материалов и их свойствами, что открывает новые возможности их применения и разработки [5-8].

В данной работе исследованы современные подходы к применению машинного обучения для симуляции свойств композиционного материала с матрицей из базальтопластика и нашивками из углепластика на основе данных о характеристиках входящих компонентов, предоставленных Центром НТИ «Цифровое материаловедение: новые материалы и вещества» (МГТУ им. Н.Э, Баумана).

Имеются данные о начальных свойствах компонентов композиционного материала (количество связующего, наполнителя, температурный режим отверждения и т.д.). В результате необходимо спрогнозировать ряд конечных свойств получаемых композиционных материалов.

Представлены основные алгоритмы и методы, а также проанализированы примеры успешного использования машинного обучения. Исследование направлено на то, чтобы продемонстрировать, как машинное обучение может стать ключевым инструментом в создании более эффективных и надежных материалов, способствуя тем самым развитию будущих технологий. Обработка данных и разработка моделей машинного обучения осуществлялись с помощью языка программирования Python с применения библиотек.

Описание признаков набора данных приведено на рисунке 1. Все признаки имеют тип float64 (числа с плавающей точкой), кроме «Угол нашивки, град». «Угол нашивки, град» имеет тип int64 (целочисленные значения). Все признаки, кроме «Угол нашивки, град», являются непрерывными, количественными. «Угол нашивки, град» принимает только два значения. Пропусков в данных нет (Рис. 1).

**

*Рис. 1 — Описание признаков набора данных*

По результатам анализа визуализаций можно сделать вывод, что почти все признаки, кроме «Угол нашивки, град», распределены ненормально и принимают неотрицательные значения. «Угол нашивки, град» принимает значения только – 0 и 90. Дополнительно выполнен анализ описательной статистики данных.

Рассмотрены 2 метода выявления выбросов для имеющихся признаков: метод 3-х сигм и метод межквартильного интервала. Воспользовавшись этими методами было найдено:

* 3 выброса методом 3-х сигм;
* 59 выбросов методом межквартильного интервала расстояний.

Для наиболее эффективной работы моделей воспользуемся методом межквартильного интервала. Значения, определенные как выбросы удаляем. После чего удаляем строки с пропущенными значениями. После очистки от выбросов в наборе данных осталось 11 строк. Выявить связь между признаками возможно при помощи матрицы корреляции. Несмотря на небольшое количество имеющихся данных, для обучения моделей они могут быть использованы в таком количестве, так как присутствует выраженная корреляция между рядом признаков. При этом необходимо отметить, что полученные результаты будут непригодны для промышленного использования без дополнительного исследования с использованием большего массива данных.

Предсказание значений вещественной, непрерывной переменной — это задача регрессии. Эта зависимая переменная должна иметь связь с одной или несколькими независимыми переменными, называемых также предикторами или регрессорами. Регрессионный анализ помогает понять, как «типичное» значение зависимой переменной изменяется при изменении независимых переменных. В настоящей работе рассмотрены следующие методы машинного обучения: линейная регрессия (LinearRegression), Лассо (LASSO) регрессия, случайный лес (RandomForest), метод k-ближайших соседей (KNeighborsRegressor), метод опорных векторов для регрессии (SVR), градиентный бустинг, деревья решений (Decision Trees), нейронная сеть.

С целью повышения качества обучения моделей была выполнена предобработка данных - нормализация, таким образом, чтобы все значения находились в диапазоне от -1 до 1. После использования метода MinMaxScaler, данные были нормализованы (шкалированы/масштабированы), теперь признаки могут быть переданы на вход машинному обучению.

Следующий этап выполнения работы — это разработка и обучение предсказательных моделей. Исходные данные были разделены по целевым признакам, а также на тренировочные и тестовые.

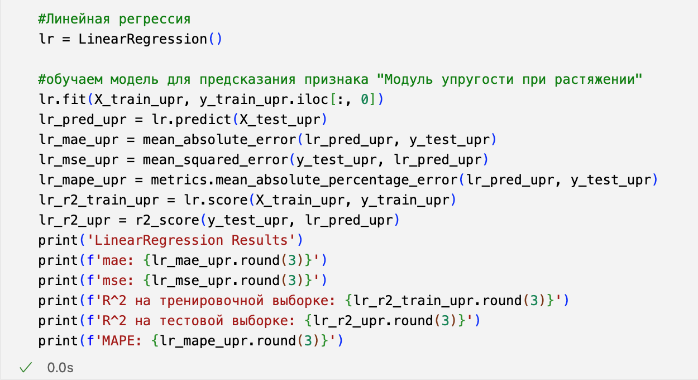
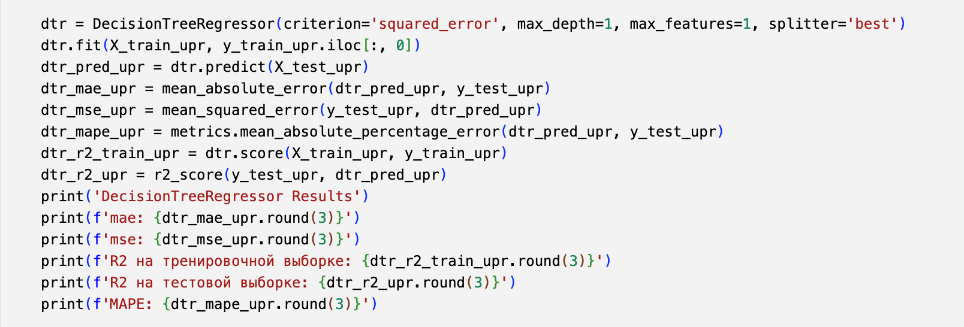
Опробованы восемь методов (включая нейросеть) для прогнозирования двух целевых признаков: «Модуль упругости при растяжении, ГПа» и «Прочность при растяжении, МПа».

Для того чтобы воспользоваться методами обучения моделей, необходимо разделить набор данных на входную часть, в которую входят все данные кроме целевых. Необходимо исключить весь столбец, который будет прогнозироваться и сохранить это множество как входную (X) часть. Затем нужно наоборот, оставить только столбец, в который входят все целевые значения, и сохранить данное множество как выходную (y) часть. Эти манипуляции позволят применить метод train\_test\_split из библиотеки sklearn, который разделит наши части, на тренировочные и тестовые множества. Все вышеописанные действия разделения выборки на входную, выходную, а также, тренировочную и тестовую часть представлены на рис. 2.



*Рис. 2 — Разделение выборки на входное, выходное, тренировочное и тестовое множество*

Некоторые методы машинного обучения является непараметрическими, то есть полностью зависят от данных и не имеют гибких настроек, что не позволяет нам улучшить результаты модели для конкретной задачи путём подбора гиперпараметров. В настоящей работе используются и те, и другие методы, поэтому взглянем на примеры процесса обучения тех и других моделей. На рис. 3а представлен пример обучения модели методом классической линейной регрессии (LinearRegression). Далее, на рис. 3б, представлен пример обучения модели методом дерева решений (DecisionTreeRegressor), уже с подбором гиперпараметров.

**

а) б)

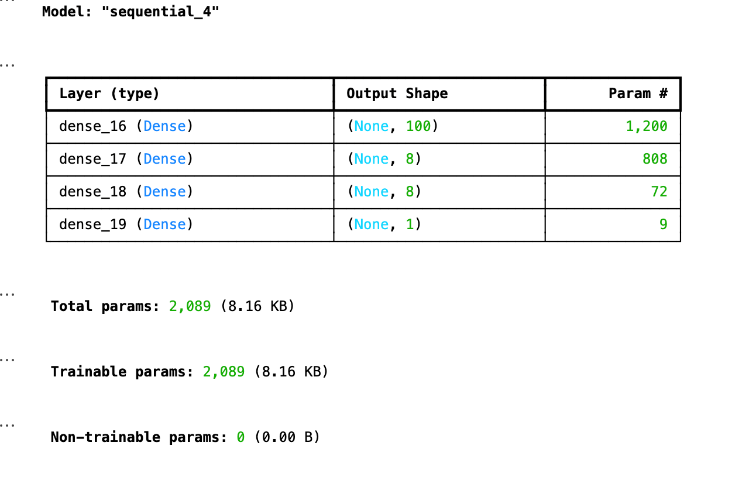
*Рис. 3 — Обучение модели методом линейной регрессии для предсказания признака «Модуль упругости при растяжении, ГПа»*

Для поиска гиперпараметров по сетке использован метод GridSearchCV из библиотеки sklearn. Он получает на вход метод, выбранный нами для решения нашей задачи (регрессор) и набор параметров, для поочерёдной передачи их регрессору, затем выполняет обучение и определяет лучшую комбинацию, что называется гиперпараметрами. Полученные гиперпараметры были установлены в регрессор, после этого выполнено сохранение модели.

Для построения нейросети воспользуемся библиотекой tensorflow. Строим полносвязную нейронную сеть, в которую входит 3 скрытых слоя: один с 100 нейронами и два по 8 нейронов, с функцией активации «relu» и один выходной слой c функцией активации «linear». Данная архитектура была выбрана исходя из анализа литературных данных и статей, где были приведены примеры решения подобных задач прогназирования.

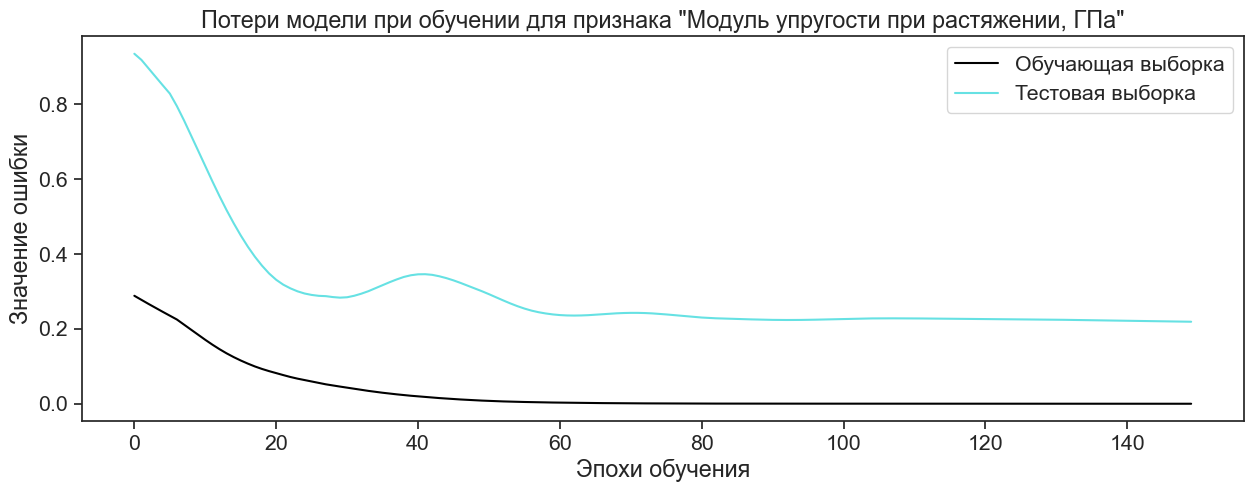
В качестве оптимизатора берём «Adam», в качестве функции потерь берём «mse».

Пример архитектуры нейросети приведен на рис. 4.



*Рис. 4 — Архитектура нейросети*

Запуск обучения нейросети выполнен со следующими параметрами: пропорция разбиения данных на тестовые и валидационные 10%, количество эпох 150. График обучения нейросети приведен на рис. 5.



*Рис. 5 — График обучения нейросети*

Существует множество различных метрик качества. В данной работе использованы: MAE (Mean Absolute Error) средняя абсолютная ошибка, MSE (Mean Squared Error) среднеквадратичная ошибка, MAPE (Mean Absolute Percentage Error) или средняя абсолютная процентная ошибка, R2 или коэффициент детерминации.

R2 в норме принимает положительные значения. Эту метрику надо максимизировать. Отрицательные значение коэффициента детерминации означают плохую объясняющую способность модели.

Результаты работы моделей для прогноза признаков: «Модуль упругости при растяжении, ГПа», «Прочность при растяжении, МПа», представлены в виде таблице 1.

*Таблица 1 — Результаты обучения моделей для прогноза целевых признаков*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Метод | МАЕ | MSE | R2\_train | R2 | MAPE |
| Модуль упругости при растяжении, ГПа | | | | | |
| RandomForest | 0.00 | 0.00 | 1.00 | 1.00 | 0.00 |
| K\_Neighbors | 0.00 | 0.00 | 1.00 | 1.00 | 0.00 |
| LCV | 0.01 | 0.01 | 0.99 | 0.99 | 0.50 |
| LinearRegression | 0.04 | 0.01 | 1.00 | 0.98 | 0.52 |
| Neural Network | 0.20 | 0.05 | 0.87 | 0.76 | 0.68 |
| GradientBoosting | 0.45 | 0.45 | 0.19 | 0.17 | 1.02 |
| DecisionTree | 0.50 | 0.50 | 0.56 | -0.25 | 0.66 |
| SVR | 0.50 | 0.50 | -0.44 | -0.64 | 4.99 |
| Прочность при растяжении, МПа | | | | | |
| RandomForest | 0.00 | 0.00 | 1.00 | 1.00 | 0.00 |
| K\_Neighbors | 0.00 | 0.00 | 1.00 | 1.00 | 0.00 |
| LCV | 0.01 | 0.01 | 0.99 | 0.99 | 0.50 |
| LinearRegression | 0.04 | 0.01 | 1.00 | 0.98 | 0.52 |
| Neural Network | 0.17 | 0.04 | 0.93 | 0.81 | 0.68 |
| GradientBoosting | 0.37 | 0.37 | 0.62 | 0.44 | 1.02 |
| DecisionTree | 0.33 | 0.33 | 0.22 | 0.33 | 0.66 |
| SVR | 0.50 | 0.50 | -0.44 | -0.64 | 4.99 |

Проведенное исследование позволило разработать рабочие модели, которые демонстрируют определённый потенциал в решении задач прогнозирования конечных свойств композиционных материалов. Однако, ограниченность исходных данных существенно сказывается на надежности и обобщаемости полученных результатов. Для достижения более точных и устойчивых моделей необходимо провести дополнительные исследования с использованием значительно большего объема данных. Это позволит не только улучшить качество моделей, но и расширить их применение в практических сценариях.

**Список использованных источников**

1. Композиционные материалы : учебное пособие для вузов / Д. А. Иванов, А. И. Ситников, С. Д. Шляпин ; под редакцией А. А. Ильина. — Москва : Издательство Юрайт, 2019 — 253 с.

2. Akinpelu S.B. et al. Interpretable machine learning methods to predict the 1mechanical properties of ABX3 perovskites, Results in Physics, 65 (2024) 107978

3. Вережбовский Г.Б. Прогнозирование физических свойств многокомпонентных композитных материалов с полимерной матрицей, Инженерный вестник Дона, №7 (2022)

4. Ле Тхи Ныт Сыонг, Бондарев А.В., Бондарева Л.И., Монакова А.С., Баршин А.В. Применение алгоритмов машинного обучения в прогнозе результата пиролитического анализа. Известия высших учебных заведений. Геология и разведка. 2020;63(6):8—19.

5. Ахметсафин Р. Д., Ахметсафина Р. З. Применение методов машинного обучения для прогноза или замещения недостающих каротажных данных // Изв. вузов. Приборостроение. 2021. Т. 64, № 7. С. 532—541.

6. Силен Дэви, Мейсман Арно, Али Мохамед. Основы Data Science и Big Data. Python и наука о данных. – СПб.: Питер, 2017. – 336 с.: ил.

7. 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.

8. Лутц М. Изучаем Python, том 1, 5-е изд.: Пер. с англ. — СПб.: ООО “Диалектика”, 2019. 832 с.