



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана

Выпускная квалификационная работа по курсу "Data Science"

Тема: Прогнозирование конечных свойств
новых материалов
(композиционных материалов).

Докладчик: Запорожец Марина Владимировна



Объединение датасетов и описательная статистика

Датасет состоит из двух файлов: X_br и X_nir.

Файл X_br содержит 10 признаков и индекс; 1023 строки.

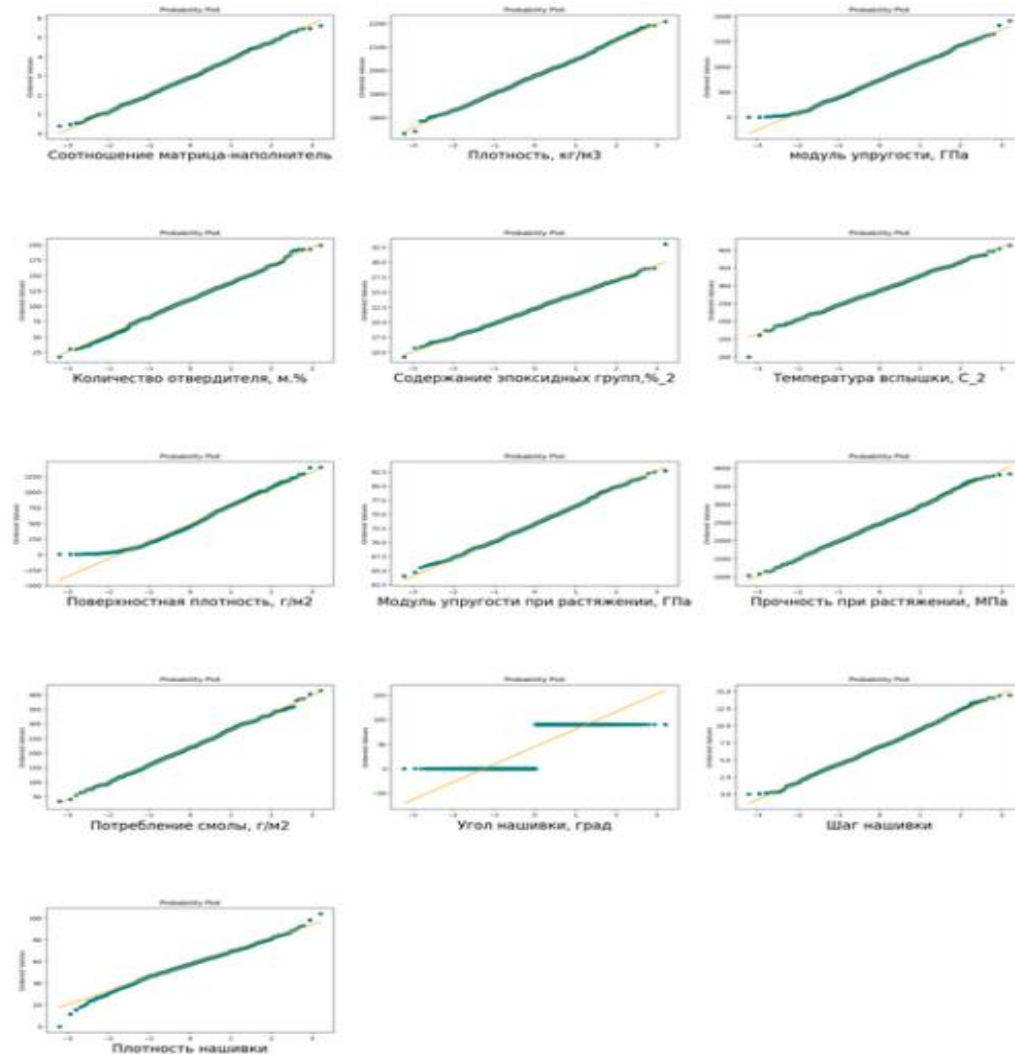
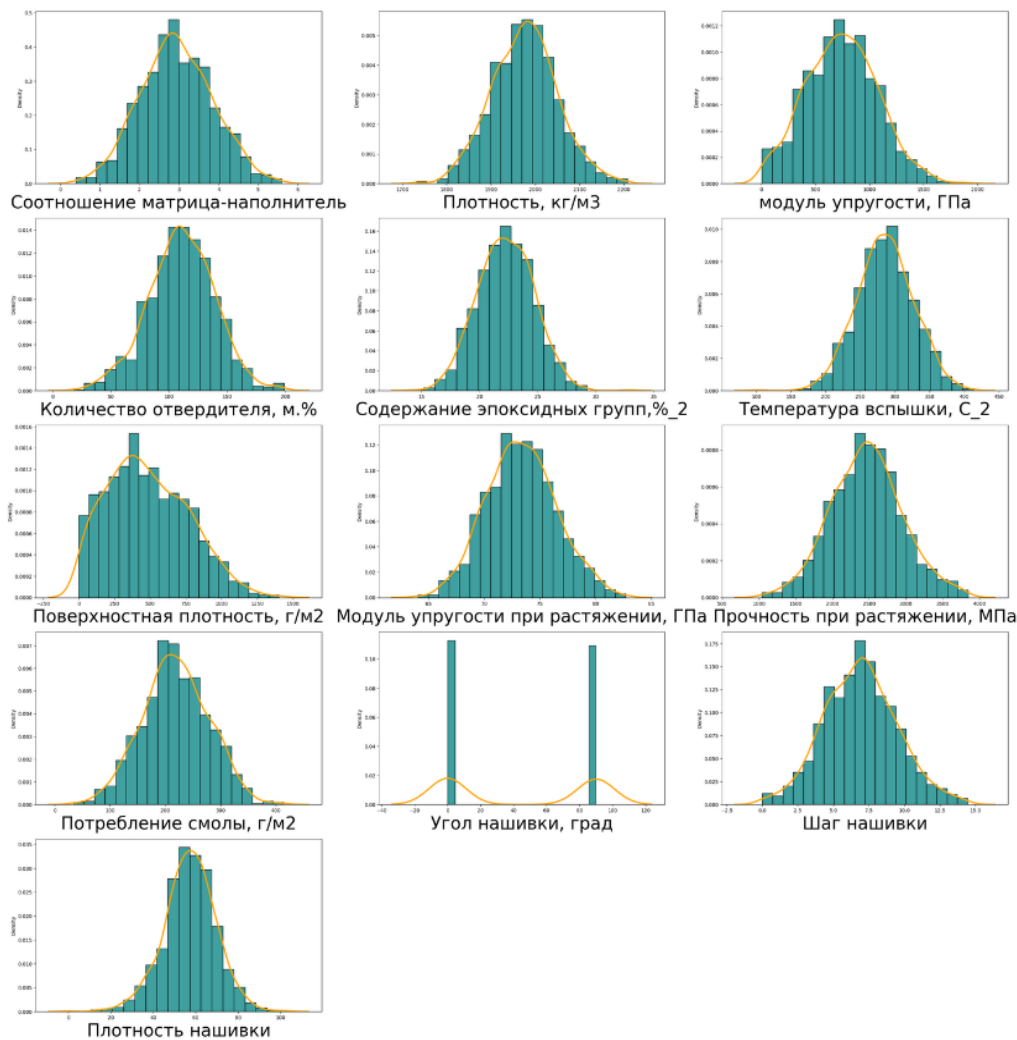
Файл X_nir содержит 3 признака и индекс; строк: 1040.

Файлы объединены с типом INNER по индексу. Дальнейшие исследования проводим с объединенным датасетом, содержащим 13 признаков и 1023 строк или объектов.

	Соотношение матрица-наполнитель	Плотность, кг/м3	модуль упругости, ГПа	Количество отвердителя, м.%	Содержание эпоксидных групп,%_2	Температура вспышки, С_2	Поверхностная плотность, г/м2	Модуль упругости при растяжении, ГПа	Прочность при растяжении, МПа	Потребление смолы, г/м2	Угол нашивки, град	Шаг нашивки	Плотность нашивки
count	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00	1023.00
mean	2.93	1975.73	739.92	110.57	22.24	285.88	482.73	73.33	2466.92	218.42	44.25	6.90	57.15
std	0.91	73.73	330.23	28.30	2.41	40.94	281.31	3.12	485.63	59.74	45.02	2.56	12.35
min	0.39	1731.76	2.44	17.74	14.25	100.00	0.60	64.05	1036.86	33.80	0.00	0.00	0.00
25%	2.32	1924.16	500.05	92.44	20.61	259.07	266.82	71.25	2135.85	179.63	0.00	5.08	49.80
50%	2.91	1977.62	739.66	110.56	22.23	285.90	451.86	73.27	2459.52	219.20	0.00	6.92	57.34
75%	3.55	2021.37	961.81	129.73	23.96	313.00	693.23	75.36	2767.19	257.48	90.00	8.59	64.94
max	5.59	2207.77	1911.54	198.95	33.00	413.27	1399.54	82.68	3848.44	414.59	90.00	14.44	103.99

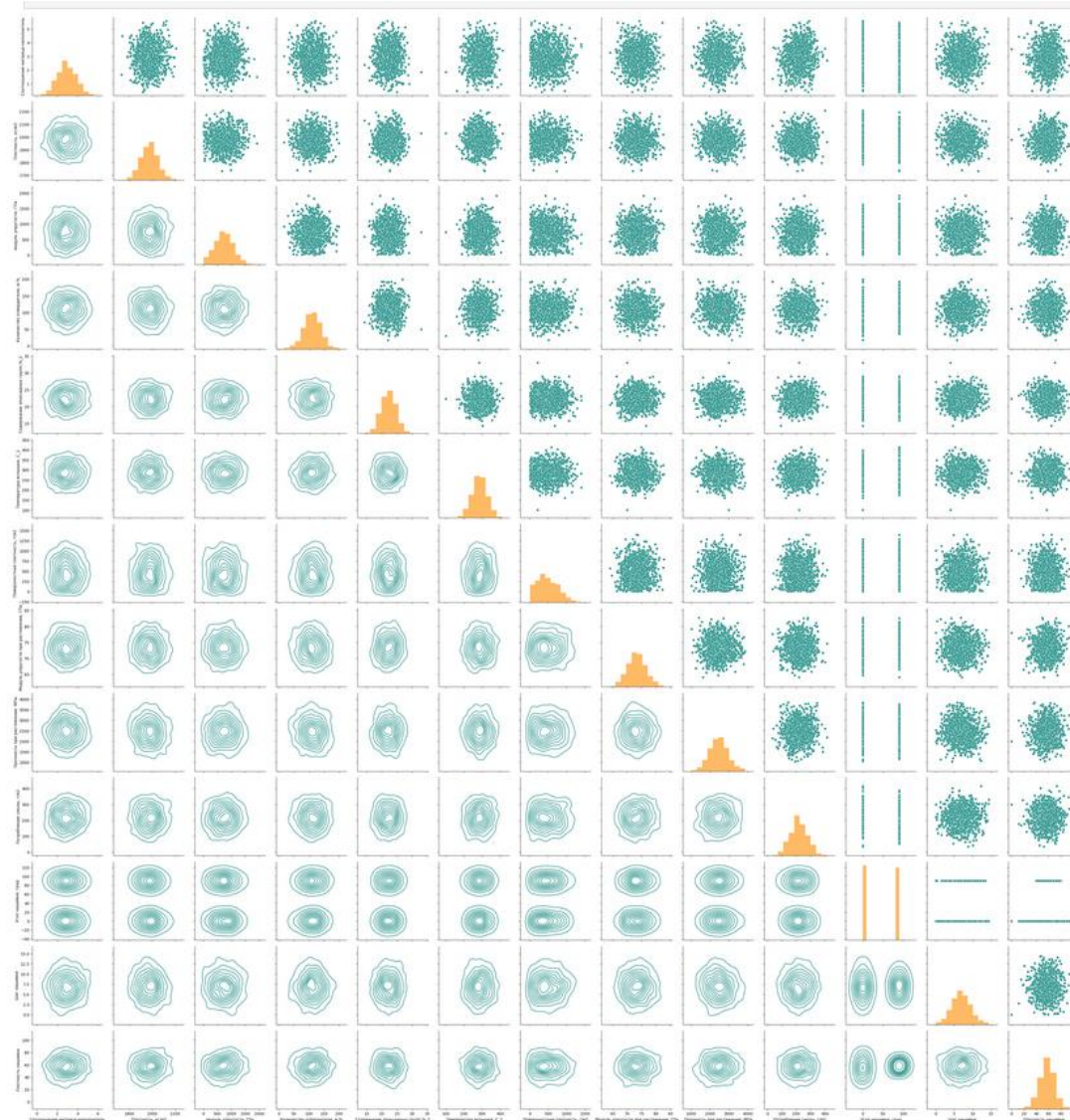
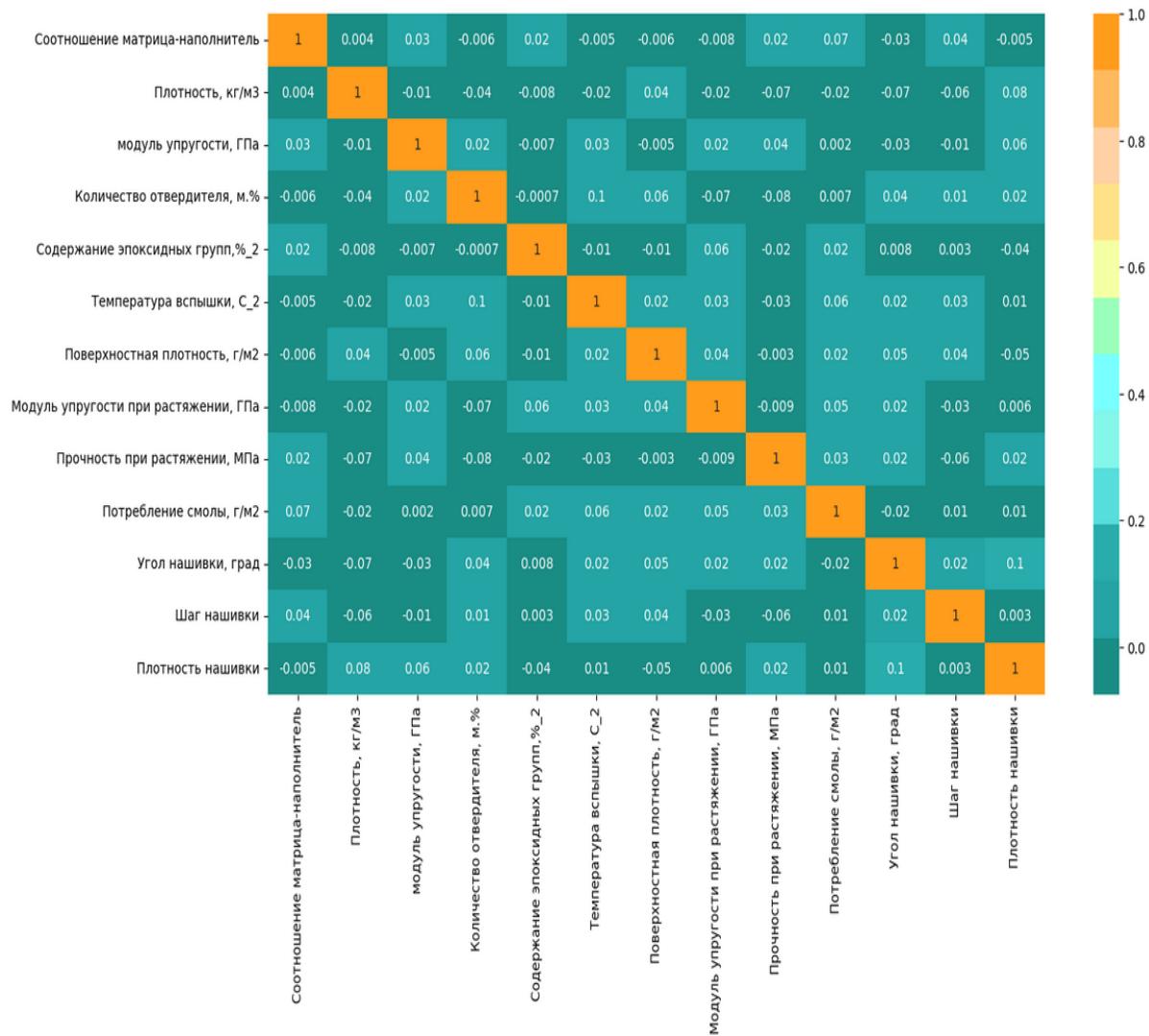


Гистограммы распределения и графики Q-Q





Тепловая карта и парная диаграмма. Зависимостей не обнаружено.





Описательная статистика очищенных данных

	Соотношение матрица-наполнитель	Плотность, кг/м3	модуль упругости, ГПа	Количество отвердителя, м.%	Содержание эпоксидных групп,%_2	Температура вспышки, С_2	Поверхностная плотность, г/м2	Модуль упругости при растяжении, ГПа	Прочность при растяжении, МПа	Потребление смолы, г/м2	Угол нашивки, град	Шаг нашивки	Плотность нашивки
count	922.00	922.00	922.00	922.00	922.00	922.00	922.00	922.00	922.00	922.00	922.00	922.00	922.00
mean	2.93	1974.12	736.12	111.14	22.20	286.18	482.43	73.30	2461.49	218.05	45.98	6.93	57.56
std	0.90	71.04	327.61	26.75	2.39	39.42	280.44	3.03	453.56	57.14	45.01	2.51	11.12
min	0.55	1784.48	2.44	38.67	15.70	179.37	0.60	65.79	1250.39	72.53	0.00	0.04	28.66
25%	2.32	1923.32	498.54	92.86	20.56	259.21	264.35	71.24	2148.18	179.88	0.00	5.14	50.28
50%	2.91	1977.32	736.18	111.16	22.18	286.22	457.73	73.25	2455.97	218.70	90.00	6.97	57.58
75%	3.55	2020.05	956.96	130.11	23.96	313.01	695.53	75.31	2751.23	256.62	90.00	8.61	64.84
max	5.31	2161.57	1628.00	181.83	28.96	386.07	1291.34	81.20	3654.43	359.05	90.00	13.73	86.01

Выделение целевых переменных «Модуль упругости при растяжении, ГПа» и «Прочность при растяжении, МПа»

```
X_no_norm = data[['Соотношение матрица-наполнитель', 'Плотность, кг/м3',  
                  'модуль упругости, ГПа', 'Количество отвердителя, м.%',  
                  'Содержание эпоксидных групп,%_2', 'Температура вспышки, С_2',  
                  'Поверхностная плотность, г/м2',  
                  'Потребление смолы, г/м2', 'Угол нашивки, град',  
                  'Шаг нашивки', 'Плотность нашивки']]  
  
y_1 = data['Модуль упругости при растяжении, ГПа']  
y_2 = data['Прочность при растяжении, МПа']
```

Посмотрим размерность.

```
X_no_norm.shape, y_1.shape, y_2.shape
```

```
((922, 11), (922,), (922,))
```

Выделение целевой переменной «Соотношение матрица – наполнитель»

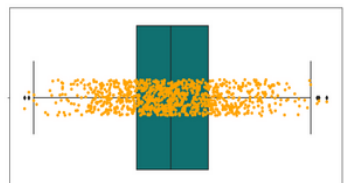
```
X_no_norm = data[['Плотность, кг/м3',  
                  'модуль упругости, ГПа', 'Количество отвердителя, м.%',  
                  'Содержание эпоксидных групп,%_2', 'Температура вспышки, С_2',  
                  'Поверхностная плотность, г/м2', 'Модуль упругости при растяжении, ГПа',  
                  'Прочность при растяжении, МПа', 'Потребление смолы, г/м2',  
                  'Угол нашивки, град', 'Шаг нашивки', 'Плотность нашивки']]  
  
y = data['Соотношение матрица-наполнитель']
```

```
X_no_norm.shape, y.shape
```

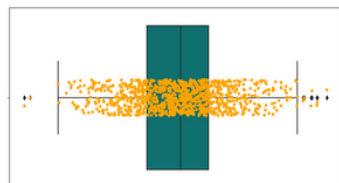
```
((922, 12), (922,))
```



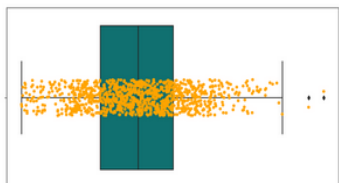
Графики Voxplot с выбросами (слева) и Voxplot без выбросов (справа).



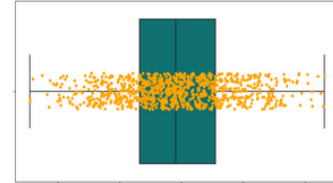
Соотношение матрица-наполнитель



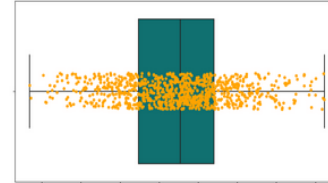
Плотность, кг/м3



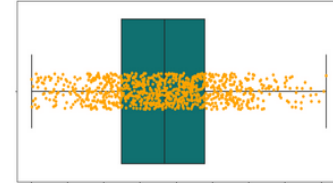
модуль упругости, ГПа



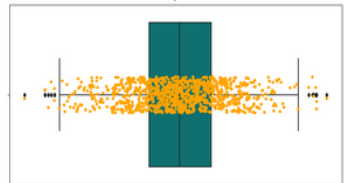
Соотношение матрица-наполнитель



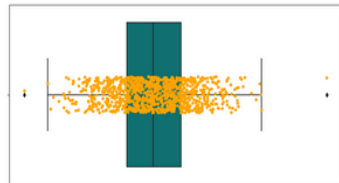
Плотность, кг/м3



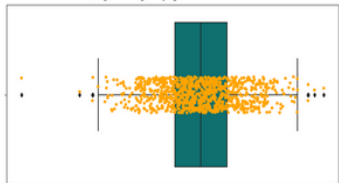
модуль упругости, ГПа



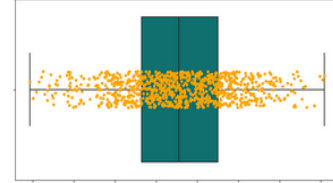
Количество отвердителя, м.%



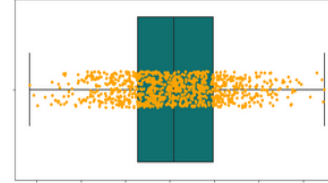
Содержание эпоксидных групп,%_2



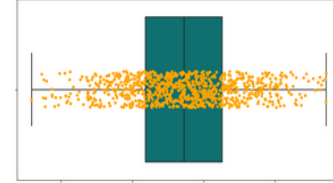
Температура вспышки, С_2



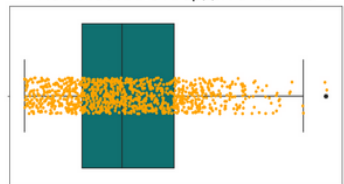
Количество отвердителя, м.%



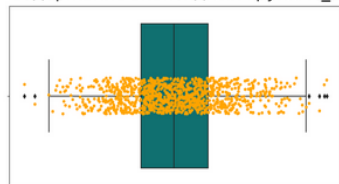
Содержание эпоксидных групп,%_2



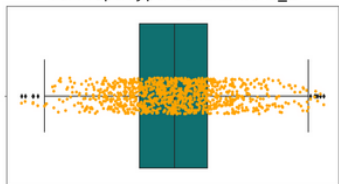
Температура вспышки, С_2



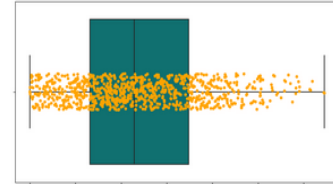
Поверхностная плотность, г/м2



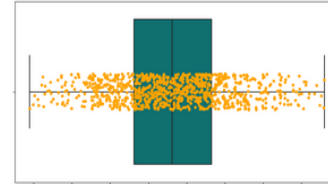
Модуль упругости при растяжении, ГПа



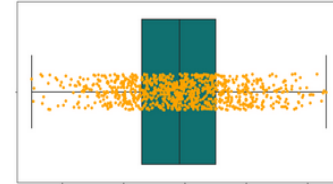
Прочность при растяжении, МПа



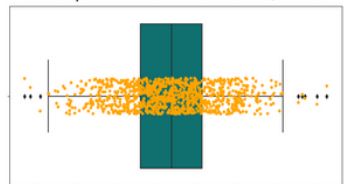
Поверхностная плотность, г/м2



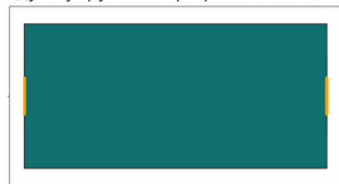
Модуль упругости при растяжении, ГПа



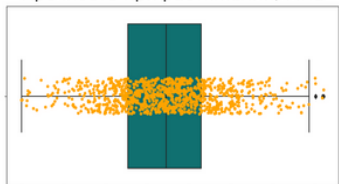
Прочность при растяжении, МПа



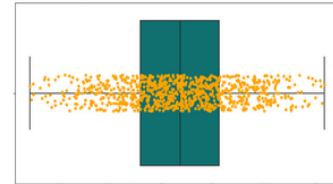
Потребление смолы, г/м2



Угол нашивки, град



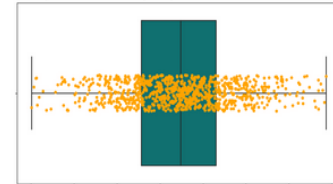
Шаг нашивки



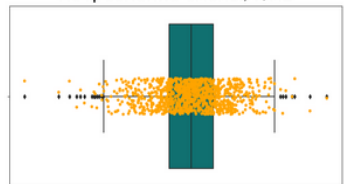
Потребление смолы, г/м2



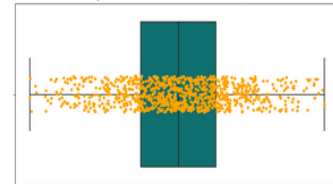
Угол нашивки, град



Шаг нашивки



Плотность нашивки



Плотность нашивки



Используемые модели и метрики

Используемые модели регрессии

- Линейная регрессия
- Метод k-ближайших соседей
- Случайный лес
- Метод опорных векторов для регрессии
- Градиентный бустинг
- Нейронная сеть

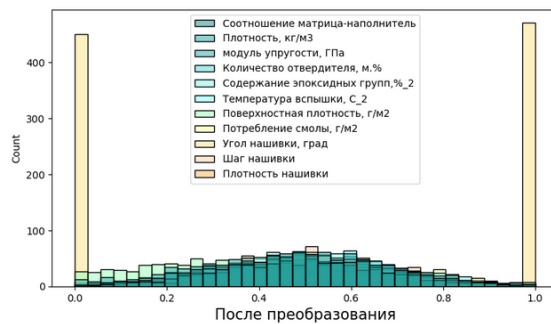
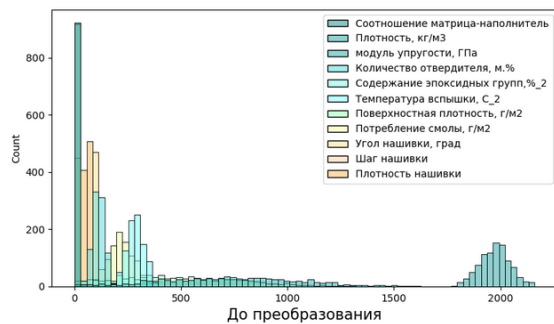
Метрики качества работы моделей

- R^2 - коэффициент детерминации
- MSE – (Mean Squared Error) средняя квадратичная ошибка
- RMSE (Root Mean Squared Error) - корень из средней квадратичной ошибки
- MAE (Mean Absolute Error) - средняя абсолютная ошибка
- MAPE (Mean Absolute Percentage Error) - средняя абсолютная ошибка в процентах

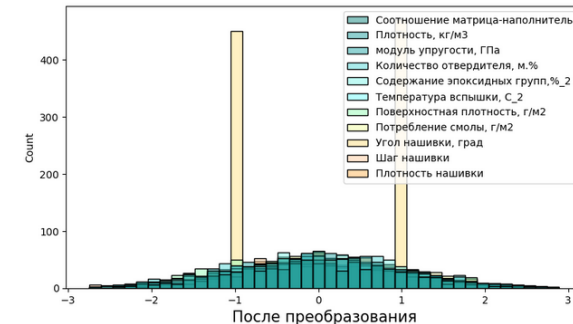
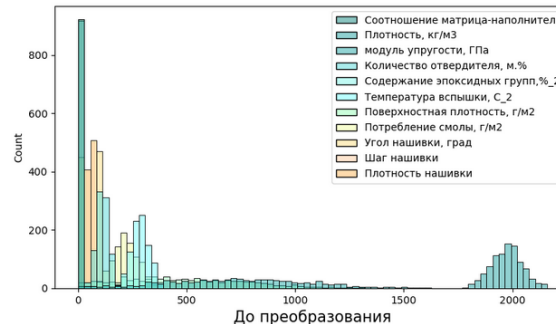


Графики преобразований при применении нормализаторов и стандартизаторов (выбор GridSearch и RandomSearch)

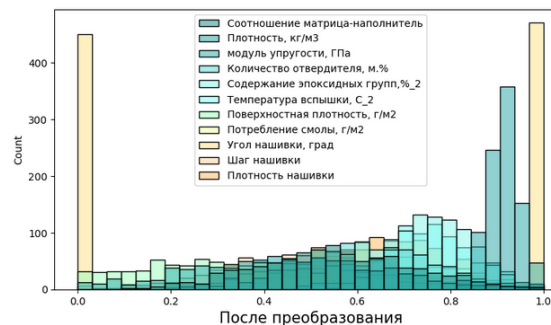
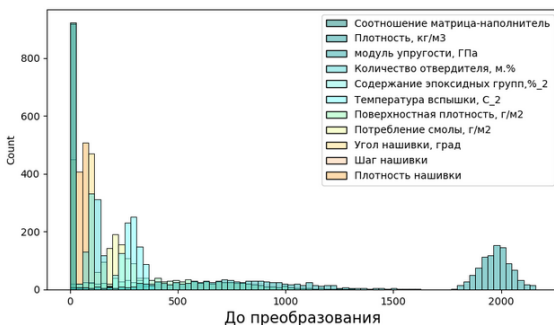
MinMaxScaler()



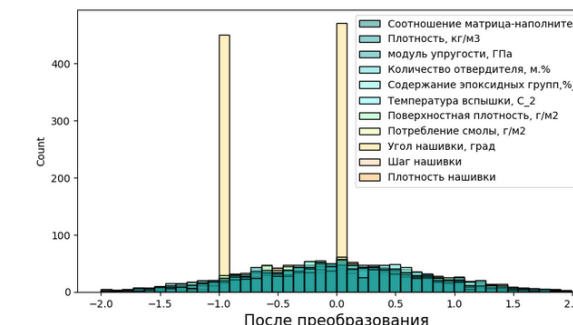
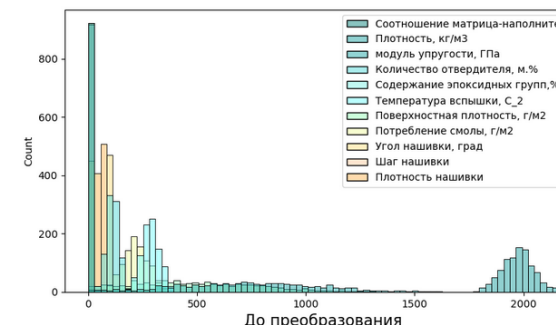
StandardScaler()



MaxAbsScaler()



RobustScaler()





ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГУ им. Н.Э. Баумана

Работа моделей предсказания Модуля упругости при растяжении (слева) и Прочности при растяжении (справа)

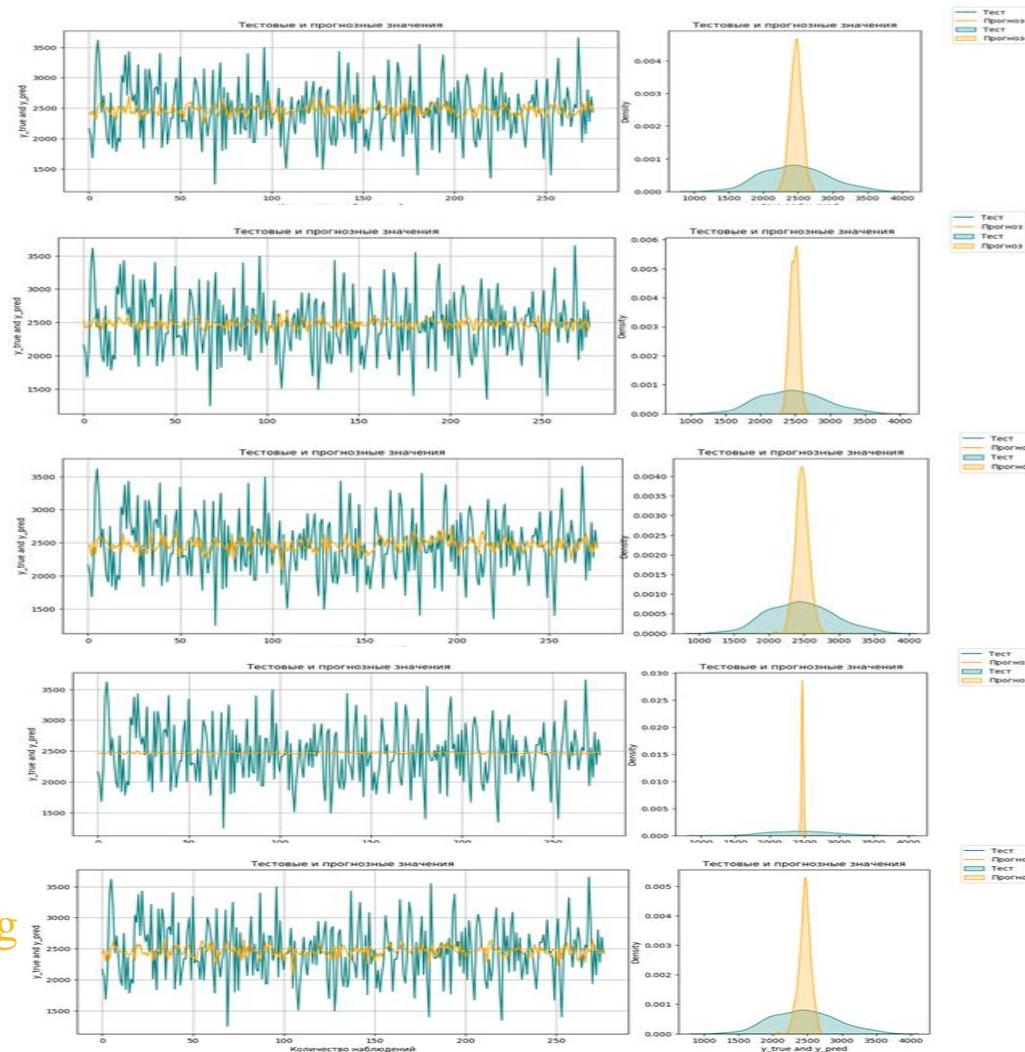
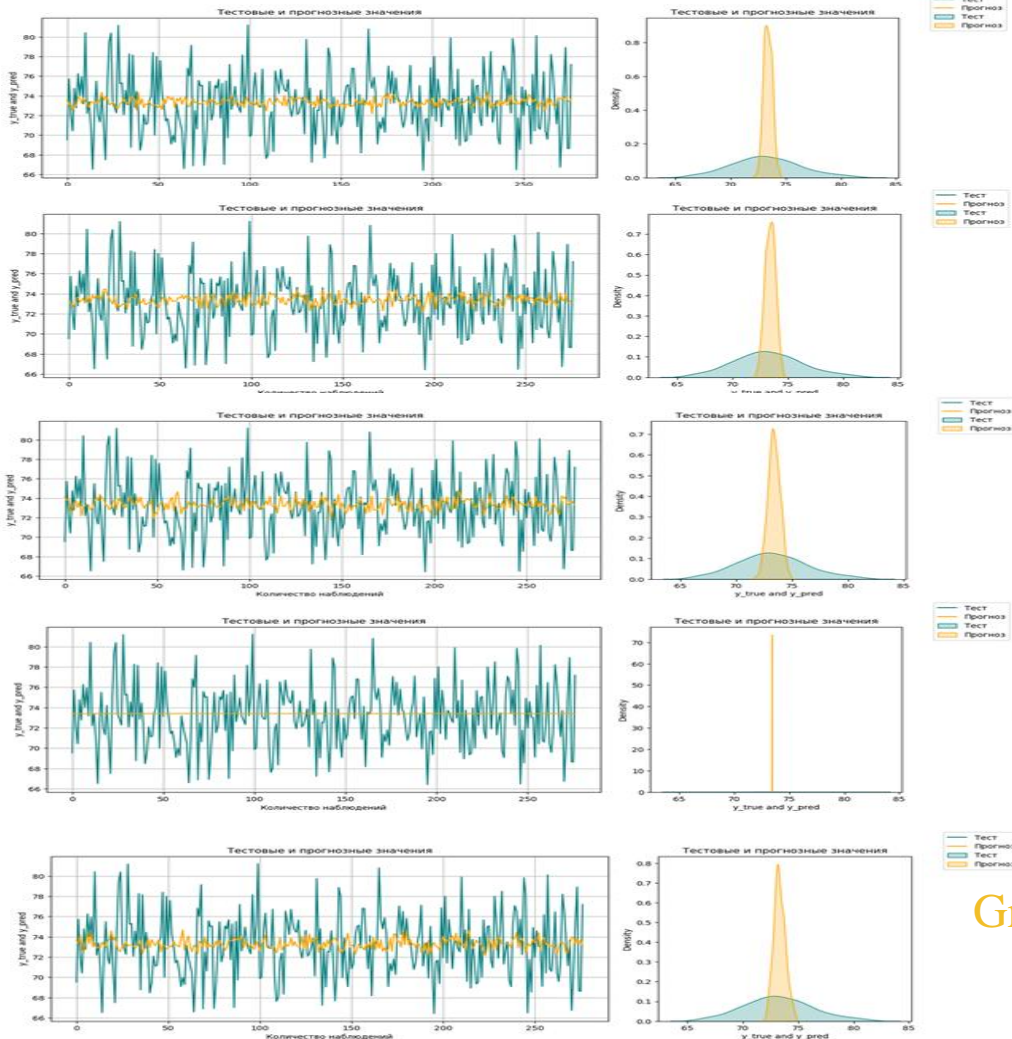
Linear
Regression

Kneighbors
Regressor

Random
Forest

Support Vector
Regression

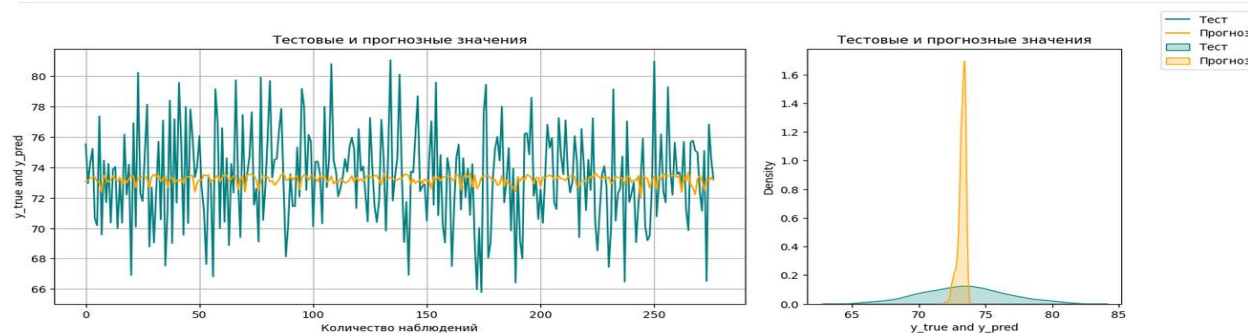
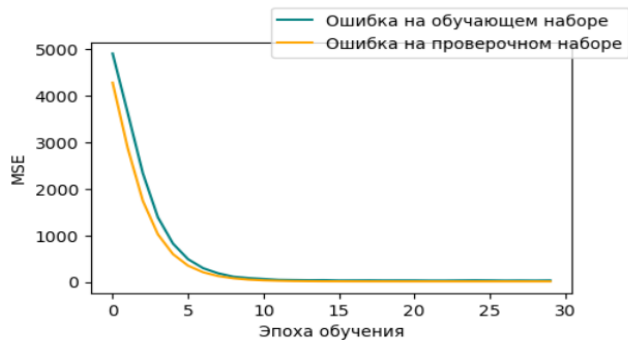
Gradient Boosting
Regressor



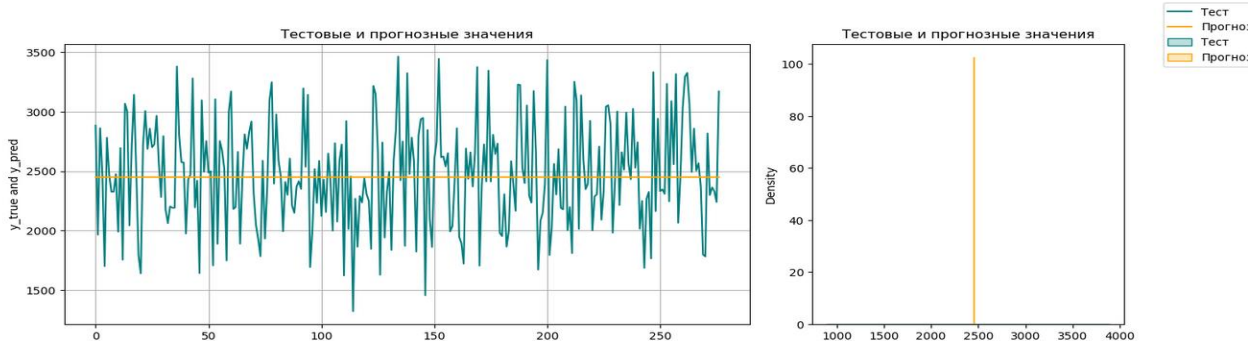
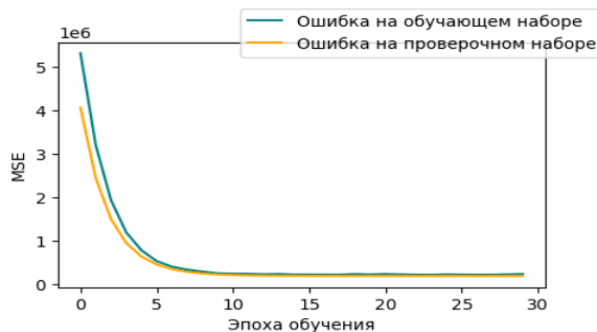


Работа полносвязной нейронной сети

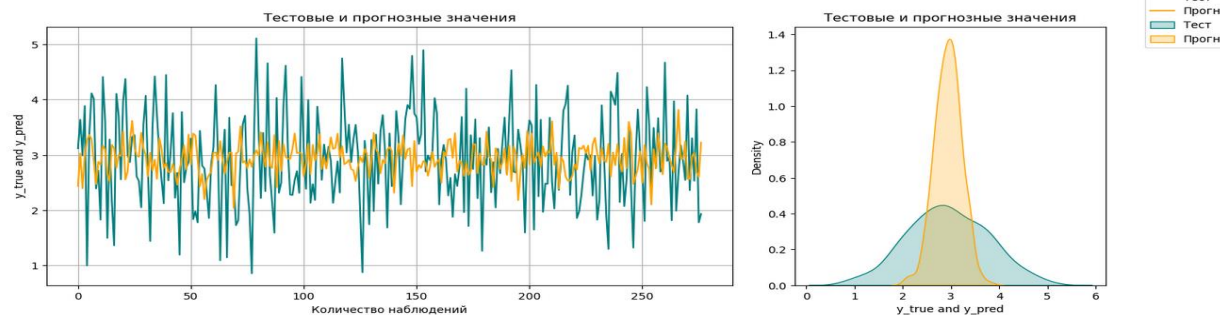
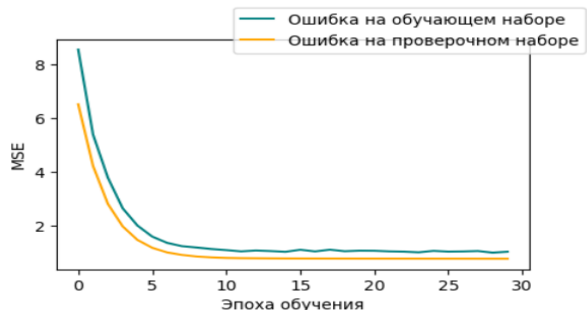
Модуль
упругости при
растяжении



Прочность при
растяжении



Соотношение
матрица -
наполнитель





Метрики работы моделей

Модуль упругости при растяжении

	Model_regr	R2	MSE	RMSE	MAE	MAPE
0	LinearRegression	-0.03	9.71	3.12	2.48	0.03
1	KNeighborsRegressor	-0.01	9.57	3.09	2.48	0.03
2	RandomForestRegressor	-0.05	9.94	3.15	2.48	0.03
3	SVR	-0.01	9.51	3.08	2.46	0.03
4	GradientBoostingRegressor	-0.04	9.83	3.14	2.50	0.03
5	Neural_net_minmax	0.01	9.75	3.12	2.50	0.03
6	Neural_net_stand_scal	-0.01	9.98	3.16	2.51	0.03
7	Neural_net_max_abs_scal	-0.01	9.89	3.14	2.52	0.03
8	Neural_net_robust_sc	-0.08	10.58	3.25	2.58	0.04

Лучший показатель R2 = 0.009 в модели Neural_net_minmax
Лучший показатель MSE = 9.5116 в модели SVR
Лучший показатель MAE = 2.4591 в модели SVR

Прочность при растяжении

	Model_regr	R2	MSE	RMSE	MAE	MAPE
0	LinearRegression	-0.01	214500.17	463.14	372.07	0.16
1	KNeighborsRegressor	0.01	209932.30	458.18	369.54	0.16
2	RandomForestRegressor	0.01	209452.24	457.66	366.52	0.16
3	SVR	0.00	211079.56	459.43	370.15	0.16
4	GradientBoostingRegressor	-0.02	215203.10	463.90	371.32	0.16
5	Neural_net_minmax	-0.00	198679.67	445.74	360.30	0.15
6	Neural_net_stand_scal	-0.00	198576.58	445.62	360.28	0.15
7	Neural_net_max_abs_scal	-0.01	199034.83	446.13	360.37	0.15
8	Neural_net_robust_sc	-0.00	198663.93	445.72	360.30	0.15

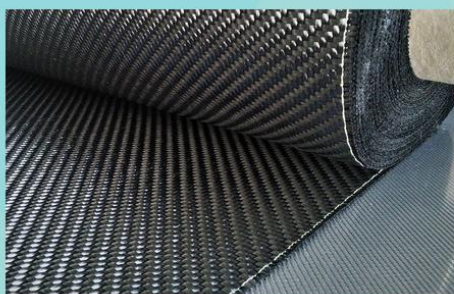
Лучший показатель R2 = 0.0091 в модели RandomForestRegressor
Лучший показатель MSE = 198576.5801 в модели Neural_net_stand_scal
Лучший показатель MAE = 360.2825 в модели Neural_net_stand_scal

Соотношение
матрица - наполнитель

R2: -0.126
MSE: 0.774
RMSE: 0.88
MAE: 0.717
MAPE: 0.294



<https://composite-predict.onrender.com>



Предсказание конечных свойств получаемых композиционных материалов по имеющимся измерениям.

На вход подаются данные о начальных свойствах компонентов композиционных материалов (количество связующего, наполнителя, температурный режим отверждения и т.д.).

Матрица-наполнитель

Модуль упругости при растяжении

Прочность при растяжении

Композиционные материалы - это искусственно созданные материалы, состоящие из нескольких других с четкой границей между ними. Композиты обладают теми свойствами, которые не наблюдаются у компонентов по отдельности. При этом композиты являются монолитным материалом, т.е. компоненты материала неотделимы друг от друга без разрушения конструкции в целом. 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).

Соотношение матрица-наполнитель для введенных параметров:
[[2.065465]]

Расчет соотношения матрица-наполнитель

Введите параметры

Плотность, кг/м3

Модуль упругости, ГПа

Количество отвердителя, м. %

Содержание эпоксидных групп, %₂

Температура вспышки, C₂

Поверхностная плотность, г/м2

Модуль упругости при растяжении, ГПа

Прочность при растяжении, МПа

Потребление смолы, г/м2

Угол нашивки, град

Шаг нашивки

Плотность нашивки

Рассчитать

Сброс

[Вернуться на главную страницу](#)



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана



do.bmstu.ru