

Исходный датасет

	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	Шаг нашивки	Плотность нашивки	Угол нашивки, град	Модуль упругости при растяжении, ГПа	Прочность при растяжении, МПа
0	1.857143	2030.000000	738.736842	30.000000	22.267857	100.000000	210.000000	220.000000	4.000000	57.000000	0	70.000000	3000.000000
1	1.857143	2030.000000	738.736842	50.000000	23.750000	284.615385	210.000000	220.000000	4.000000	60.000000	0	70.000000	3000.000000
2	1.857143	2030.000000	738.736842	49.900000	33.000000	284.615385	210.000000	220.000000	4.000000	70.000000	0	70.000000	3000.000000
3	1.857143	2030.000000	738.736842	129.000000	21.250000	300.000000	210.000000	220.000000	5.000000	47.000000	0	70.000000	3000.000000
4	2.771331	2030.000000	753.000000	111.860000	22.267857	284.615385	210.000000	220.000000	5.000000	57.000000	0	70.000000	3000.000000
1018	2.271346	1952.087902	912.855545	86.992183	20.123249	324.774576	209.198700	125.007669	9.076380	47.019770	90	73.090961	2387.292495
1019	3.444022	2050.089171	444.732634	145.981978	19.599769	254.215401	350.660830	117.730099	10.565614	53.750790	90	72.920827	2360.392784
1020	3.280604	1972.372865	416.836524	110.533477	23.957502	248.423047	740.142791	236.606764	4.161154	67.629684	90	74.734344	2662.906040
1021	3.705351	2066.799773	741.475517	141.397963	19.246945	275.779840	641.468152	197.126067	6.313201	58.261074	90	74.042708	2071.715856
1022	3.808020	1890.413468	417.316232	129.183416	27.474763	300.952708	758.747882	194.754342	6.078902	77.434468	90	74.309704	2856.328932
1023 ro	ws × 13 columns												



Характеристики исходного датасета

Таблица основных статистических характеристик атрибутов

	Соотношение матрица- наполнитель	Плотность, кг/м3	Модуль упругости, ГПа	Количество отвердителя, м.%	Содержание эпоксидных групп,%_2	Температура вспышки, C_2	Поверхностная плотность, г/ м2	Потребление смолы, г/м2	Шаг нашивки	Плотность нашивки	Угол нашивки, град	Модуль упругости при растяжении, ГПа	Прочность при растяжении, МПа
count	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000
mean	2.930366	1975.734888	739.923233	110.570769	22.244390	285.882151	482.731833	218.423144	6.899222	57.153929	44.252199	73.328571	2466.922843
std	0.913222	73.729231	330.231581	28.295911	2.406301	40.943260	281.314690	59.735931	2.563467	12.350969	45.015793	3.118983	485.628006
min	0.389403	1731.764635	2.436909	17.740275	14.254985	100.000000	0.603740	33.803026	0.000000	0.000000	0.000000	64.054061	1036.856605
25%	2.317887	1924.155467	500.047452	92.443497	20.608034	259.066528	266.816645	179.627520	5.080033	49.799212	0.000000	71.245018	2135.850448
50%	2.906878	1977.621657	739.664328	110.564840	22.230744	285.896812	451.864365	219.198882	6.916144	57.341920	0.000000	73.268805	2459.524526
75%	3.552660	2021.374375	961.812526	129.730366	23.961934	313.002106	693.225017	257.481724	8.586293	64.944961	90.000000	75.356612	2767.193119
max	5.591742	2207.773481	1911.536477	198.953207	33.000000	413.273418	1399.542362	414.590628	14.440522	103.988901	90.000000	82.682051	3848.436732



Характеристики исходного датасета

Коэффициенты корреляции

	•			•									
	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	Шаг нашивки	Плотность нашивки	Угол нашивки, град	Модуль упругости при растяжении, ГПа	Прочность при растяжении, МПа
Соотношение матрица- наполнитель	1.000000	0.003841	0.031700	-0.006445	0.019766	-0.004776	-0.006272	0.072531	0.036437	-0.004652	-0.031073	-0.008411	0.024148
Плотность, кг/м3	0.003841	1.000000	-0.009647	-0.035911	-0.008278	-0.020695	0.044930	-0.015937	-0.061015	0.080304	-0.068474	-0.017602	-0.069981
Модуль упругости, ГПа	0.031700	-0.009647	1.000000	0.024049	-0.006804	0.031174	-0.005306	0.001840	-0.009875	0.056346	-0.025417	0.023267	0.041868
Количество отвердителя, м.%	-0.006445	-0.035911	0.024049	1.000000	-0.000684	0.095193	0.055198	0.007446	0.014887	0.017248	0.038570	-0.065929	-0.075375
Содержание эпоксидных групп,%_2	0.019766	-0.008278	-0.006804	-0.000684	1.000000	-0.009769	-0.012940	0.015165	0.003022	-0.039073	0.008052	0.056828	-0.023899
Температура вспышки, С_2	-0.004776	-0.020695	0.031174	0.095193	-0.009769	1.000000	0.020121	0.059954	0.025795	0.011391	0.020695	0.028414	-0.031763
Поверхностная плотность, г/м2	-0.006272	0.044930	-0.005306	0.055198	-0.012940	0.020121	1.000000	0.015692	0.038332	-0.049923	0.052299	0.036702	-0.003210
Потребление смолы, г/м2	0.072531	-0.015937	0.001840	0.007446	0.015165	0.059954	0.015692	1.000000	0.013394	0.012239	-0.015334	0.050938	0.028602
Шаг нашивки	0.036437	-0.061015	-0.009875	0.014887	0.003022	0.025795	0.038332	0.013394	1.000000	0.003487	0.023616	-0.029468	-0.059547
Плотность нашивки	-0.004652	0.080304	0.056346	0.017248	-0.039073	0.011391	-0.049923	0.012239	0.003487	1.000000	0.107947	0.006476	0.019604
Угол нашивки, град	-0.031073	-0.068474	-0.025417	0.038570	0.008052	0.020695	0.052299	-0.015334	0.023616	0.107947	1.000000	0.023003	0.023398
Модуль упругости при растяжении, ГПа	-0.008411	-0.017602	0.023267	-0.065929	0.056828	0.028414	0.036702	0.050938	-0.029468	0.006476	0.023003	1.000000	-0.009009



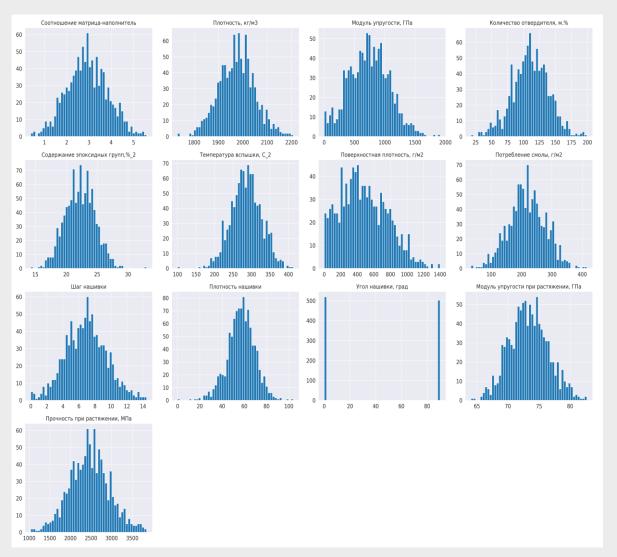
Характеристики исходного датасета

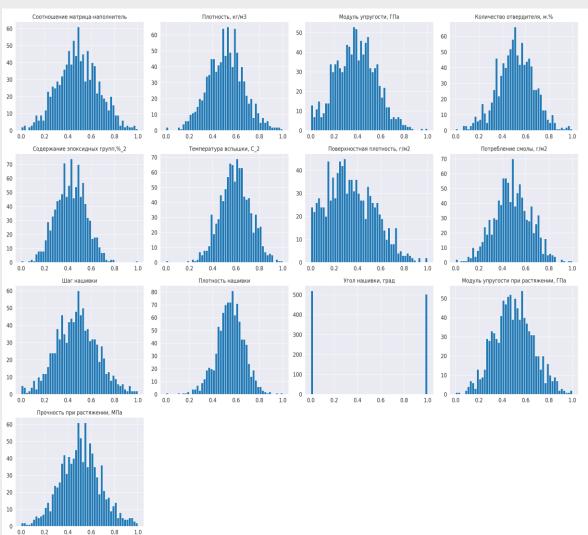
Характеристики данных по принадлежности к атрибутам датасета

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1023 entries, 0 to 1022
Data columns (total 13 columns):
    Column
                                        Non-Null Count Dtype
    Соотношение матрица-наполнитель
                                        1023 non-null float64
                                        1023 non-null float64
    Плотность, кг/м3
2 Модуль упругости, ГПа
                                        1023 non-null float64
    Количество отвердителя, м.%
                                        1023 non-null float64
    Содержание эпоксидных групп,%_2
                                        1023 non-null float64
   Температура вспышки, С 2
                                        1023 non-null float64
    Поверхностная плотность, г/м2
                                        1023 non-null float64
    Потребление смолы, г/м2
                                        1023 non-null float64
                                        1023 non-null float64
    Шаг нашивки
                                        1023 non-null float64
    Плотность нашивки
 10 Угол нашивки, град
                                        1023 non-null int64
    Модуль упругости при растяжении, ГПа 1023 non-null
                                                       float64
12 Прочность при растяжении, МПа
                                        1023 non-null float64
dtypes: float64(12), int64(1)
memory usage: 104.0 KB
```



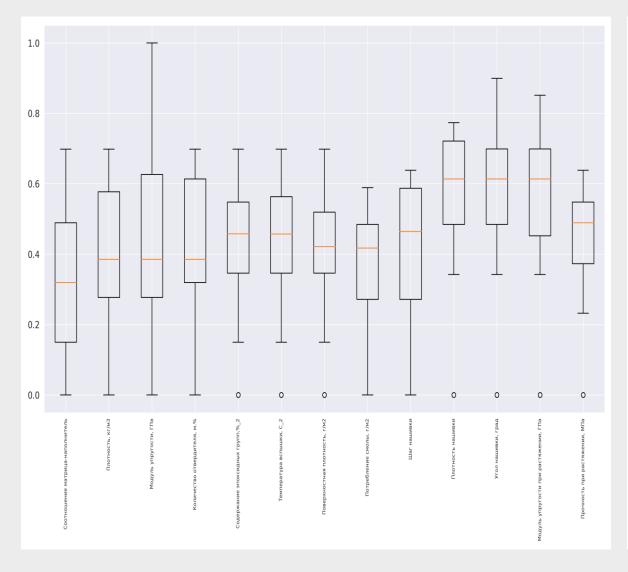
Графические интерпретации исходного и нормализованного датасета Гистограммы распределения по атрибутам

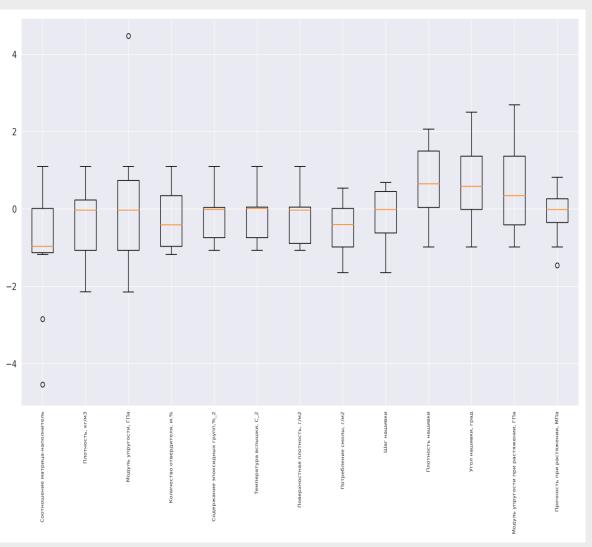






Графические интерпретации нормализованного и стандартизированного датасета Диаграммы ящиков с усами



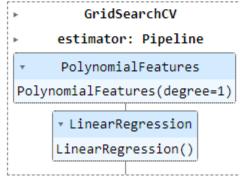




Разработка и обучение моделей

▼ Полномиальная регрессия

```
# низашло предыдушее. что ж давайте еще парочку мат моделей катанем, а то, что
    # зря что ли часы с г-жой Тверской прошли)))
    # напомню я плохо понимаю что конкретный атрибут делает, но я как настоящий DS
    # рандомно перебираю модели
    # главное делать это с уверенным видом
    def PolynomialRegression(degree=1, **kwargs):
        return make pipeline(PolynomialFeatures(degree), LinearRegression(**kwargs))
    # никто не задумывался, что **kwargs звучит как боевой клич берсерков? или орков?
    # вобщем где-то из этой серии
    # Поиск по сетке
    param grid = {'polynomialfeatures degree': [1, 2, 3], 'linearregression fit intercept': [True, False]}
    grid_search = GridSearchCV(PolynomialRegression(), param_grid, cv=10, scoring='neg mean squared error')
    grid search.fit(X train dex, y train dex)
Ľ→
              GridSearchCV
```





Подбор глобальных параметров моделей

• Рандомный лес. ВАУ

```
▼ Случайный лес
       # Поиск по сетке
        model_forest = RandomForestRegressor()
        #model_forest = RandomForestRegressor(max_features = 2, n_estimators = 30, random_state = 0)
        grid = {'n_estimators': range(1,200, 50), #'max_features': [2, 3, 4],
                "max_features": ["auto", "sqrt", "log2"],
                'max_depth' : range (1,10,1),
                'criterion' : ['squared_error']
        grid_search = GridSearchCV(model forest, grid, n jobs= -1, cv = 10)
        grid_search.fit(X_train_dex, y_train_dex)
        #print(grid_search.best_params_)
        #print(sqrt(abs(grid_search.best_score_)))
                    GridSearchCV
         ▶ estimator: RandomForestRegressor

    RandomForestRegressor

[163] print(grid_search.best_score_)
        print(grid search.best estimator )
        -0.03661291684191641
        RandomForestRegressor(max depth=1, max features='sqrt', n estimators=51)
```



Сводная таблица ошибок для обученных моделей эпик -феил везде r2 полная грусть и отрицалово

Метод регрессии	Ошибки									
	MAE	MSE	RMSE	r ²						
Линейная регрессия	0.13610	0.02867	0.16935	-0.00698						
Полиномиальная регрессия	0.13610	0.02867	0.16935	-0.00698						
Метод Лассо	0.13601	0.02871	0.16946	-0.00835						
Эластичная сеть	0.13601	0.02871	0.16946	-0.00835						
Метод опорных векторов	0.13551	0.02854	0.16895	-0.00229						
Дерево принятия решений	0.13841	0.02994	0.17303	-0.05122						
Бэггинг	0.13781	0.02933	0.17126	-0.02976						
Случайный лес	0.14047	0.03113	0.17642	-0.09275						



Модель для рекомендации соотношения матрица наполнитель композита

• Нейроночка

```
model ns = Sequential()
model_ns.add(Dense(128, input_shape = (X_train_ns.shape[1], ), activation = 'relu'))
model ns.add(Dense(128, activation='relu'))
model ns.add(Dropout(0.18))
model ns.add(Dense(64, activation='relu'))
model ns.add(Dropout(0.18))
model ns.add(Dense(32, activation='relu'))
model ns.add(Dropout(0.18))
model ns.add(Dense(16, activation='relu'))
model ns.add(BatchNormalization())
model ns.add(Dense(1, activation = 'sigmoid'))
#компиляция
model ns.compile(optimizer = 'adam', loss = 'mean absolute error', metrics = ['mse'])
```





edu.bmstu.ru

+7 495 182-83-85

edu@bmstu.ru

Москва, Госпитальный переулок , д. 4-6, с.3

