



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана

Data Science

Выпускная квалификационная работа

по теме:

«Прогнозирование конечных свойств новых материалов
(композиционных материалов)»

Мещанова А.А.

группа 11784DS



План работы

1 Изучение теоретических основ машинного обучения

2 Разведочный анализ данных и визуализация

3 Предобработка данных

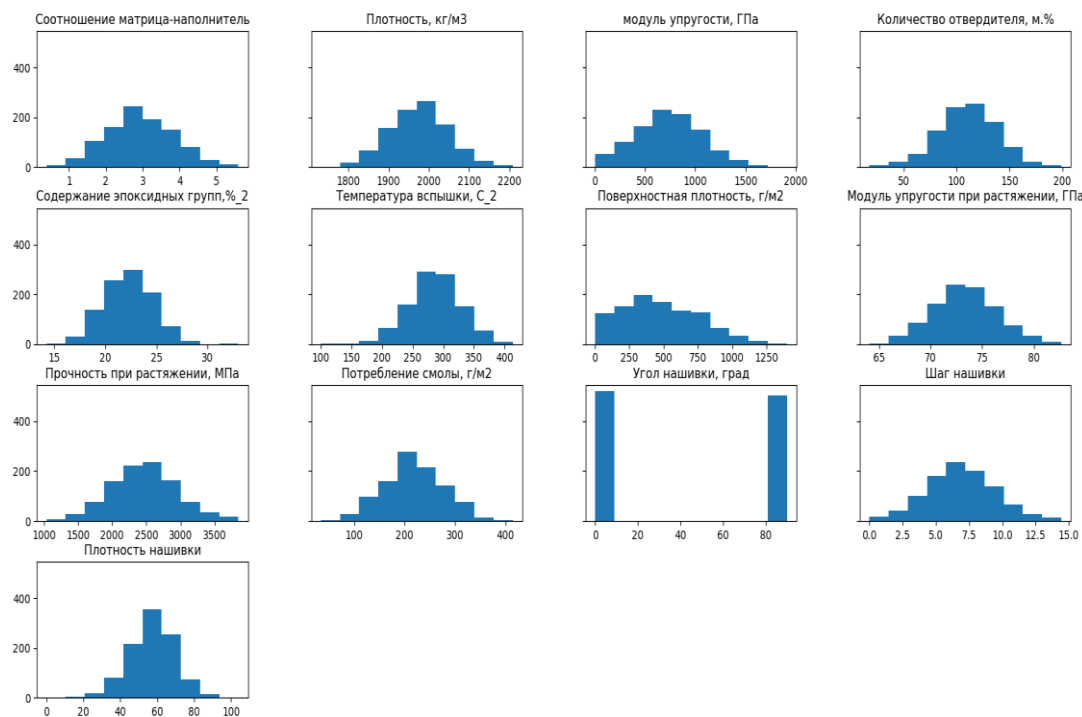
4 Разработка, обучение и оценка различных моделей

5 Создание приложения для выбранной модели



Разведочный анализ данных

Описательная статистика признаков датасета, полученного в результате объединения исходных файлов

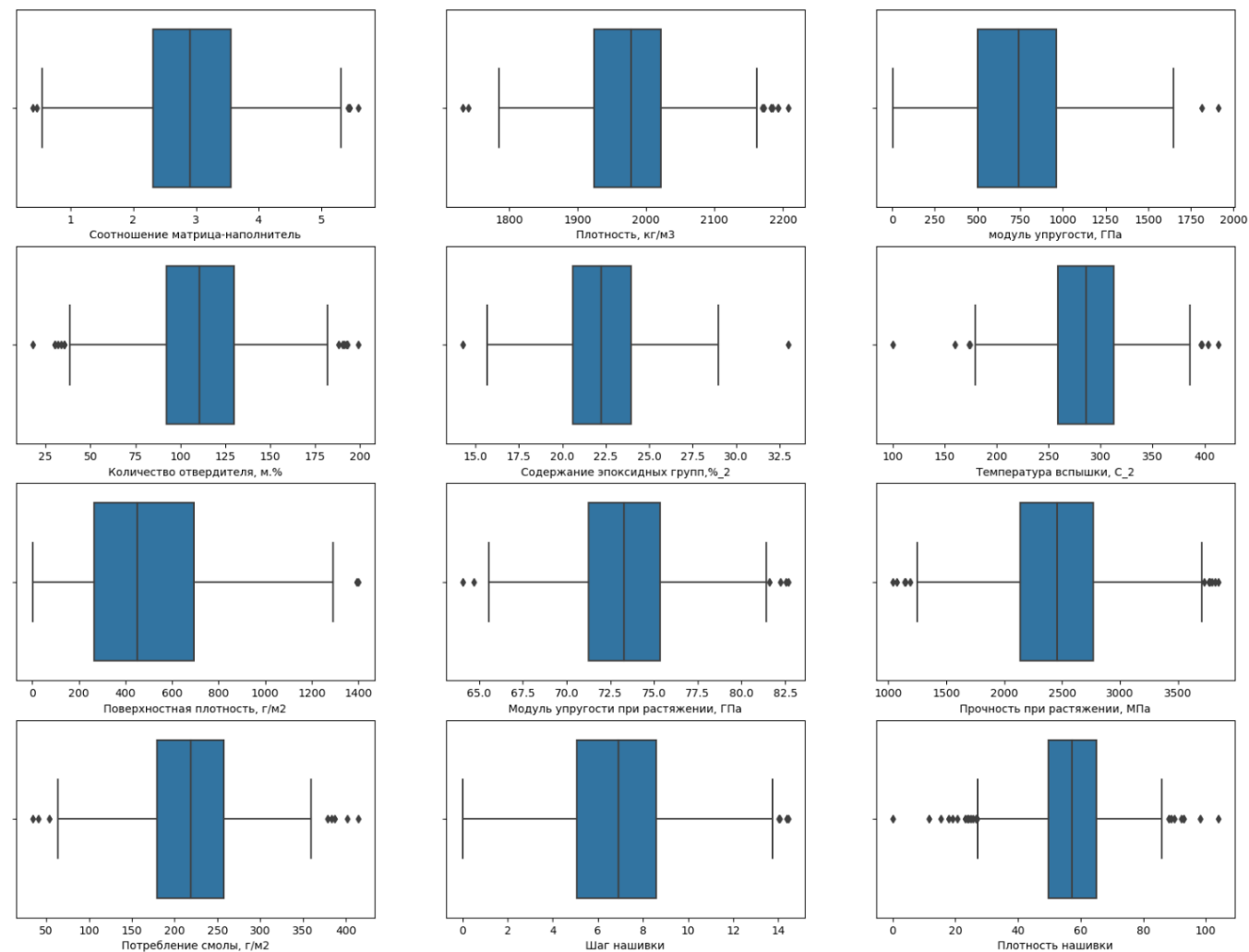
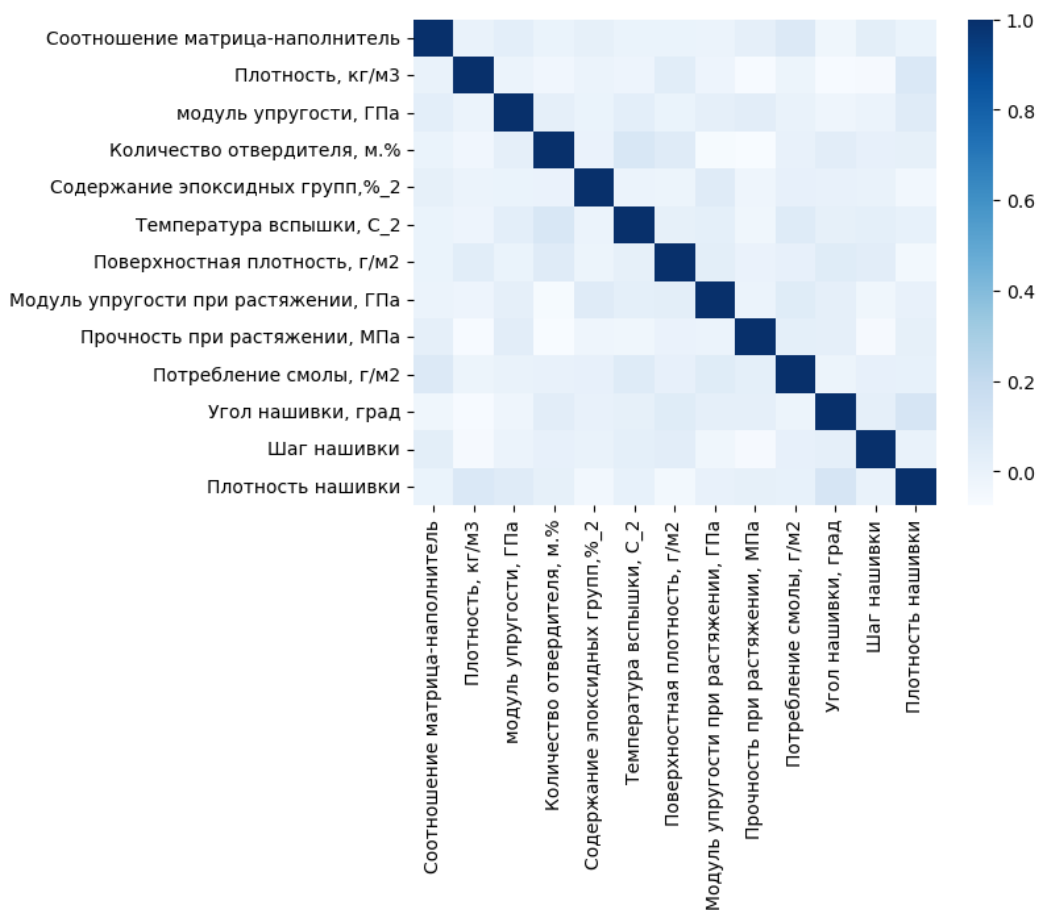


	count	mean	std	min	max
Соотношение матрица-наполнитель	1023.0	2.930366	0.913222	0.389403	5.591742
Плотность, кг/м3	1023.0	1975.734888	73.729231	1731.764635	2207.773481
модуль упругости, ГПа	1023.0	739.923233	330.231581	2.436909	1911.536477
Количество отвердителя, м.%	1023.0	110.570769	28.295911	17.740275	198.953207
Содержание эпоксидных групп, %_2	1023.0	22.244390	2.406301	14.254985	33.000000
Температура вспышки, C_2	1023.0	285.882151	40.943260	100.000000	413.273418
Поверхностная плотность, г/м2	1023.0	482.731833	281.314690	0.603740	1399.542362
Модуль упругости при растяжении, ГПа	1023.0	73.328571	3.118983	64.054061	82.682051
Прочность при растяжении, МПа	1023.0	2466.922843	485.628006	1036.856605	3848.436732
Потребление смолы, г/м2	1023.0	218.423144	59.735931	33.803026	414.590628
Угол нашивки, град	1023.0	44.252199	45.015793	0.000000	90.000000
Шаг нашивки	1023.0	6.899222	2.563467	0.000000	14.440522
Плотность нашивки	1023.0	57.153929	12.350969	0.000000	103.988901



Разведочный анализ данных

Отсутствие видимых корреляций и наличие выбросов

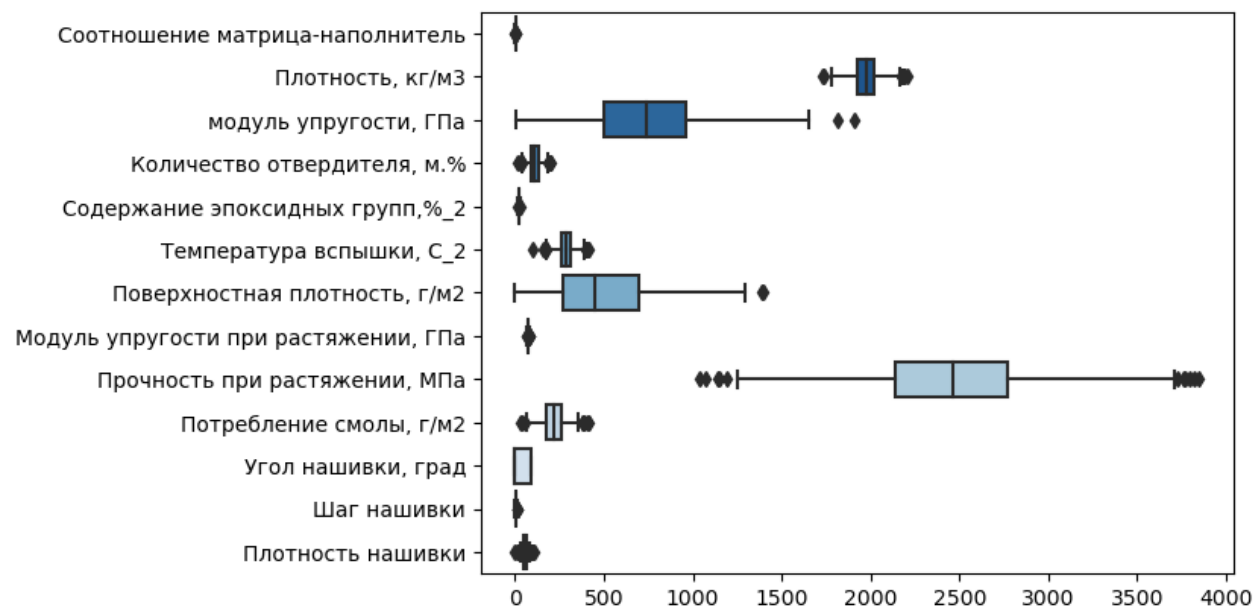




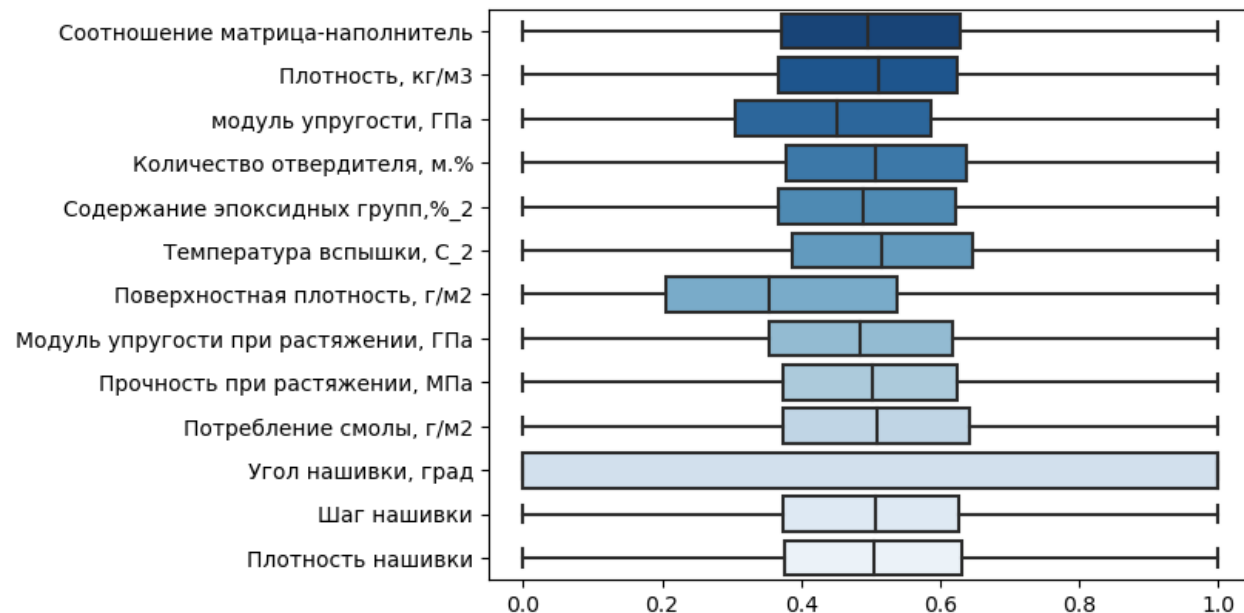
Предобработка данных

Сравнение боксплотов до и после очистки и нормализации данных

До



После





Оценки для моделей предсказания модуля упругости при растяжении

	Модель регрессии	MAE	MSE	R2
0	Linear	2.6109355505901184	10.152935631185066	-0.024476898005094228
1	Ridge	2.5998136023821914	10.084913812910749	-0.017613190412159918
2	Lasso	2.5948108078737024	10.09572202309215	-0.018703787461314736
3	ElasticNet	2.5801929025662833	9.998351572705044	-0.00887867080603466
4	GradientBoosting	2.6998404476642666	11.112730470692213	-0.12132451682377177
5	KNeighbors	2.6225621597024773	10.325154738425082	-0.04185458098976058
6	DecisionTree	2.6342990118377716	10.509499850482133	-0.06045583243299113
7	RandomForest	2.576793337919035	10.024106937432755	-0.011477503017891744
8	AdaBoost	2.6353405208350638	10.432782936591103	-0.052714750560453716
9	NeuralNetwork	2.580495886956309	9.966996628786234	-0.005714815853147925

Лучшие модели:
NeuralNetwork
ElasticNet

Оценки для моделей предсказания прочности при растяжении

	Модель регрессии	MAE	MSE	R2
0	Linear	372.05238450611114	214801.90979576955	-0.003529768950333745
1	Ridge	369.8832593534413	213955.87645116603	0.0004228013393881014
2	Lasso	368.9355592328838	213322.56630912074	0.0033815533408245724
3	ElasticNet	368.8669394882762	214058.58303888745	-5.703200261408803e-05
4	GradientBoosting	382.86326924208987	225302.159467887	-0.05258572537734851
5	KNeighbors	400.581653879789	250632.24504259886	-0.1709249661618053
6	DecisionTree	400.581653879789	250632.24504259886	-0.1709249661618053
7	RandomForest	368.94893692224065	213550.7865727877	0.0023153345689093108
8	AdaBoost	380.4897225346174	224358.12034276538	-0.04817528337516097
9	NeuralNetwork	1213.5428445429625	1686732.611086863	-6.880220381161292

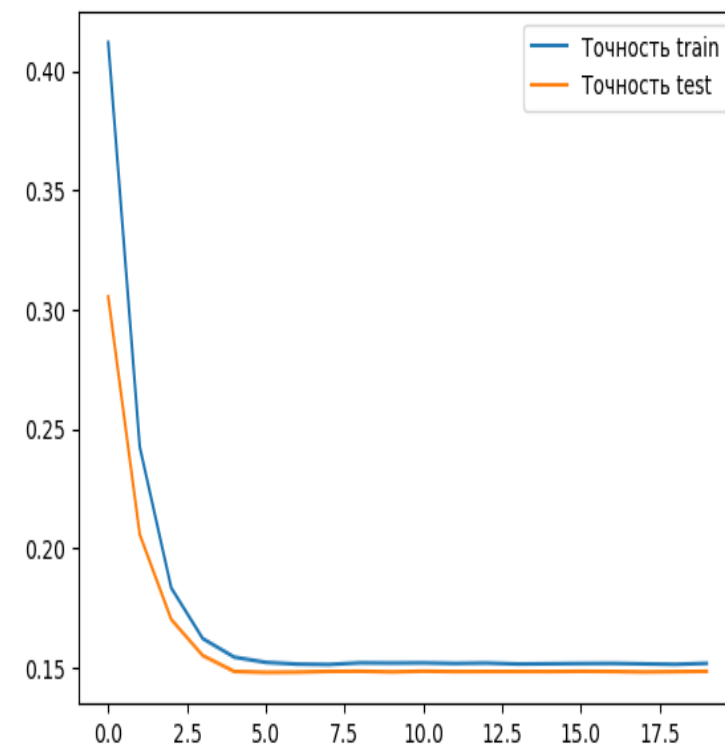
Лучшие модели:
RandomForest
Lasso
Ridge



Оценки моделей искусственной нейронной сети для предсказания соотношения матрица-наполнитель

Визуализация обучения
выбранной модели

	Версия нейросети	MAE	MSE	R2
0	Нейросеть 1	2.386060661491529	6.45774411548042	-7.4474735694538445
1	Нейросеть 2	0.7137162029574031	0.7679923222381272	-0.004622469957904052
2	Нейросеть 3	0.7104158222844613	0.7644645654075289	-7.757443281297682e-06
3	Нейросеть 4	0.7143640262384645	0.775958358616658	-0.015042963640772733
4	Нейросеть 5	0.7116793262948142	0.7653706815770234	-0.001193061824467545
5	Нейросеть 6	0.7131466493755869	0.7667935002491898	-0.003054272630024002
6	Нейросеть 7	1.19195047989152	1.9860870722335884	-1.598030790521507
7	Нейросеть 8	0.7101724567930656	0.7637227176227177	0.0009626649585182667
8	Нейросеть 9	0.711454842236497	0.7661455688434341	-0.0022067036760575753
9	Нейросеть 10	0.7133339033046622	0.7730611124592092	-0.011253031754097309





ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана



do.bmstu.ru