



**ОБРАЗОВАТЕЛЬНЫЙ
ЦЕНТР** МГТУ им. Н. Э. Баумана

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА по курсу «Data Science»

Прогнозирование конечных свойств новых материалов (композиционных материалов)

Докладчик: **Желдак Евгений Михайлович**,
слушатель 2 потока 11107 Образовательного центра МГТУ им. Н. Э. Баумана
под управлением МИЦ «Композиты России»

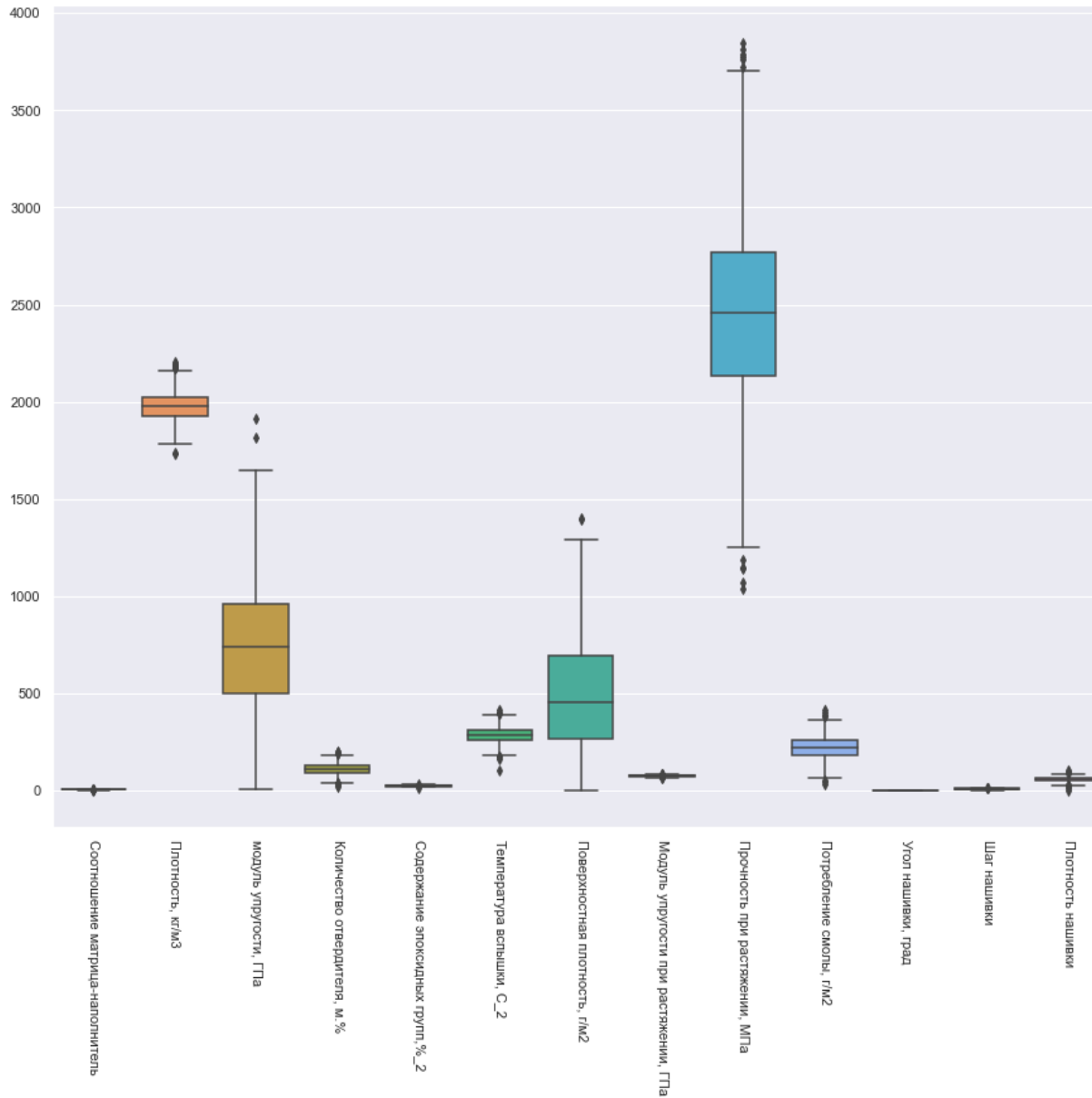
Решаемые задачи и исходные данные

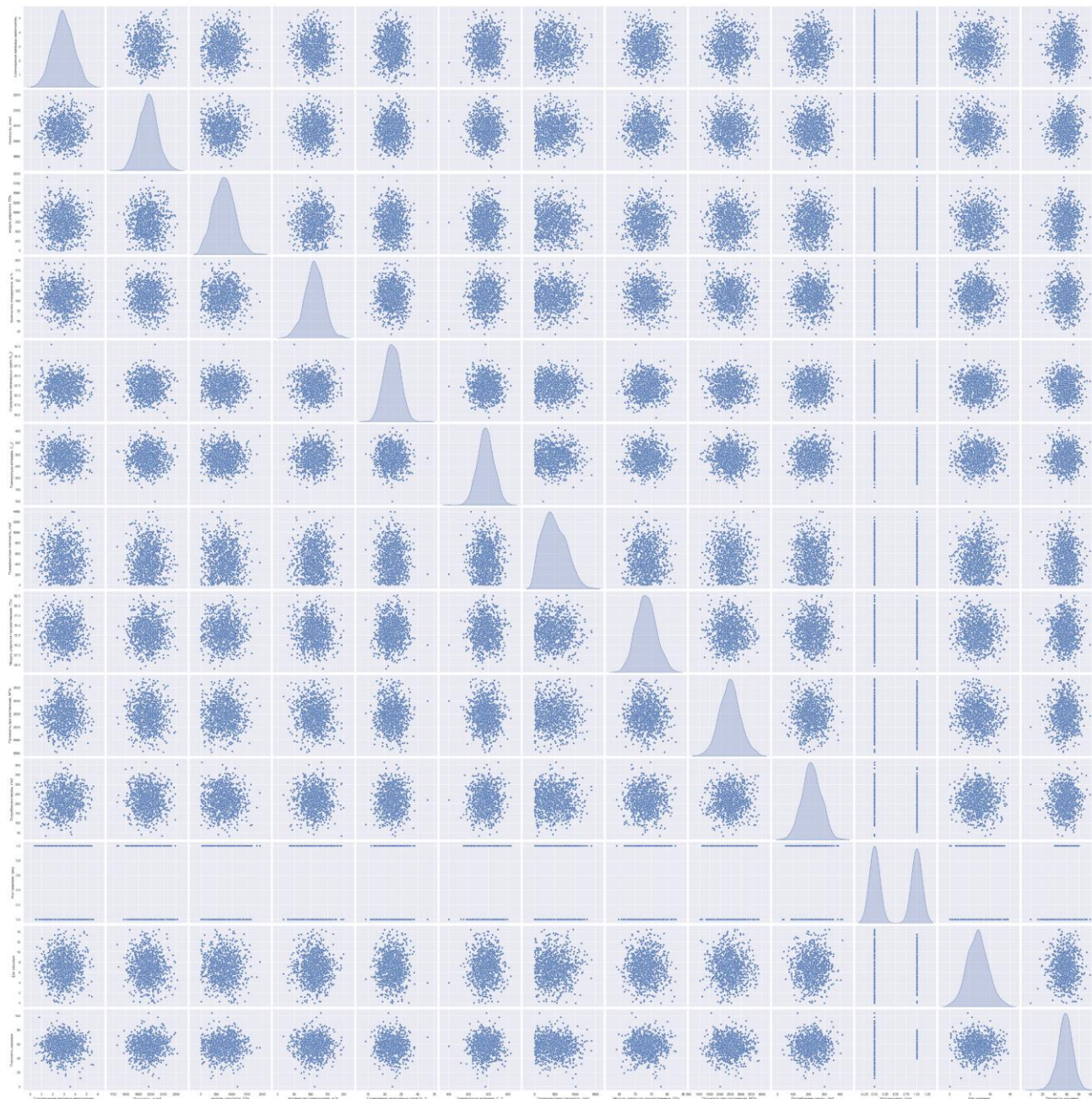
- Разведочный анализ данных
- Выбор регрессионной модели и прогноз значений модуля упругости при растяжении и прочности при растяжении
- Построение нейронной сети для прогноза соотношения матрица-наполнитель
- Разработка web-приложения для прогноза соотношения матрица-наполнитель

Исходные данные:

- 13 признаков по 1023 значения
- 1 категориальный признак
- Пропуски отсутствуют
- Распределение данных близко к нормальному
- Линейной зависимости между признаками нет

Данные до предобработки



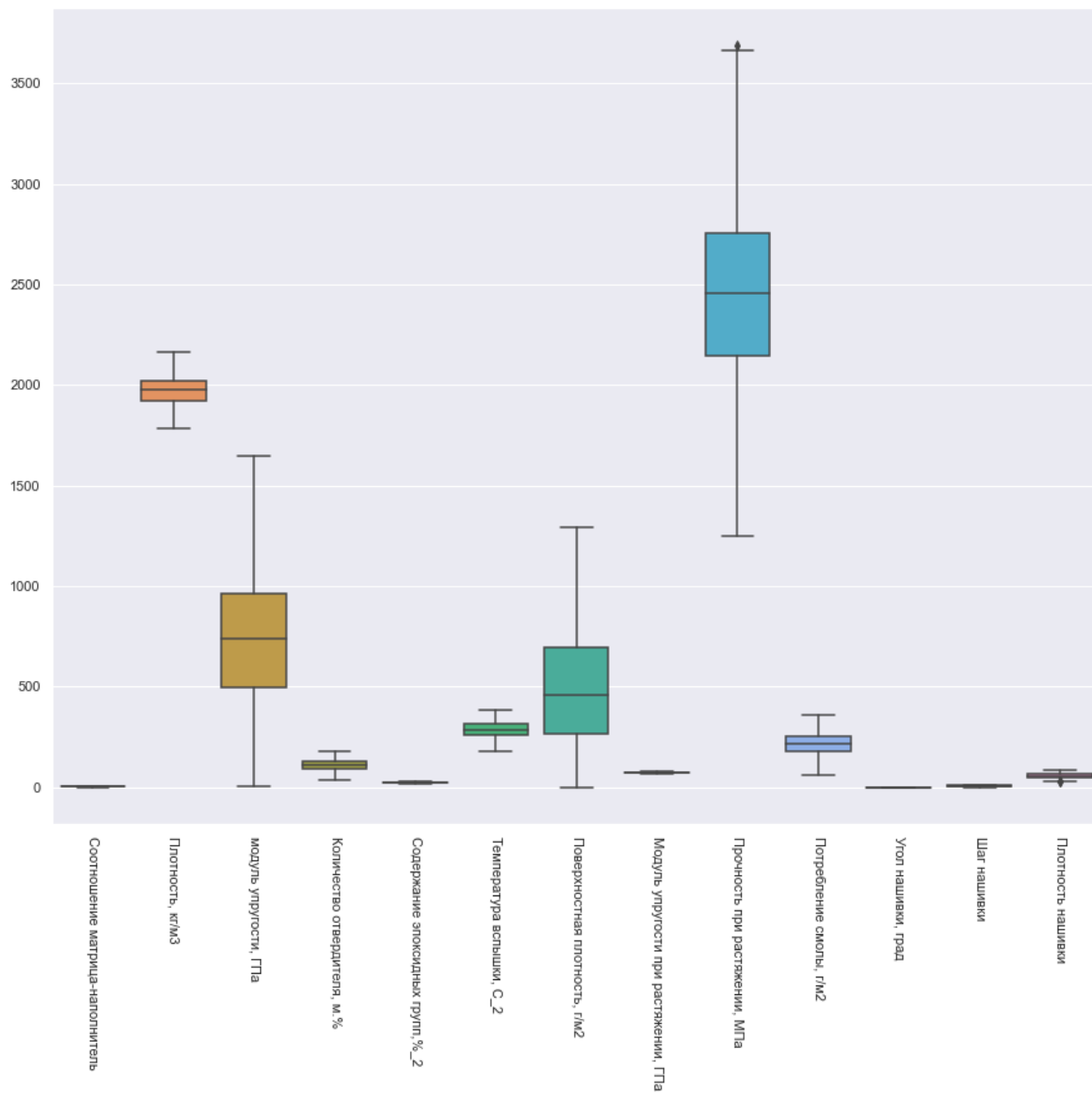


Попарные графики до
предобработки

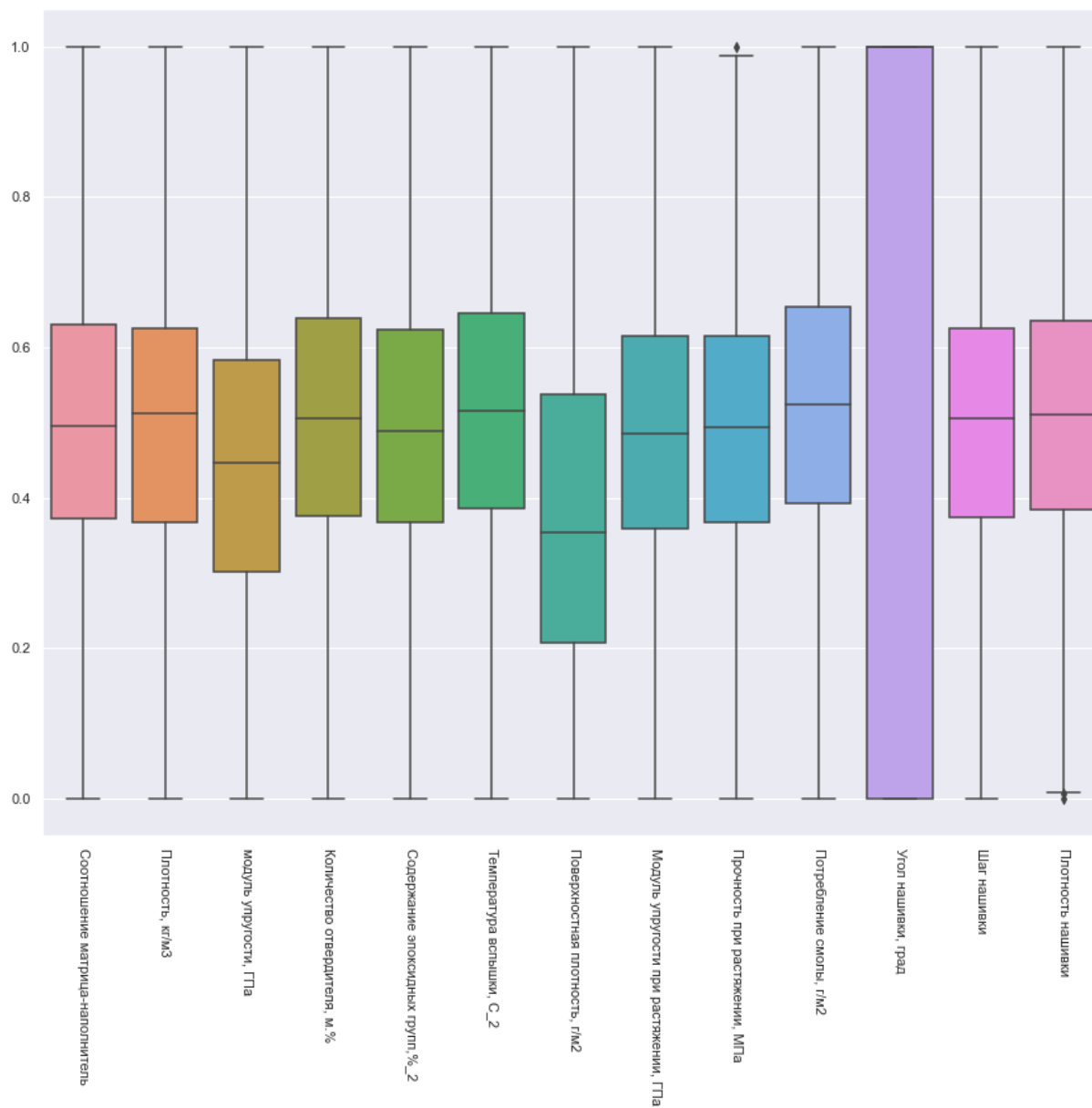


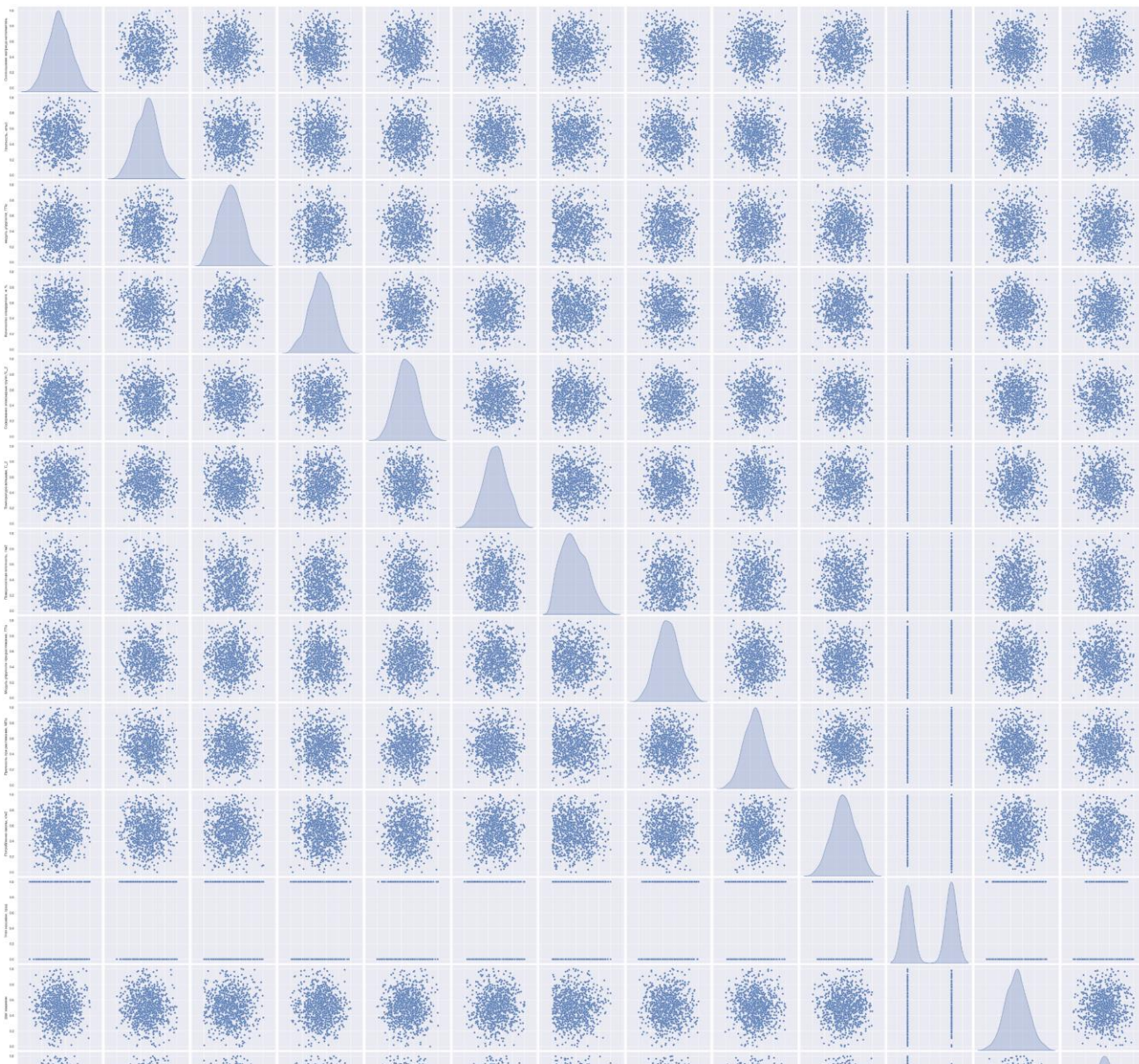
**ОБРАЗОВАТЕЛЬНЫЙ
ЦЕНТР** МГТУ им. Н. З. Баумана

Данные после удаления выбросов



Данные после нормализации

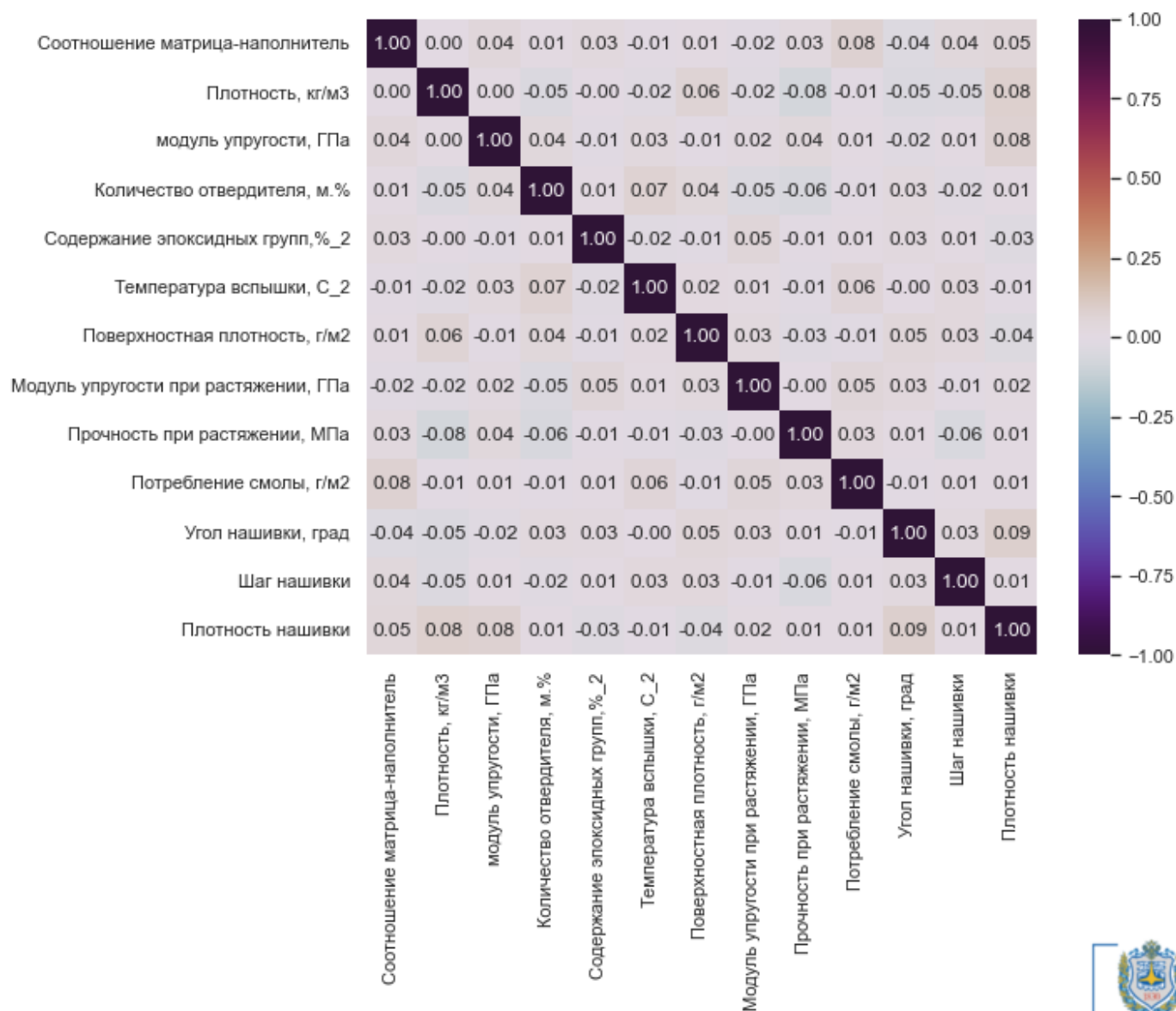




**Попарные графики
после удаления
выбросов и
нормализации**



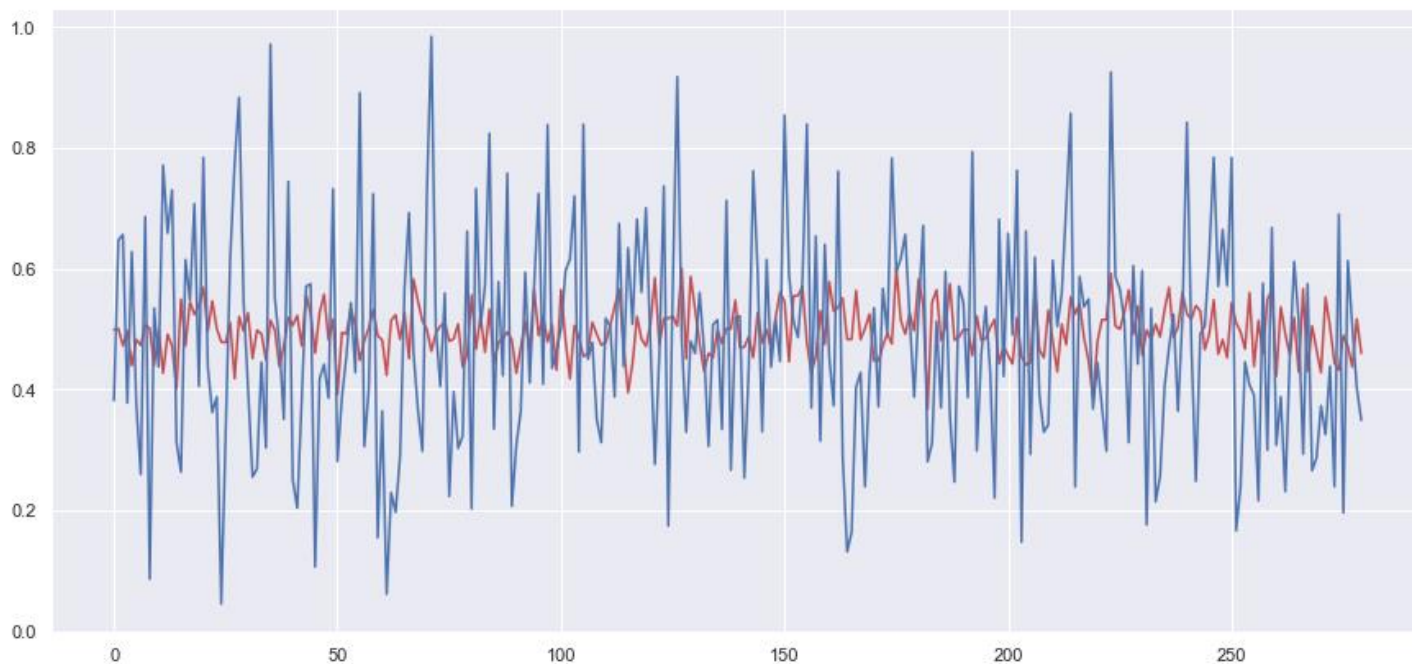
Корреляционная матрица



Разработка и обучение моделей для прогноза модуля упругости при растяжении

Модель	RMSE	r2
Linear	0.185817	-0.021807
Polynomial	0.195041	-0.125768
Spline	0.201448	-0.200949
Random Forest	0.187003	-0.034893
Multilayer Perceptron	0.204796	-0.241197

«Лучшая» модель –
случайный лес



Разработка и обучение моделей для прогноза прочности при растяжении

Модель	RMSE	r2
Linear	0.188691	0.004901
Polynomial	0.193815	-0.049882
Spline	0.191584	-0.025843
Random Forest	0.186723	0.025550
Multilayer Perceptron	0.217275	-0.319415

«Лучшая» модель –
случайный лес

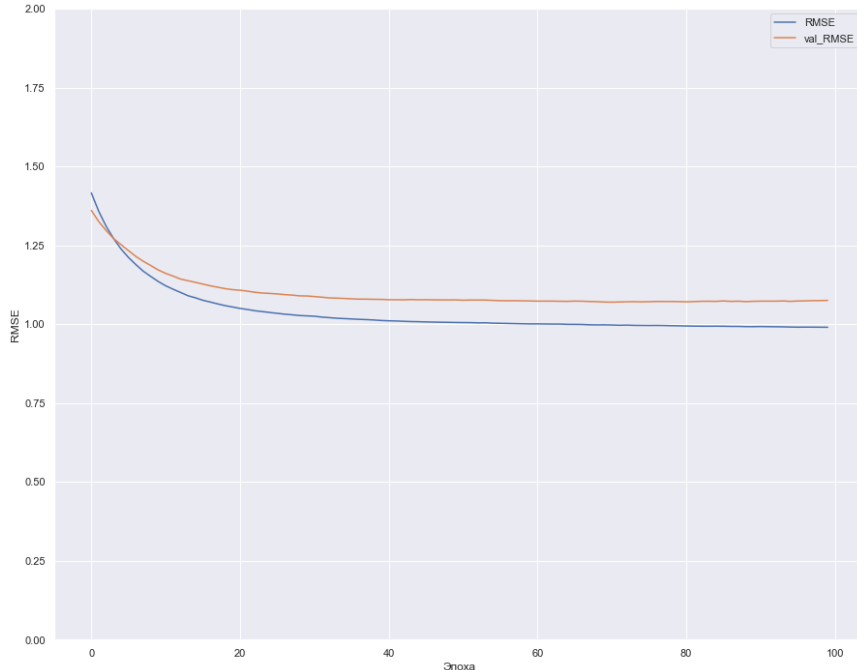


Создание нейронной сети для рекомендации соотношения матрица-наполнитель

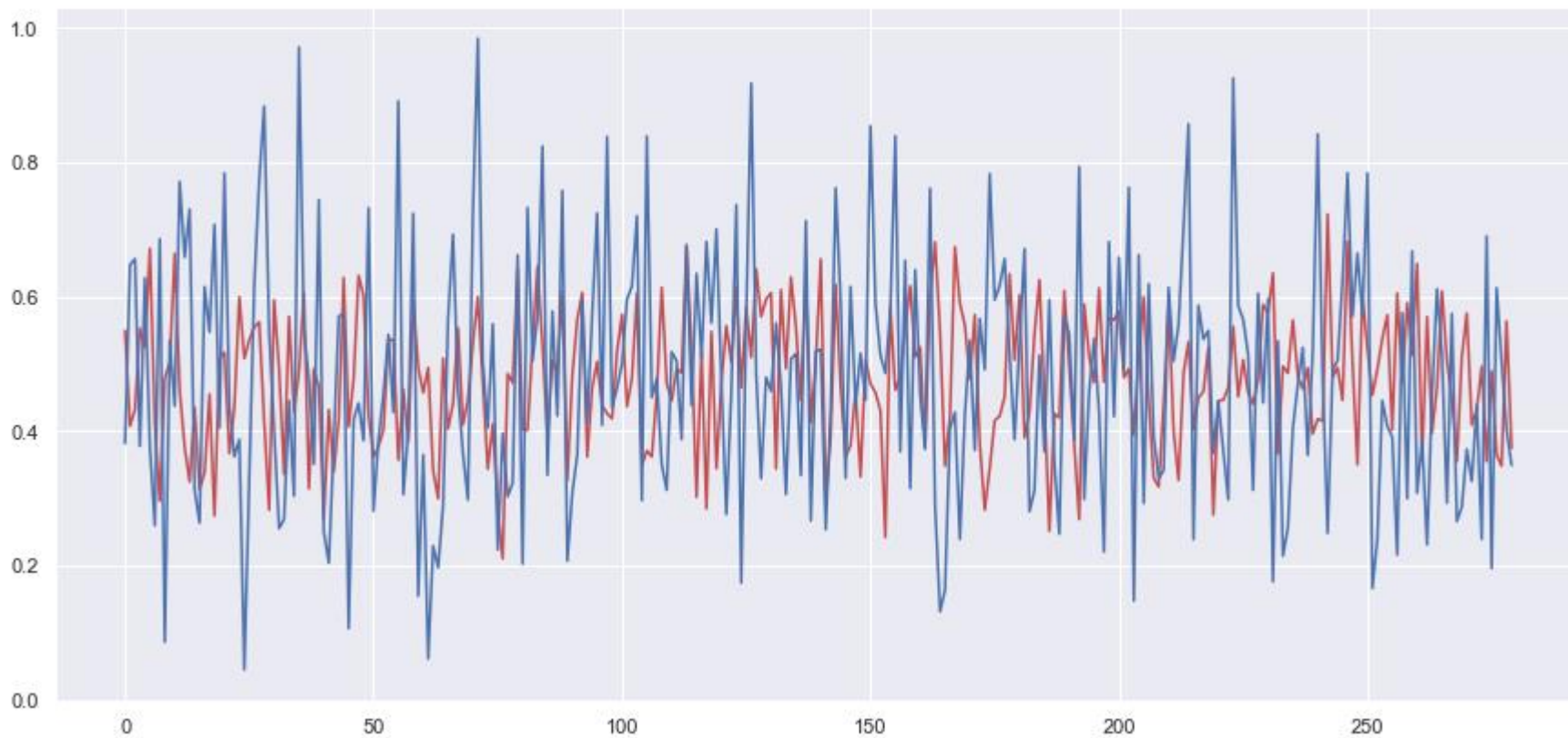
```
normalizer = tf.keras.layers.Normalization(axis=-1)
normalizer.adapt(np.array(X))
def build_and_compile_model(norm):
    model = keras.Sequential([
        norm,
        layers.Dense(6, activation='relu'),
        layers.Dense(1)])
    model.compile(loss='mean_absolute_error',

optimizer=tf.keras.optimizers.Adam(0.001),

metrics=[tf.keras.metrics.RootMeanSquaredError()])
    return model
```



Сравнение прогнозных и проверочных данных



Разработка web-приложения

← → ↻ 127.0.0.1:5000/nn_model/

Прогнозирование соотношения матрица-наполнитель

Плотность, кг/м3 (1700...2300)

Модуль упругости, ГПа (2...2000)

Количество отвердителя, м.% (17...200)

Содержание эпоксидных групп,%_2 (14...34)

Температура вспышки, С_2 (100...414)

Поверхностная плотность, г/м2 (0.6...1400)

Модуль упругости при растяжении, ГПа (64...83)

Прочность при растяжении, МПа (1036...3849)

Потребление смолы, г/м2 (33...414)

Угол нашивки, град (0 или 90)

Шаг нашивки (0...15)

Плотность нашивки (0...104)

Входные переменные:

	Соотношение матрица-наполнитель	Плотность, кг/м3	модуль упругости, ГПа	Количество отвердителя, м.%	Содержание эпоксидных групп,%_2	Температура вспышки, С_2	П
0	0	1996.159145	525.057774	77.506883	18.126107	223.408685	28.6

Результат модели:

Соотношение матрица-наполнитель
2.902847183234564

- Используем сохранённые нормализатор и нейронную сеть
- Для начала работы нужно ввести значения всех свойств.
- Нажать на кнопку «Submit Query».
- Отобразятся входные переменные и соответствующее им прогнозное значение соотношения матрица-наполнитель

Выводы

- коэффициенты корреляции между признаками стремятся к нулю;
- примененные модели линейной и полиномиальной регрессии, случайный лес и нейронная сеть не показали высокой эффективности в прогнозировании свойств композитов, необходимы дополнительные вводные данные для улучшения моделей;
- лучшие метрики – критерии качества у модели случайного леса.

СПАСИБО ЗА ВНИМАНИЕ!

