

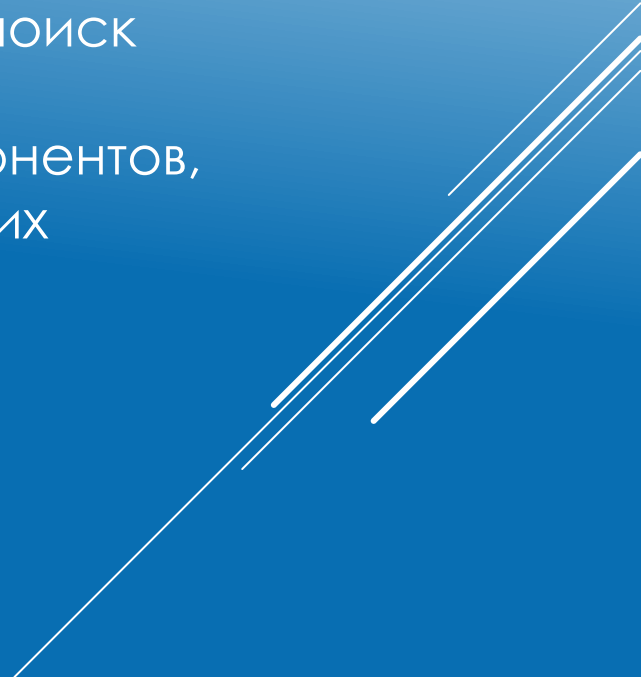
Выпускная квалификационная работа по курсу
«Data Science»

ПРОГНОЗИРОВАНИЕ КОНЕЧНЫХ СВОЙСТВ
НОВЫХ МАТЕРИАЛОВ
(КОМПОЗИЦИОННЫХ МАТЕРИАЛОВ)

Орлянский А.Л.

ЦЕЛИ РАБОТЫ:

Задача, исследуемая в данной работе, направлена на поиск способов прогнозирования свойств композиционного материала путем оценки свойств входящих в него компонентов, а также их структуры, соотношения и применяемых при их совмещении технологических процессов.

Several white lines of varying lengths and orientations are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

ИСХОДНЫЕ ДАННЫЕ

В качестве исходных данных для решения задачи представлено два массива данных, содержащие информацию о следующих характеристиках материалов, применяемых при изготовлении композитного материала, а также их структуры и взаимного соотношения:

- соотношение матрица-наполнитель
- плотность
- модуль упругости
- количество отвердителя
- содержание эпоксидных групп
- температура вспышки
- поверхностная плотность
- модуль упругости при растяжении
- прочность при растяжении
- потребление смолы
- угол нашивки
- шаг нашивки
- плотность нашивки

На основании имеющихся исходных данных необходимо научиться прогнозировать следующие конечные свойства композиционного материала в зависимости от имеющихся характеристик:

- модуль упругости при растяжении
- модуль прочности при растяжении
- соотношение матрица-наполнитель

Первый датасет имеет 1023 записи, второй – 1040, данные не имеют пропусков. По условиям задачи датасеты объединяются в один датасет по индексу с типом объединения INNER.

В связи с незначительностью потери данных при их объединении датасеты объединены в один с количеством записей 1023.

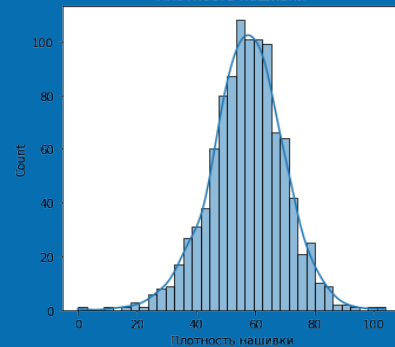
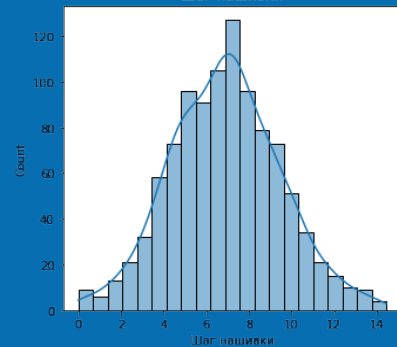
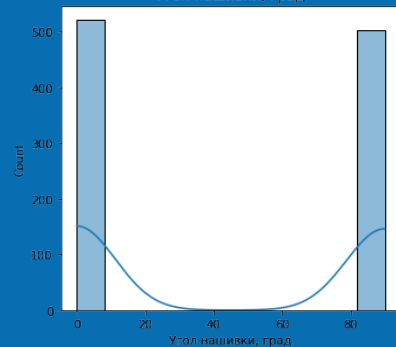
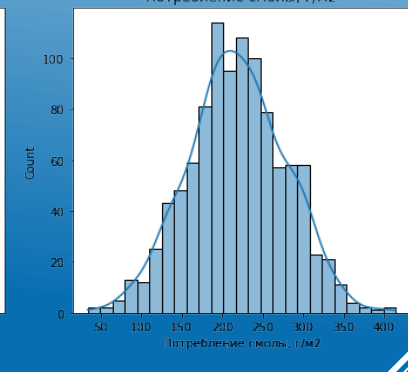
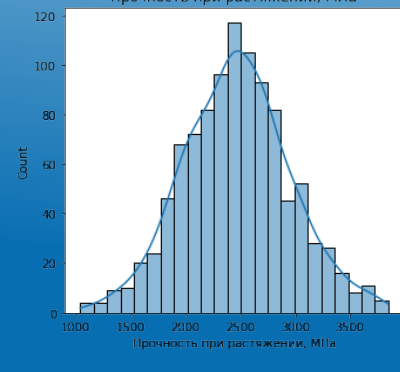
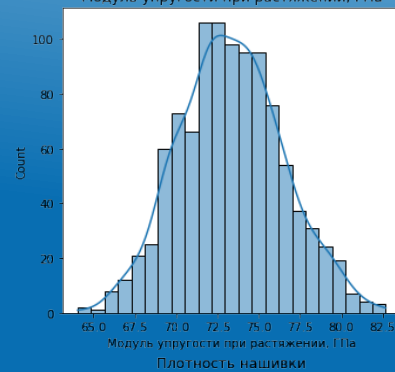
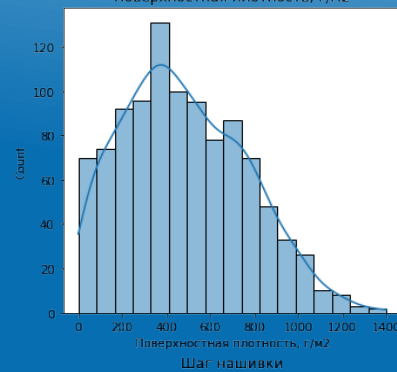
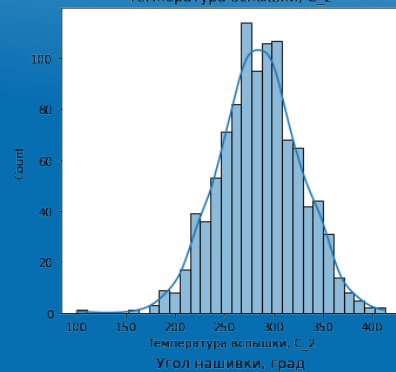
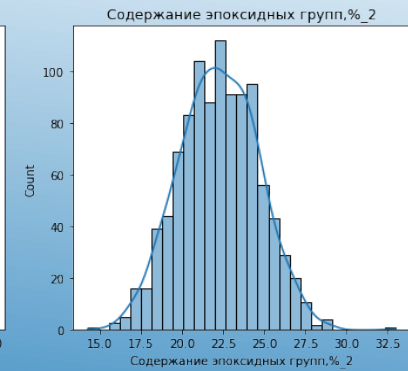
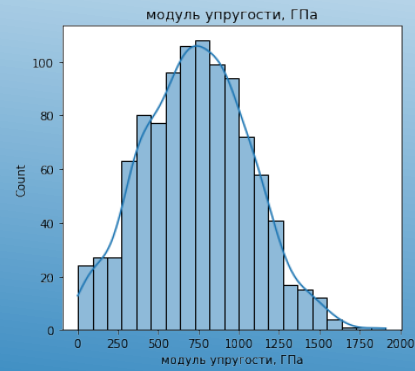
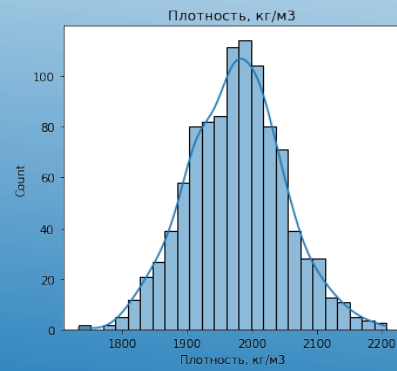
РАЗВЕДОЧНЫЙ АНАЛИЗ ДАННЫХ

В процессе разведочного анализа данных было выявлено, что данные в объединенном датасете не имеют пропусков, но имеют большой разброс значений характеристик.

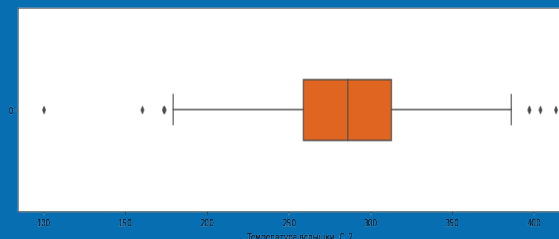
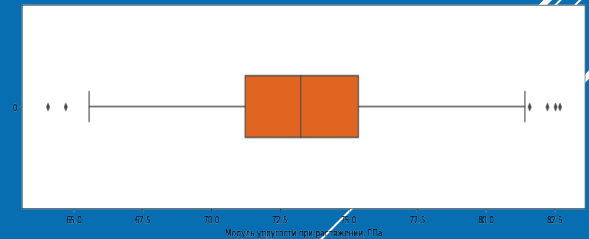
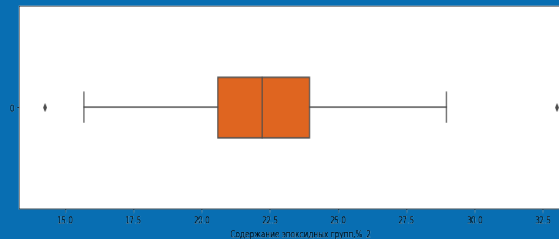
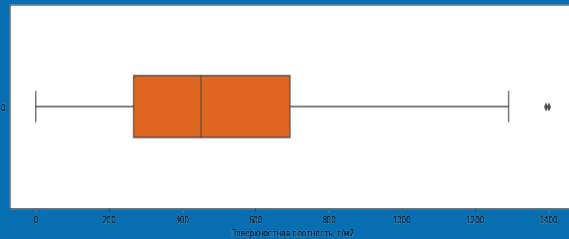
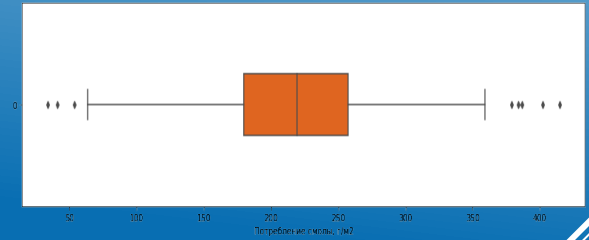
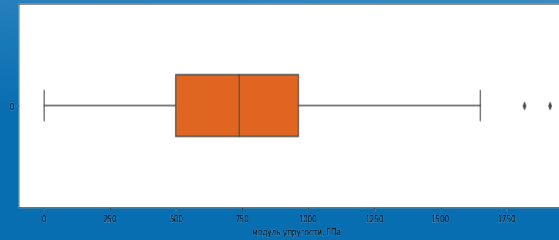
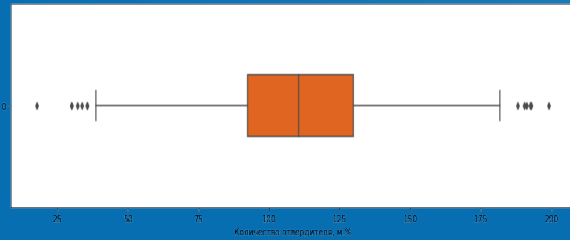
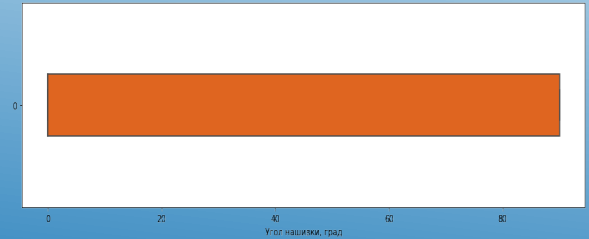
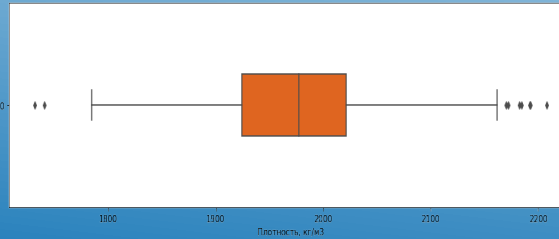
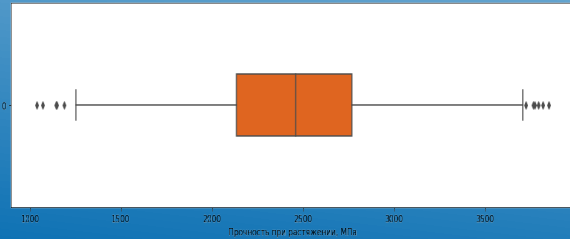
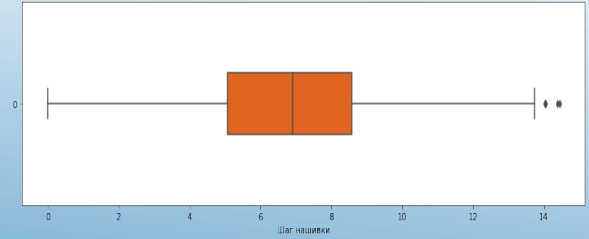
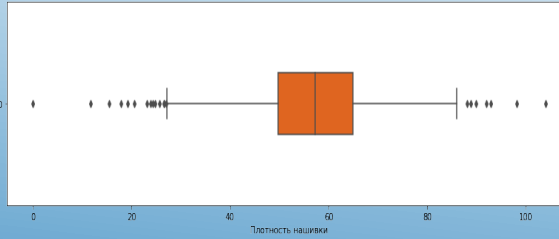
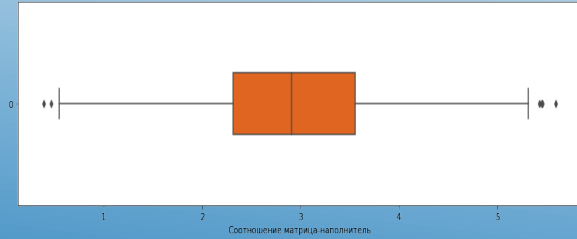
Также получены средние и медианные значения для характеристик материалов, имеющих с датасете.

#Для каждой колонки получаем среднее значение# ds_X.mean(axis=0)		[16] #Для каждой колонки получаем медианное значение# ds_X.median(axis=0)	
Соотношение матрица-наполнитель	2.930366	Соотношение матрица-наполнитель	2.906878
Плотность, кг/м3	1975.734888	Плотность, кг/м3	1977.621657
модуль упругости, ГПа	739.923233	модуль упругости, ГПа	739.664328
Количество отвердителя, м.%	110.570769	Количество отвердителя, м.%	110.564840
Содержание эпоксидных групп,%_2	22.244390	Содержание эпоксидных групп,%_2	22.230744
Температура вспышки, C_2	285.882151	Температура вспышки, C_2	285.896812
Поверхностная плотность, г/м2	482.731833	Поверхностная плотность, г/м2	451.864365
Модуль упругости при растяжении, ГПа	73.328571	Модуль упругости при растяжении, ГПа	73.268805
Прочность при растяжении, МПа	2466.922843	Прочность при растяжении, МПа	2459.524526
Потребление смолы, г/м2	218.423144	Потребление смолы, г/м2	219.198882
Угол нашивки, град	44.252199	Угол нашивки, град	0.000000
Шаг нашивки	6.899222	Шаг нашивки	6.916144
Плотность нашивки	57.153929	Плотность нашивки	57.341920
dtype: float64		dtype: float64	

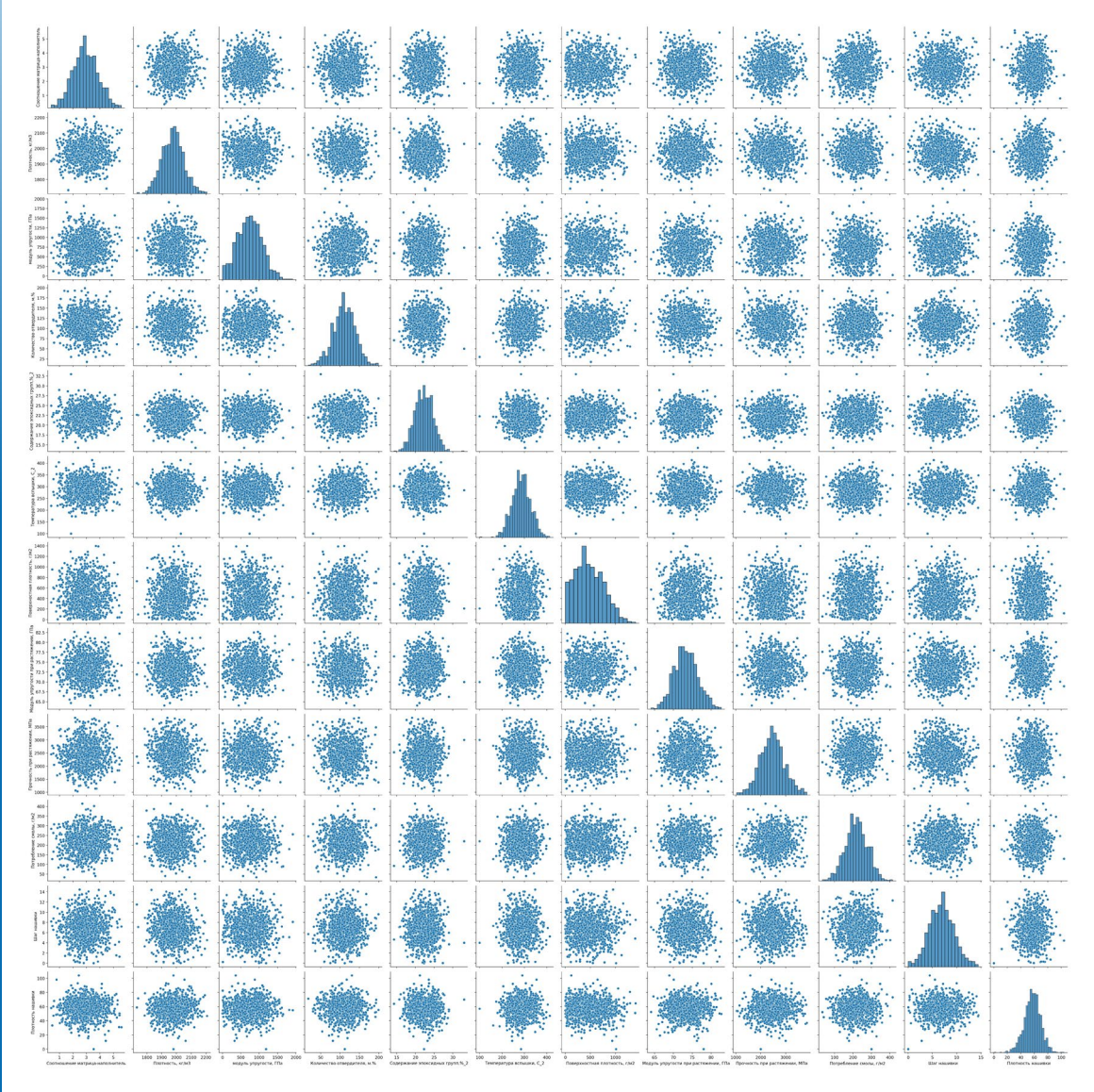
РАСПРЕДЕЛЕНИЕ ХАРАКТЕРИСТИК В ДАТАСЕТЕ



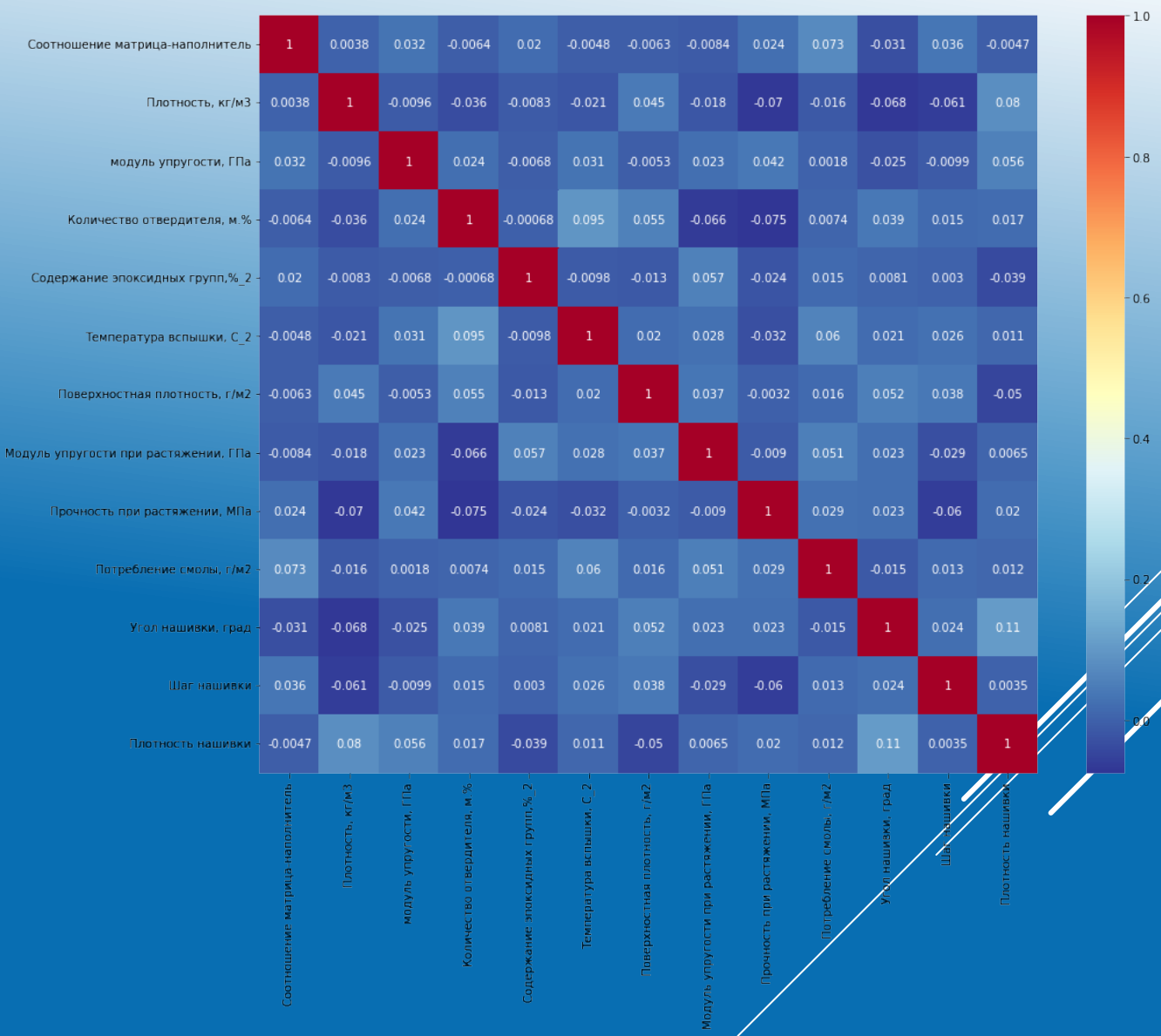
НАХОЖДЕНИЕ ВЫБРОСОВ В ДАТАСЕТЕ



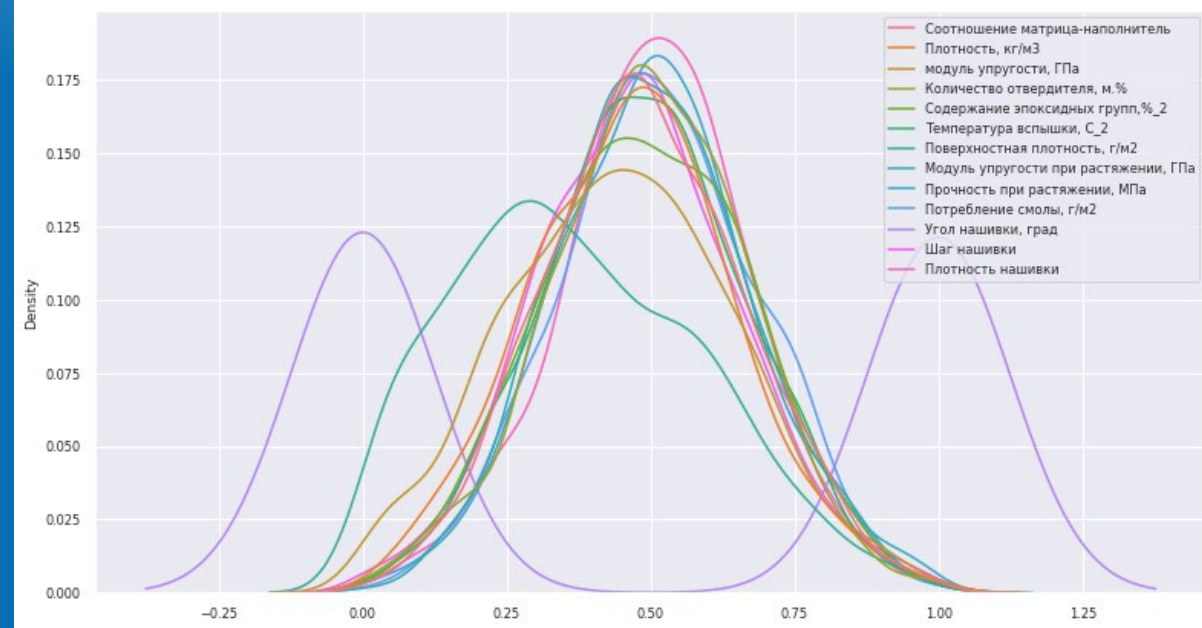
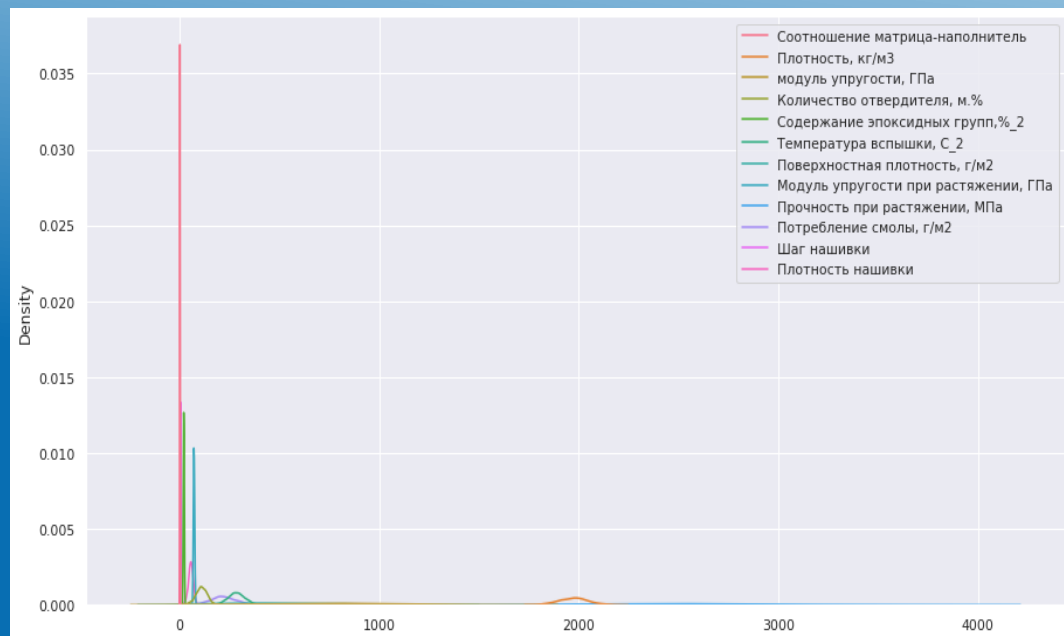
ПОПАРНЫЕ ГРАФИКИ РАССЕЯНИЯ ДЛЯ ЗНАЧЕНИЯ ХАРАКТЕРИСТИК В ДАТАСЕТЕ



ТЕПЛОВАЯ КАРТА ДЛЯ
ИССЛЕДОВАНИЯ
ЗАВИСИМОСТИ МЕЖДУ
ЗНАЧЕНИЯМИ
ХАРАКТЕРИСТИК

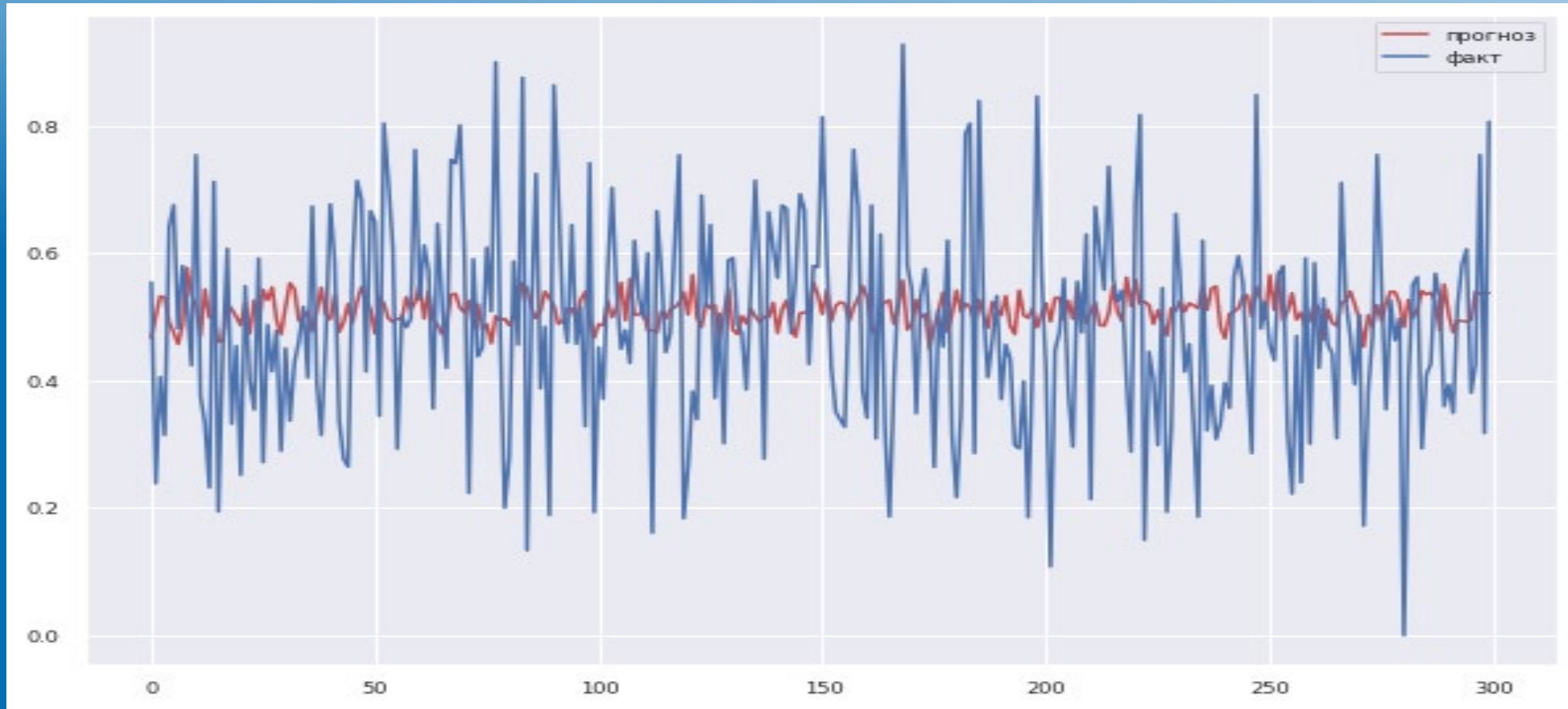


ПРЕДОБРАБОТКА ДАННЫХ



РАЗРАБОТКА И ОБУЧЕНИЕ МОДЕЛЕЙ

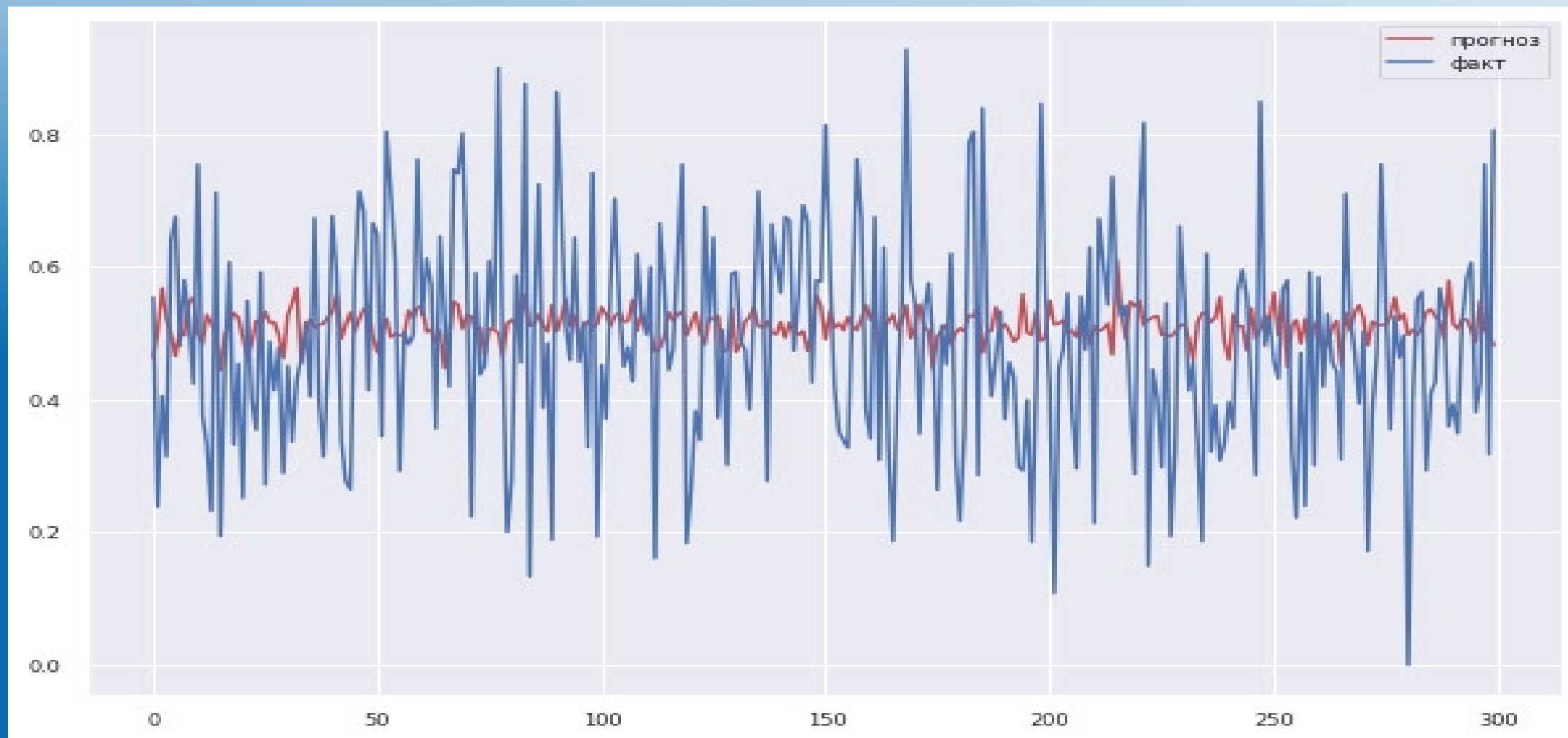
Для получения прогнозов требуемых характеристик были использованы линейная регрессия, метод к-ближайших соседей и случайный лес.



Оценка модели линейной регрессии для прогноза модуля упругости



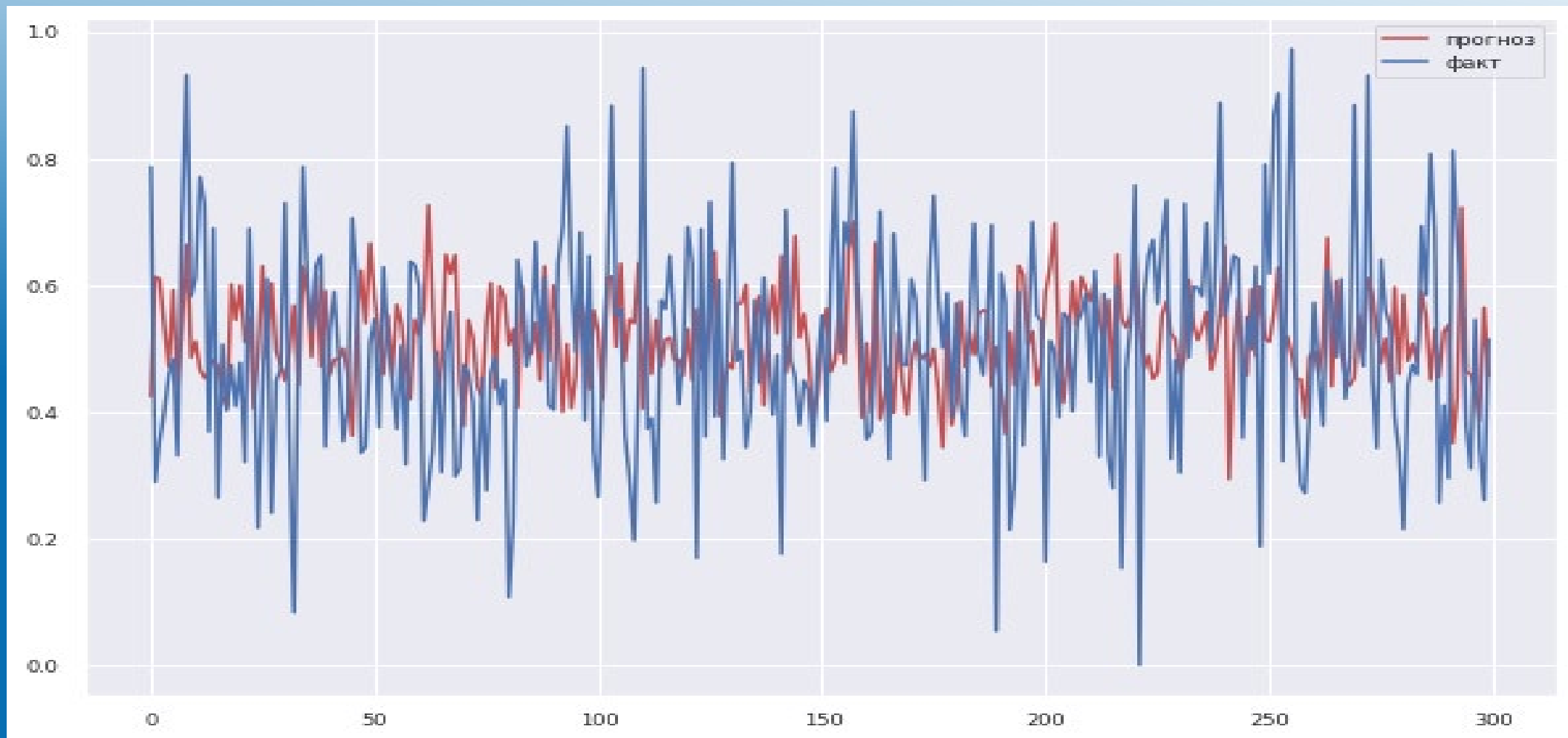
Оценка модели к-ближайших соседей для прогноза модуля упругости



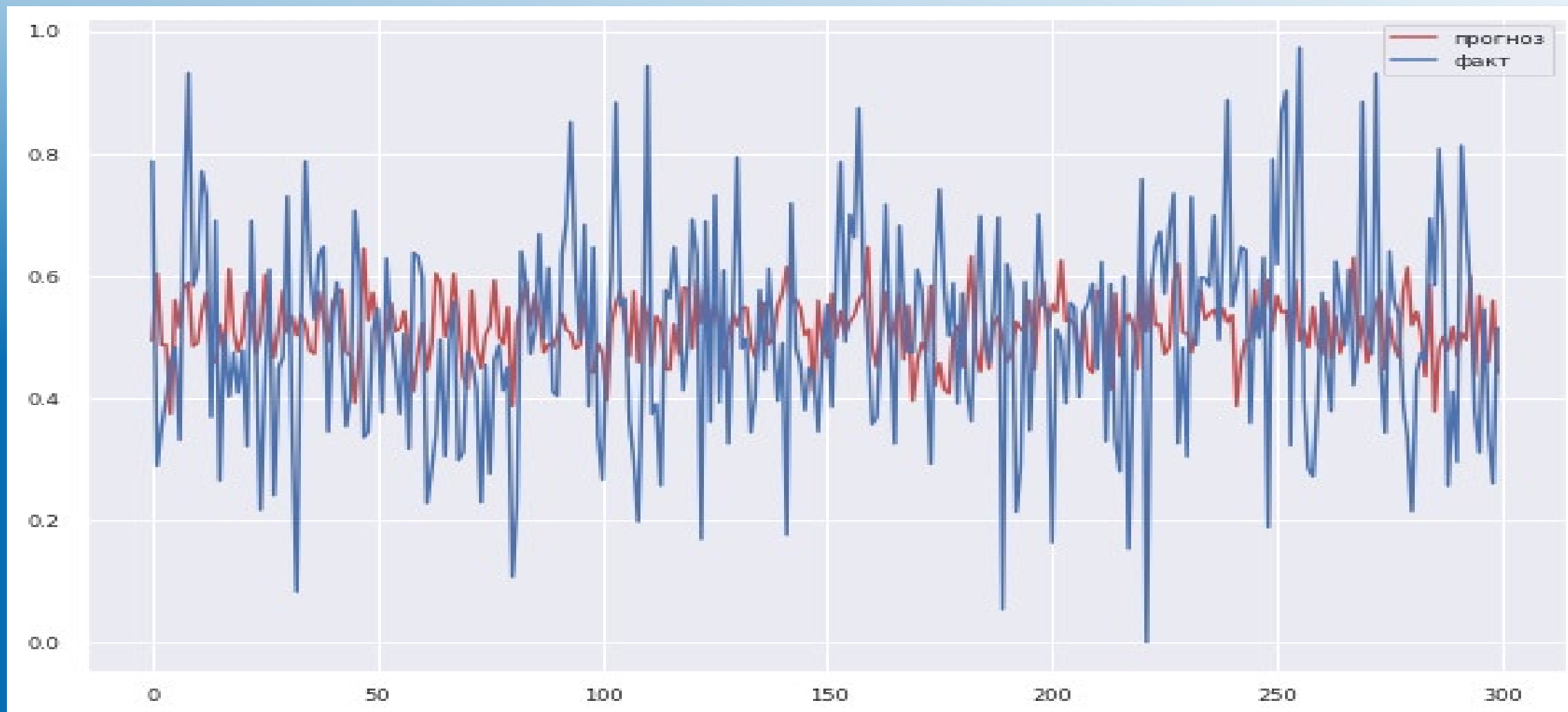
Оценка модели случайный лес для прогноза модуля упругости



Оценка модели линейной регрессии для прогноза прочности при растяжении



Оценка модели к-ближайших соседей для прогноза прочности при растяжении



Оценка модели случайный лес для прогноза прочности при растяжении