



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

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА по курсу «Data Science Pro» тема: «Прогнозирование конечных свойств новых материалов (композиционных материалов)»

Виноградов Максим Александрович

Постановка задачи

- Исходные данные: 2 датасета с общим количеством параметров 13
- Задачи:
 - - разведочный анализ данных
 - - предобработка данных
 - - обучить регрессионные модели для прогноза модуля упругости при растяжении и прочности при растяжении
 - - обучить нейронную сеть для рекомендации соотношения «матрица-наполнитель»

Разведочный анализ

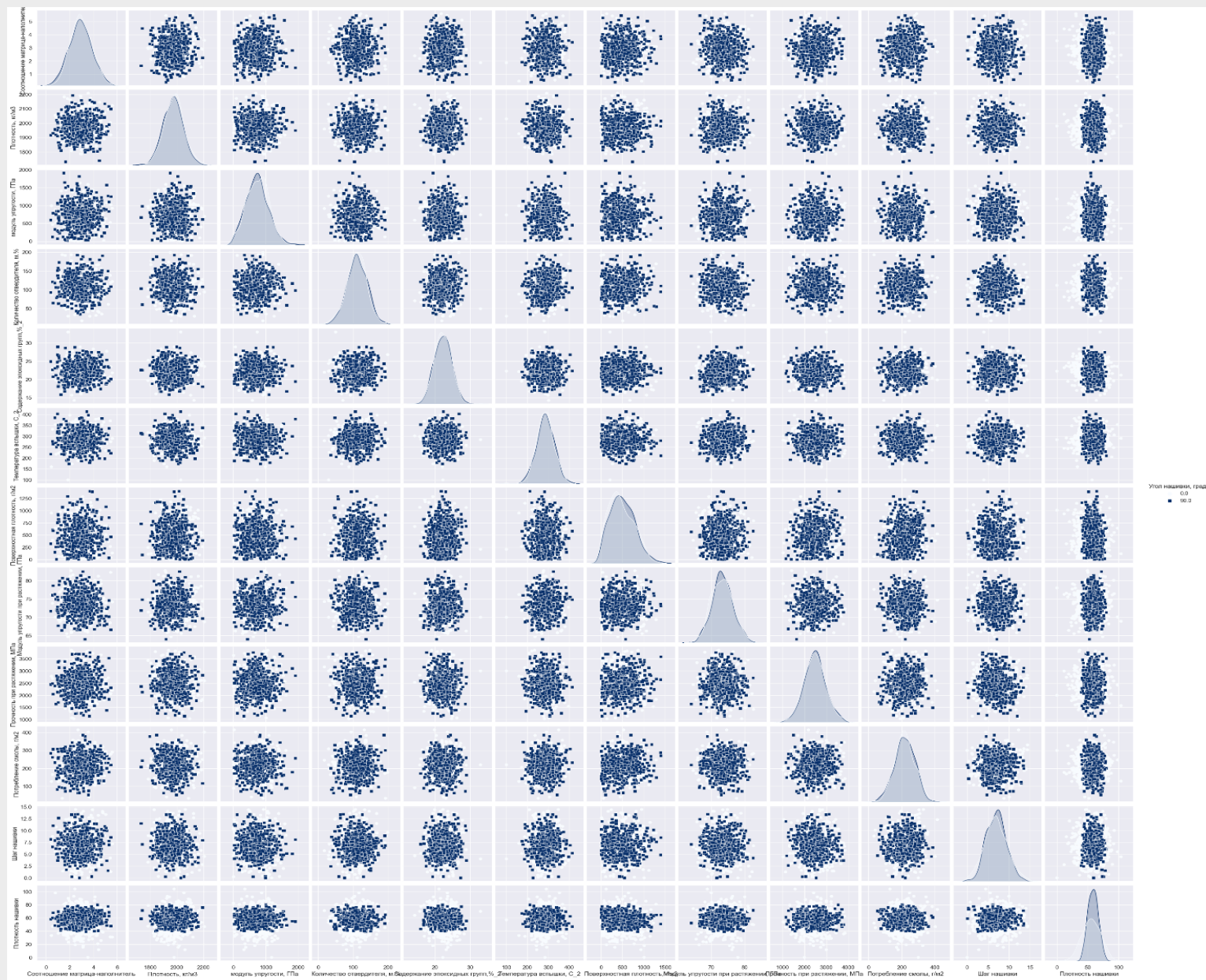
Описательная статистика датасета

	count	mean	std	min	25%	50%	75%	max
Соотношение матрица-наполнитель	1023.0	2.930366	0.913222	0.389403	2.317887	2.906878	3.552660	5.591742
Плотность, кг/м3	1023.0	1975.734888	73.729231	1731.764635	1924.155467	1977.621657	2021.374375	2207.773481
модуль упругости, ГПа	1023.0	739.923233	330.231581	2.436909	500.047452	739.664328	961.812526	1911.536477
Количество отвердителя, м.%	1023.0	110.570769	28.295911	17.740275	92.443497	110.564840	129.730366	198.953207
Содержание эпоксидных групп,%_2	1023.0	22.244390	2.406301	14.254985	20.608034	22.230744	23.961934	33.000000
Температура вспышки, C_2	1023.0	285.882151	40.943260	100.000000	259.066528	285.896812	313.002106	413.273418
Поверхностная плотность, г/м2	1023.0	482.731833	281.314690	0.603740	266.816645	451.864365	693.225017	1399.542362
Модуль упругости при растяжении, ГПа	1023.0	73.328571	3.118983	64.054061	71.245018	73.268805	75.356612	82.682051
Прочность при растяжении, МПа	1023.0	2466.922843	485.628006	1036.856605	2135.850448	2459.524526	2767.193119	3848.436732
Потребление смолы, г/м2	1023.0	218.423144	59.735931	33.803026	179.627520	219.198882	257.481724	414.590628
Угол нашивки, град	1023.0	44.252199	45.015793	0.000000	0.000000	0.000000	90.000000	90.000000
Шаг нашивки	1023.0	6.899222	2.563467	0.000000	5.080033	6.916144	8.586293	14.440522
Плотность нашивки	1023.0	57.153929	12.350969	0.000000	49.799212	57.341920	64.944961	103.988901

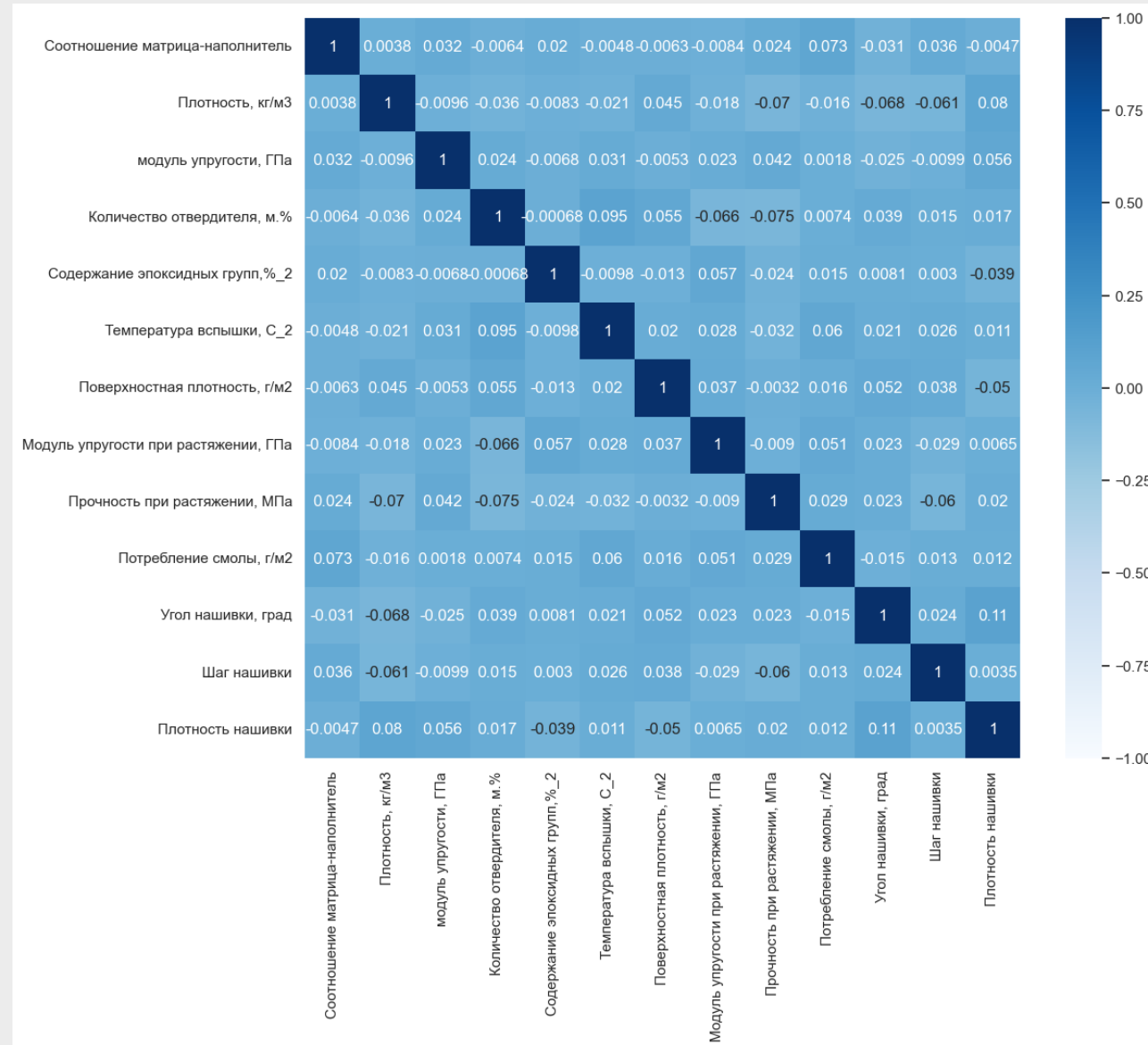
Гистограммы распределения признаков



Парные диаграммы рассеяния



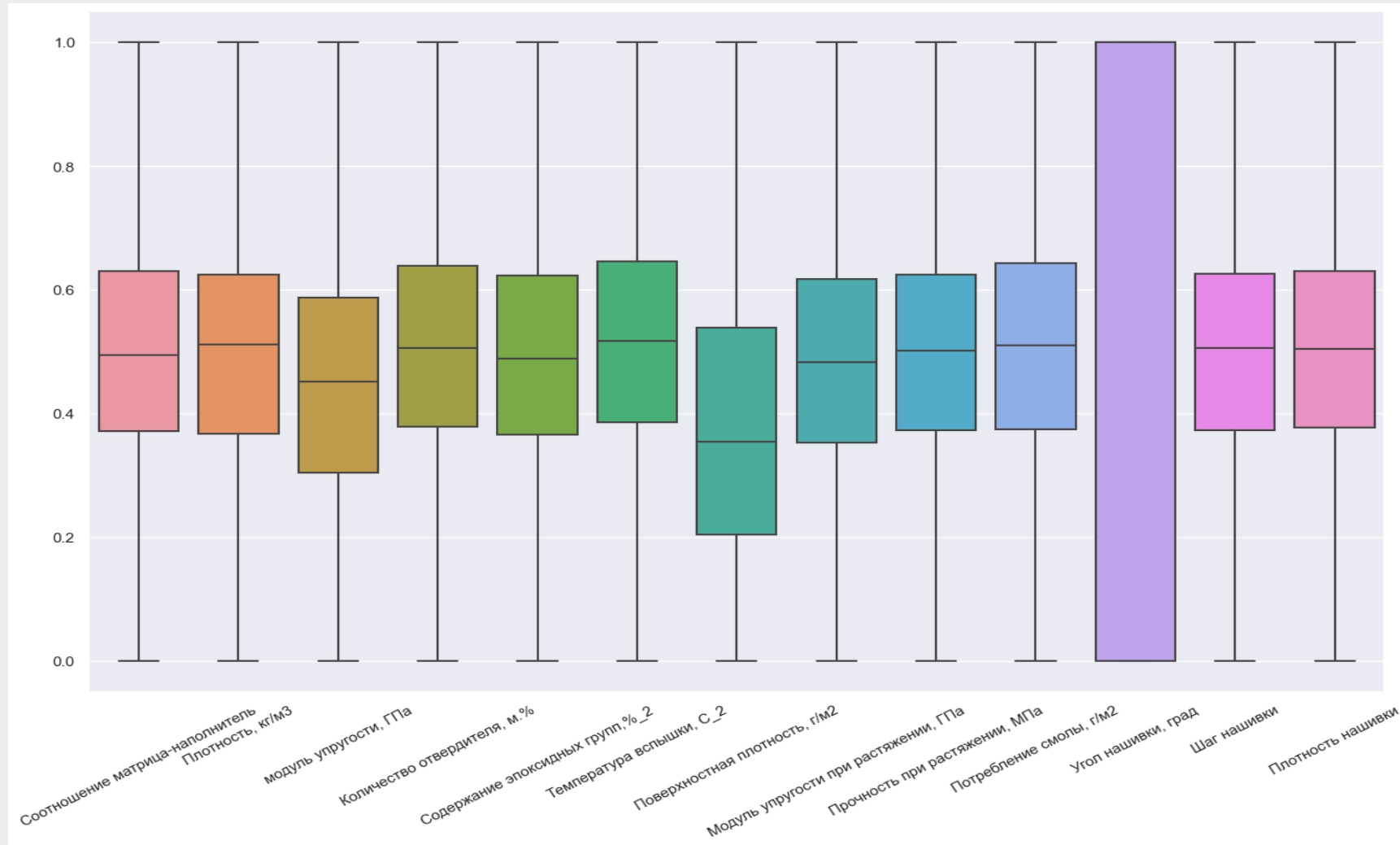
Тепловая карта коэффициентов корреляции



Описательная статистика очищенного нормализованного датасета

	count	mean	std	min	25%	50%	75%	max
Соотношение матрица-наполнитель	936.0	0.498933	0.187489	0.0	0.372274	0.494538	0.629204	1.0
Плотность, кг/м3	936.0	0.502695	0.187779	0.0	0.368517	0.511229	0.624999	1.0
модуль упругости, ГПа	936.0	0.446764	0.199583	0.0	0.301243	0.447061	0.580446	1.0
Количество отвердителя, м.%	936.0	0.504664	0.188865	0.0	0.376190	0.506040	0.637978	1.0
Содержание эпоксидных групп,%_2	936.0	0.491216	0.180620	0.0	0.367716	0.489382	0.623410	1.0
Температура вспышки, С_2	936.0	0.516059	0.190624	0.0	0.386128	0.515980	0.646450	1.0
Поверхностная плотность, г/м2	936.0	0.373733	0.217078	0.0	0.205619	0.354161	0.538683	1.0
Модуль упругости при растяжении, ГПа	936.0	0.488647	0.191466	0.0	0.359024	0.485754	0.615077	1.0
Прочность при растяжении, МПа	936.0	0.495706	0.188915	0.0	0.365149	0.491825	0.612874	1.0
Потребление смолы, г/м2	936.0	0.521141	0.195781	0.0	0.392067	0.523766	0.652447	1.0
Угол нашивки, град	936.0	0.511752	0.500129	0.0	0.000000	1.000000	1.000000	1.0
Шаг нашивки	936.0	0.502232	0.183258	0.0	0.372211	0.504258	0.624604	1.0
Плотность нашивки	936.0	0.513776	0.191342	0.0	0.390482	0.516029	0.638842	1.0

Диграммы размаха (ящики с усами) для нормализованных величин



Сравнение результатов обучения моделей регрессии

Модуль упругости

	Модель	MAE	MSE
0	Линейная регрессия	0.161937	0.039564
1	Случайный лес	0.161887	0.039612
2	k-ближайших	0.161696	0.039254
3	Метод опорных векторов	0.160573	0.039016
4	Градиентный бустинг	0.161973	0.039844

Прочность при растяжении

	Модель	MAE	MSE
0	Линейная регрессия	0.149937	0.034523
1	Случайный лес	0.150998	0.034852
2	k-ближайших	0.152226	0.035300
3	Метод опорных векторов	0.151112	0.034843
4	Градиентный бустинг	0.152481	0.035355

Архитектура нейронной сети

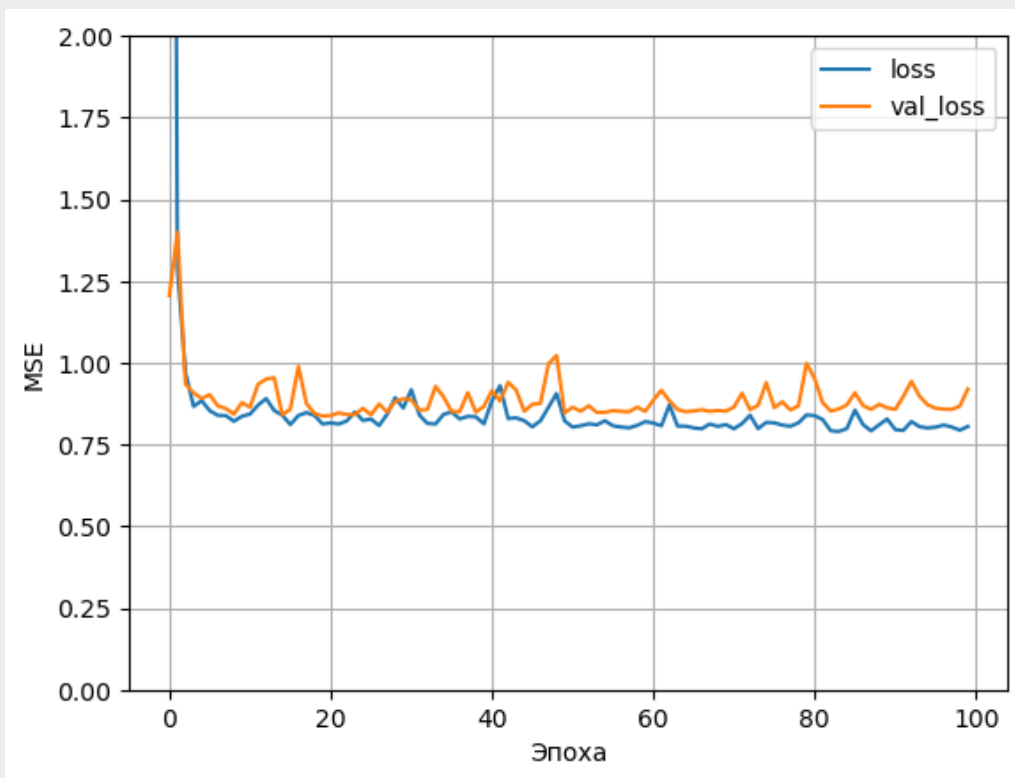
Model: "sequential"

Layer (type)	Output Shape	Param #
normalization (Normalization)	(None, 12)	3
dense (Dense)	(None, 128)	1664
dense_1 (Dense)	(None, 128)	16512
dense_2 (Dense)	(None, 128)	16512
dense_3 (Dense)	(None, 64)	8256
dense_4 (Dense)	(None, 64)	4160
dense_5 (Dense)	(None, 32)	2080
dense_6 (Dense)	(None, 16)	528
dense_7 (Dense)	(None, 1)	17

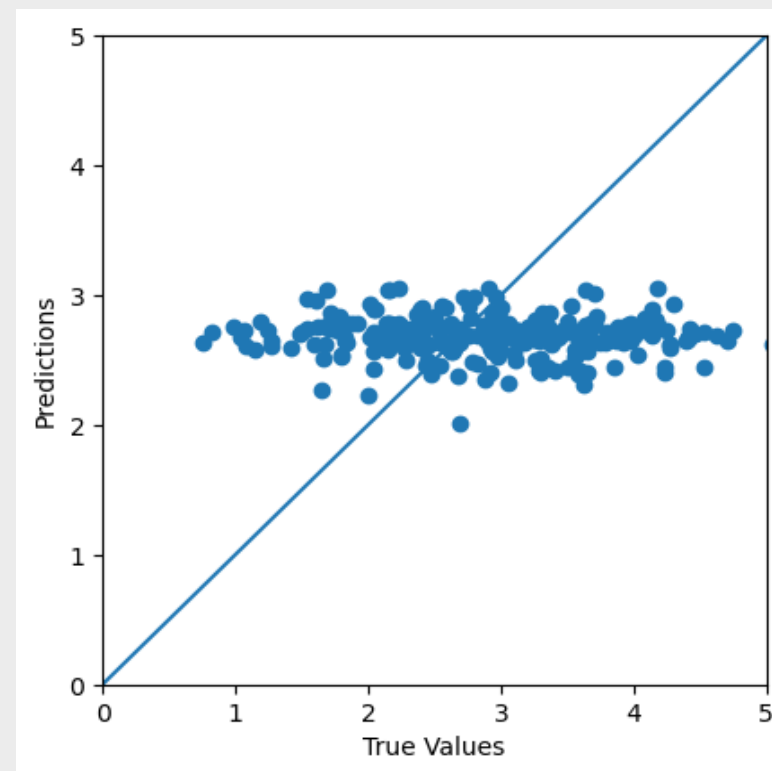
=====
Total params: 49732 (194.27 KB)
Trainable params: 49729 (194.25 KB)
Non-trainable params: 3 (16.00 Byte)

Результаты обучения нейронной сети

MSE во время обучения



Предсказания модели



Интерфейс web-приложения (flask)

Плотность, кг/м3	<input type="text" value="1975,735"/>
Модуль упругости, ГПа	<input type="text" value="739,923"/>
Количество отвердителя, м. %	<input type="text" value="110,571"/>
Содержание эпоксидных групп, %_2	<input type="text" value="22,244"/>
Температура вспышки, С_2	<input type="text" value="285,882"/>
Поверхностная плотность, г/м2	<input type="text" value="482,732"/>
Модуль упругости при растяжении, ГПа	<input type="text" value="73,329"/>
Прочность при растяжении, МПа	<input type="text" value="2461,491"/>
Потребление смолы, г/м2	<input type="text" value="218,423"/>
Угол нашивки, град	<input type="text" value="44,252"/>
Шаг нашивки	<input type="text" value="6,899"/>
Плотность нашивки	<input type="text" value="57,154"/>
<input type="button" value="Отправить"/>	

Удаленный репозиторий

Ссылка на созданный репозиторий на GitHub:

<https://github.com/vinogradov96ma/data-science-PRO>

Спасибо за внимание