



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана

Выпускная квалификационная работа по курсу «Data Science»

Тема: прогнозирование конечных свойств новых материалов (композиционных материалов)

Слушатель: Герасимова Валентина Вениаминовна



Постановка задачи

- 1 Провести разведочный анализ данных.
- 2 Провести предобработку данных.
- 3 Обучить нескольких моделей для прогноза модуля упругости при растяжении и прочности при растяжении.
- 4 Написать нейронную сеть, которая будет рекомендовать соотношение матрица-наполнитель.
- 5 Разработать приложение с графическим интерфейсом или интерфейсом командной строки.



Разведочный анализ данных

Данные по датасету

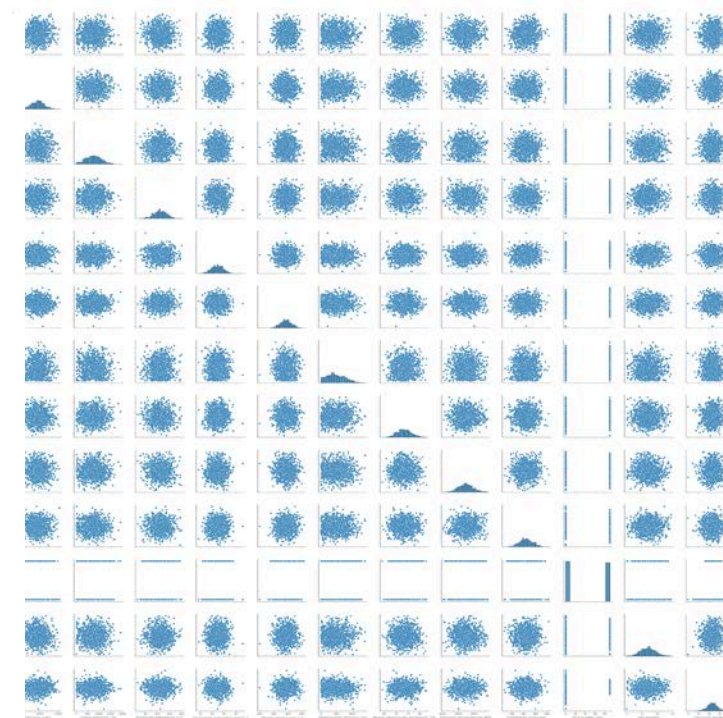
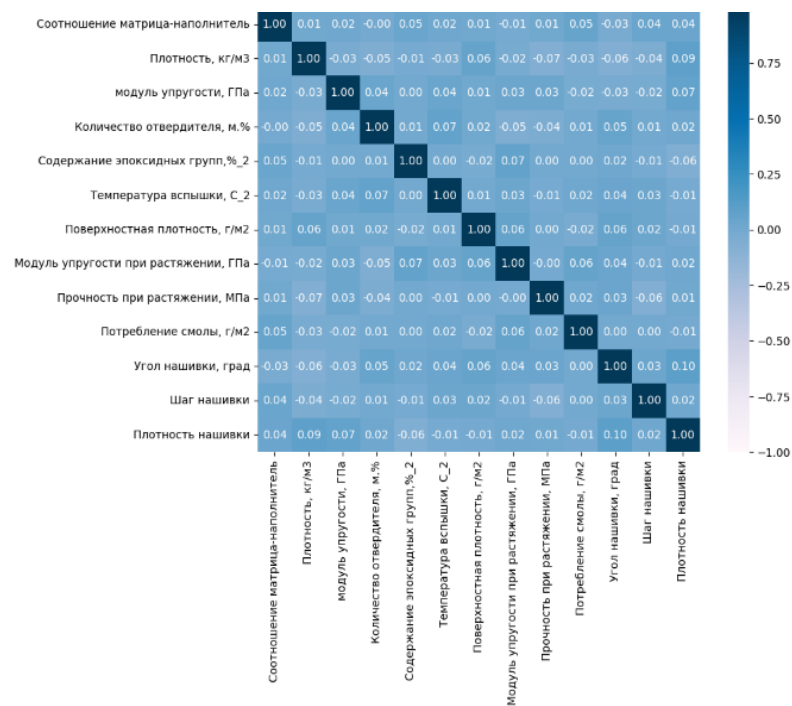
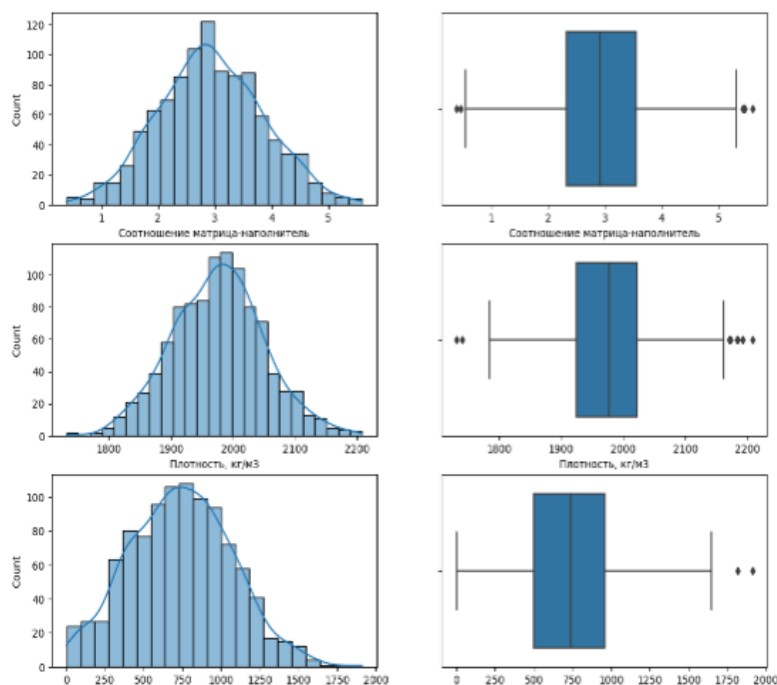
1. Размерность 1023x13
2. Пропуски отсутствуют
3. Дубликаты отсутствуют
4. Тип данных:
 1. Float64
 2. Int64

| | 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 |

Описательная статистика датасета



Разведочный анализ данных



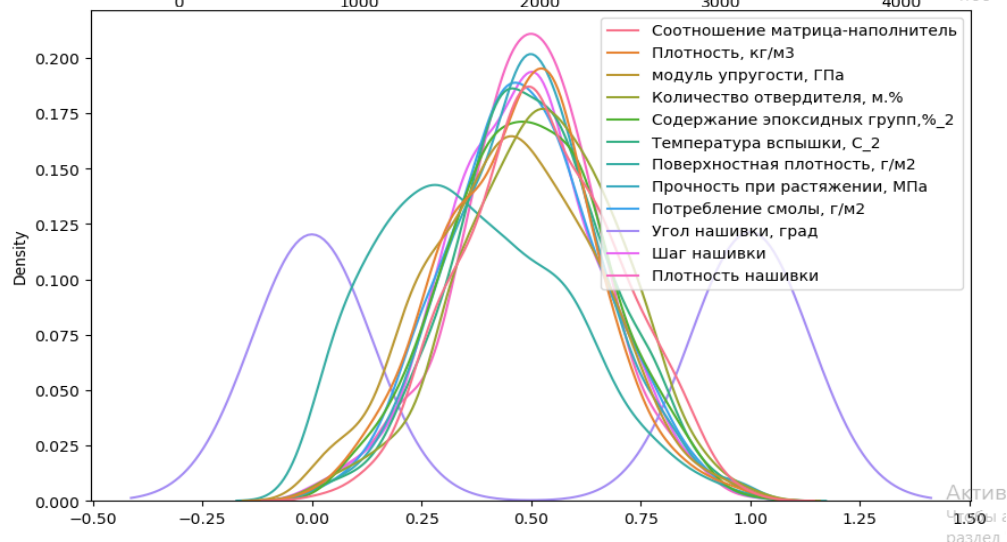
Гистограммы распределения,
диаграмма «ящик с усами»

Тепловая карта
корреляции

Попарные графики
рассеяния точек



Предобработка данных



| | count | mean | std | min | 25% | 50% | 75% | max |
|--------------------------------------|-------|----------|----------|-----|----------|----------|----------|-----|
| Соотношение матрица-наполнитель | 905.0 | 0.513209 | 0.178451 | 0.0 | 0.392589 | 0.510464 | 0.636563 | 1.0 |
| Плотность, кг/м3 | 905.0 | 0.477730 | 0.173157 | 0.0 | 0.352835 | 0.482835 | 0.587714 | 1.0 |
| модуль упругости, ГПа | 905.0 | 0.459297 | 0.194139 | 0.0 | 0.320086 | 0.457179 | 0.590467 | 1.0 |
| Количество отвердителя, м.% | 905.0 | 0.522108 | 0.171822 | 0.0 | 0.402974 | 0.519494 | 0.643543 | 1.0 |
| Содержание эпоксидных групп,%_2 | 905.0 | 0.489338 | 0.175596 | 0.0 | 0.370514 | 0.486678 | 0.619612 | 1.0 |
| Температура вспышки, С_2 | 905.0 | 0.502994 | 0.175168 | 0.0 | 0.383337 | 0.502815 | 0.620667 | 1.0 |
| Поверхностная плотность, г/м2 | 905.0 | 0.367543 | 0.209374 | 0.0 | 0.206505 | 0.348479 | 0.526737 | 1.0 |
| Модуль упругости при растяжении, ГПа | 905.0 | 0.512807 | 0.176103 | 0.0 | 0.395174 | 0.507795 | 0.630032 | 1.0 |
| Прочность при растяжении, МПа | 905.0 | 0.496558 | 0.171111 | 0.0 | 0.378564 | 0.492681 | 0.601861 | 1.0 |
| Потребление смолы, г/м2 | 905.0 | 0.495995 | 0.171316 | 0.0 | 0.384295 | 0.495770 | 0.611168 | 1.0 |
| Угол нашивки, град | 905.0 | 0.501657 | 0.500274 | 0.0 | 0.000000 | 1.000000 | 1.000000 | 1.0 |
| Шаг нашивки | 905.0 | 0.487693 | 0.176315 | 0.0 | 0.362216 | 0.490455 | 0.606312 | 1.0 |
| Плотность нашивки | 905.0 | 0.489191 | 0.166831 | 0.0 | 0.385108 | 0.492545 | 0.593333 | 1.0 |

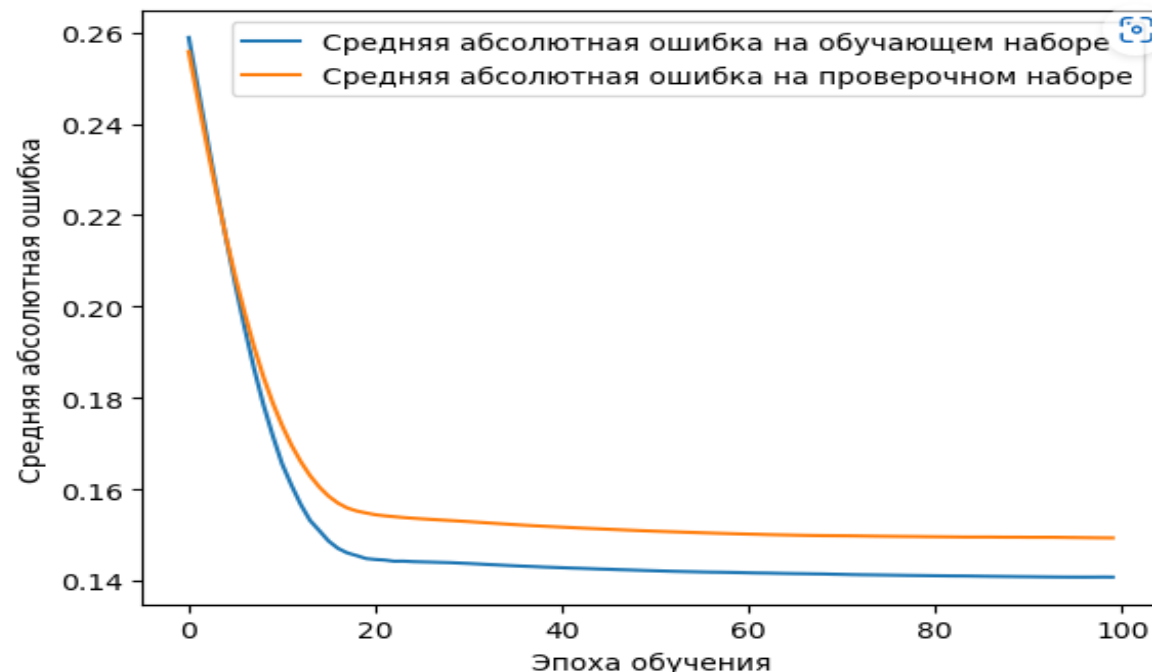
Данные после нормализации



Модели и нейронная сеть

Модели

1. Lasso
2. LinearRegression
3. BayesianRidge
4. DecisionTreeRegressor
5. GradientBoostingRegressor
6. RandomForestRegressor.
7. Нейронная сеть —
последовательная модель
Sequential





Инструменты

1. Python
2. Flask – микрофреймворк
3. Html – язык гипертекстовой разметки
4. TensorFlow SavedModel – формат сохранения модели

Введите параметры для модели

Плотность, кг/м³

Модуль упругости, ГПа

Количество отвердителя, м.%

Содержание эпоксидных групп, %₂

Температура вспышки, С₂

Поверхностная плотность, г/м²

Модуль упругости при растяжении, ГПа

Прочность при растяжении, МПа

Потребление смолы, г/м²

Угол нашивки, град

Шаг нашивки

Плотность нашивки

Рассчитать значение



ЦЕНТР
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
МГТУ им. Н.Э. Баумана



do.bmstu.ru

Спасибо за внимание!