



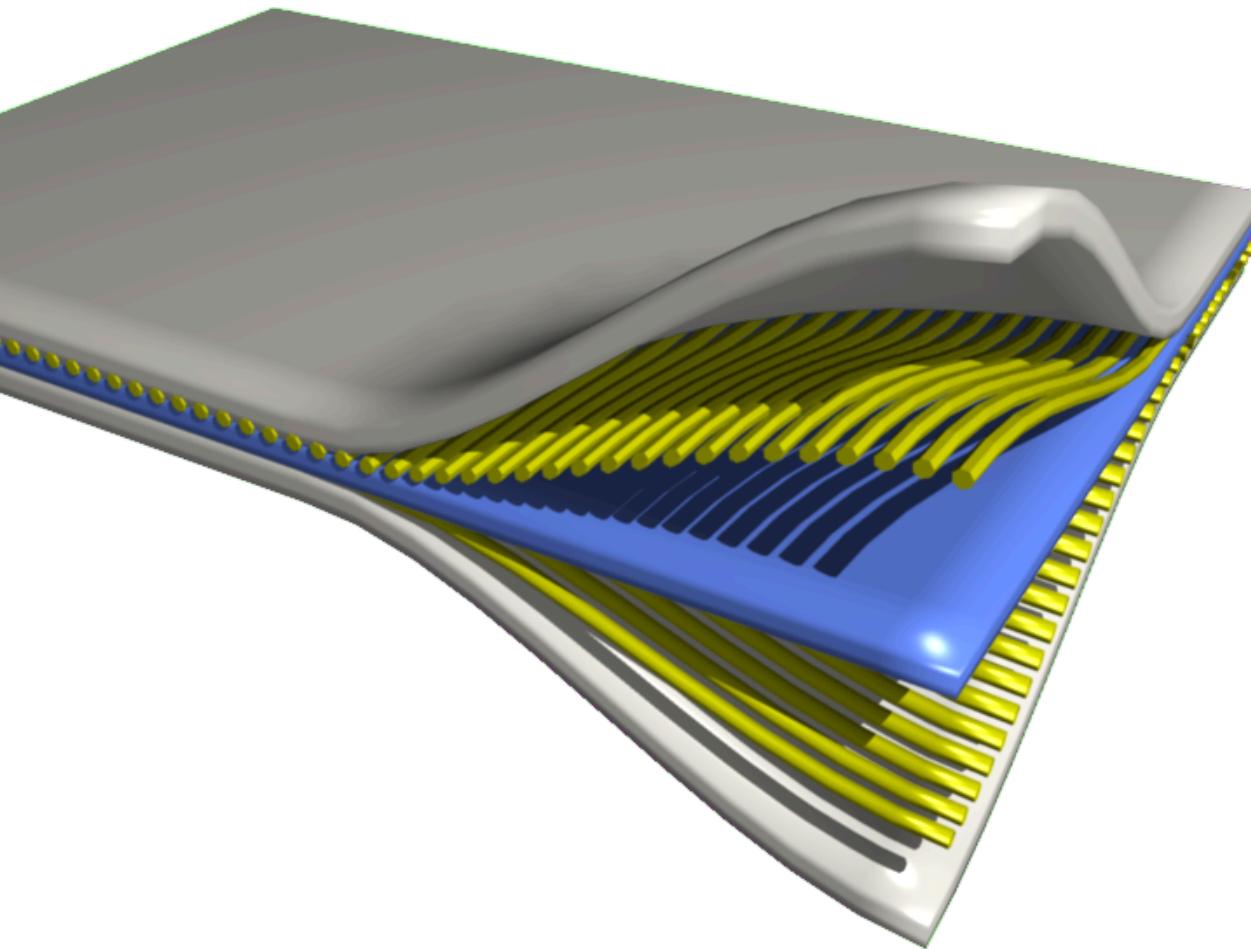
**ОБРАЗОВАТЕЛЬНЫЙ  
ЦЕНТР** МГТУ им. Н. Э. Баумана

**Выпускная квалификационная работа  
по курсу «Data Science».**

**Прогнозирование конечных свойств новых  
материалов (композиционных материалов).**

Слушатель: Лещинский Александр Викторович

## Композитные материалы



Композиционные материалы — это искусственно созданные материалы, состоящие из нескольких других с четкой границей между ними

## Проблема

Трудно определить характеристики композита, состоящего из разных по свойствам компонентов.

## Решение

Создавать прогнозные модели с целью сократить количество проводимых испытаний, а также пополнить базу данных материалов возможными новыми характеристиками материалов, и цифровыми двойниками новых композитов.

## Задача

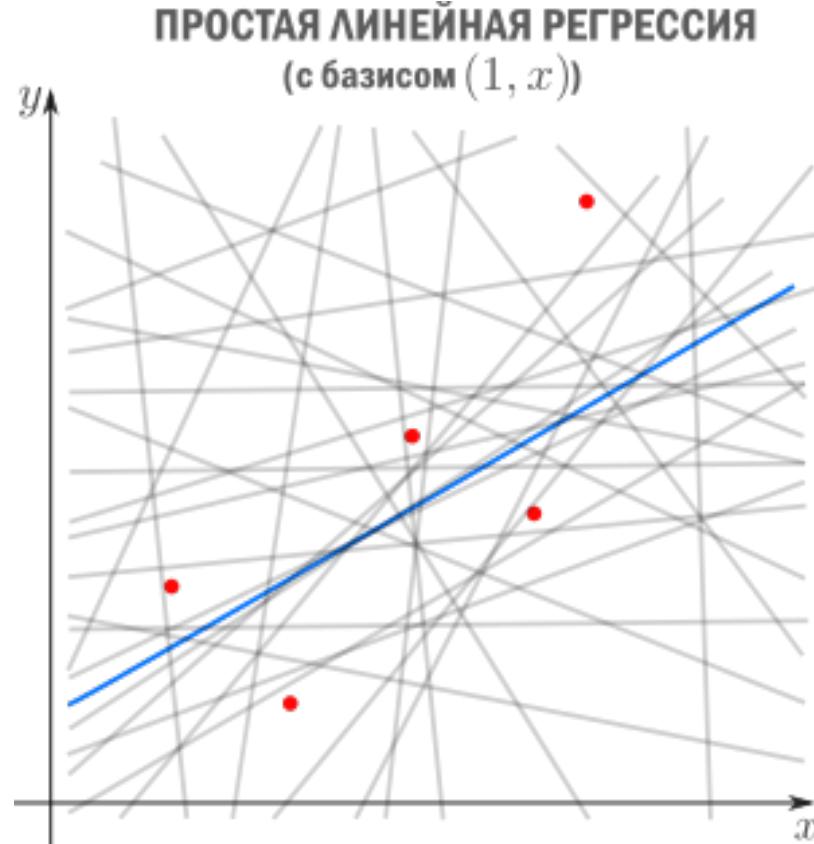
Необходимо разработать несколько моделей для прогнозирования следующих переменных:

- Модуль упругости при растяжении, ГПа;
- Прочность при растяжении, МПа.

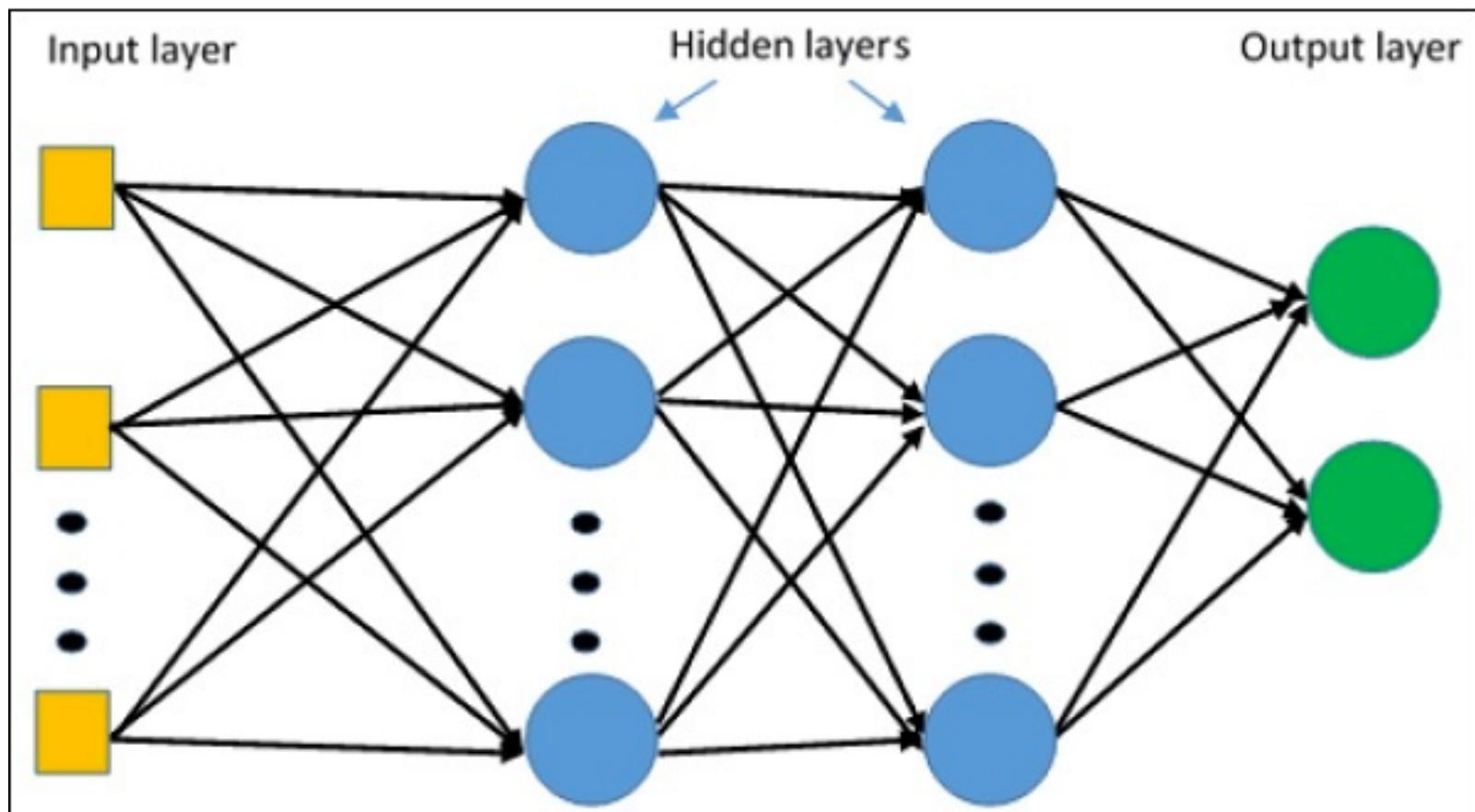
Для изучения доступен data set, состоящий из характеристик компонентов композитных материалов и характеристик производственного процесса. Общий объем dataset - 1040 измерений (строк) для каждой из 13 переменных:

- Угол нашивки, град; Шаг нашивки;
- Плотность нашивки; Соотношение матрица-наполнитель;
- Плотность, кг/м3; Модуль упругости, Гпа;
- Количество отвердителя, м.%; Содержание эпоксидных групп,%\_2;
- Температура вспышки, С\_2; Поверхностная плотность, г/м2;
- Модуль упругости при растяжении, ГПа; Прочность при растяжении, МПа;
- Потребление смолы, г/м2.

# Регрессия



## Многослойный перцептрон (нейронные сети)

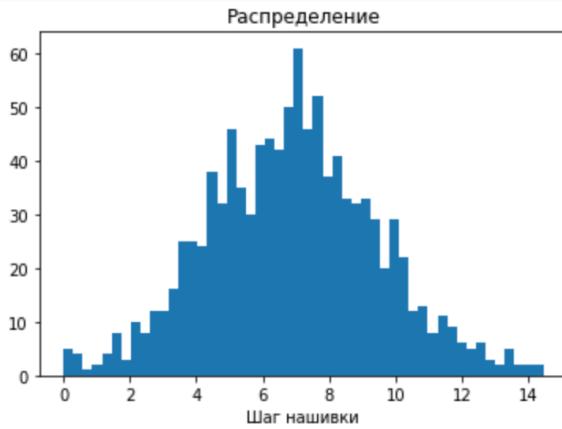


## Разведочный анализ данных

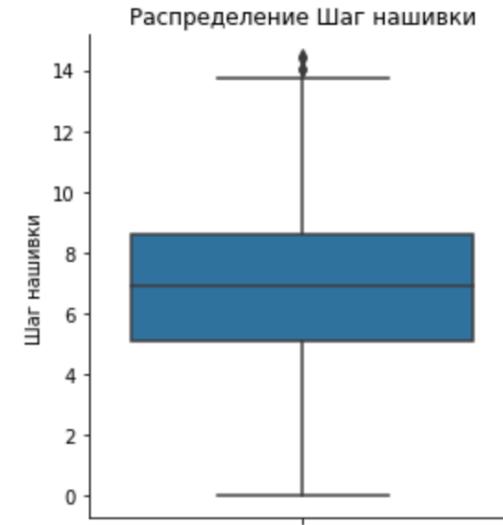
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1040 entries, 0 to 1039
Data columns (total 13 columns):
 #   Column           Non-Null Count  Dtype  
 ---  --  
 0   Угол нашивки, град      1040 non-null   int64  
 1   Шаг нашивки            1040 non-null   float64 
 2   Плотность нашивки     1040 non-null   float64 
 3   Соотношение матрица-наполнитель 1023 non-null   float64 
 4   Плотность, кг/м3        1023 non-null   float64 
 5   модуль упругости, ГПа    1023 non-null   float64 
 6   Количество отвердителя, м.% 1023 non-null   float64 
 7   Содержание эпоксидных групп,%_2 1023 non-null   float64 
 8   Температура вспышки, С_2       1023 non-null   float64 
 9   Поверхностная плотность, г/м2 1023 non-null   float64 
 10  Модуль упругости при растяжении, ГПа 1023 non-null   float64 
 11  Прочность при растяжении, МПа    1023 non-null   float64 
 12  Потребление смолы, г/м2       1023 non-null   float64 
dtypes: float64(12), int64(1)
memory usage: 105.8 KB
```

## Разведочный анализ данных

Гистограммы распределения

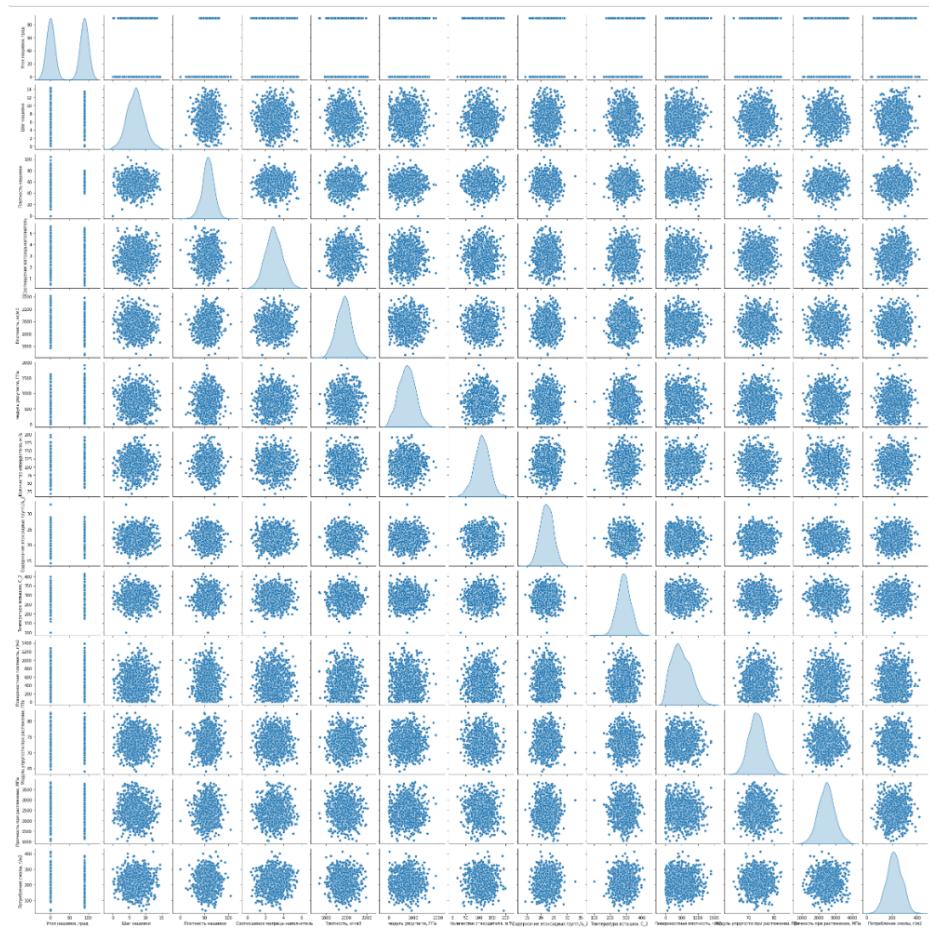


Диаграммы ящика с усами



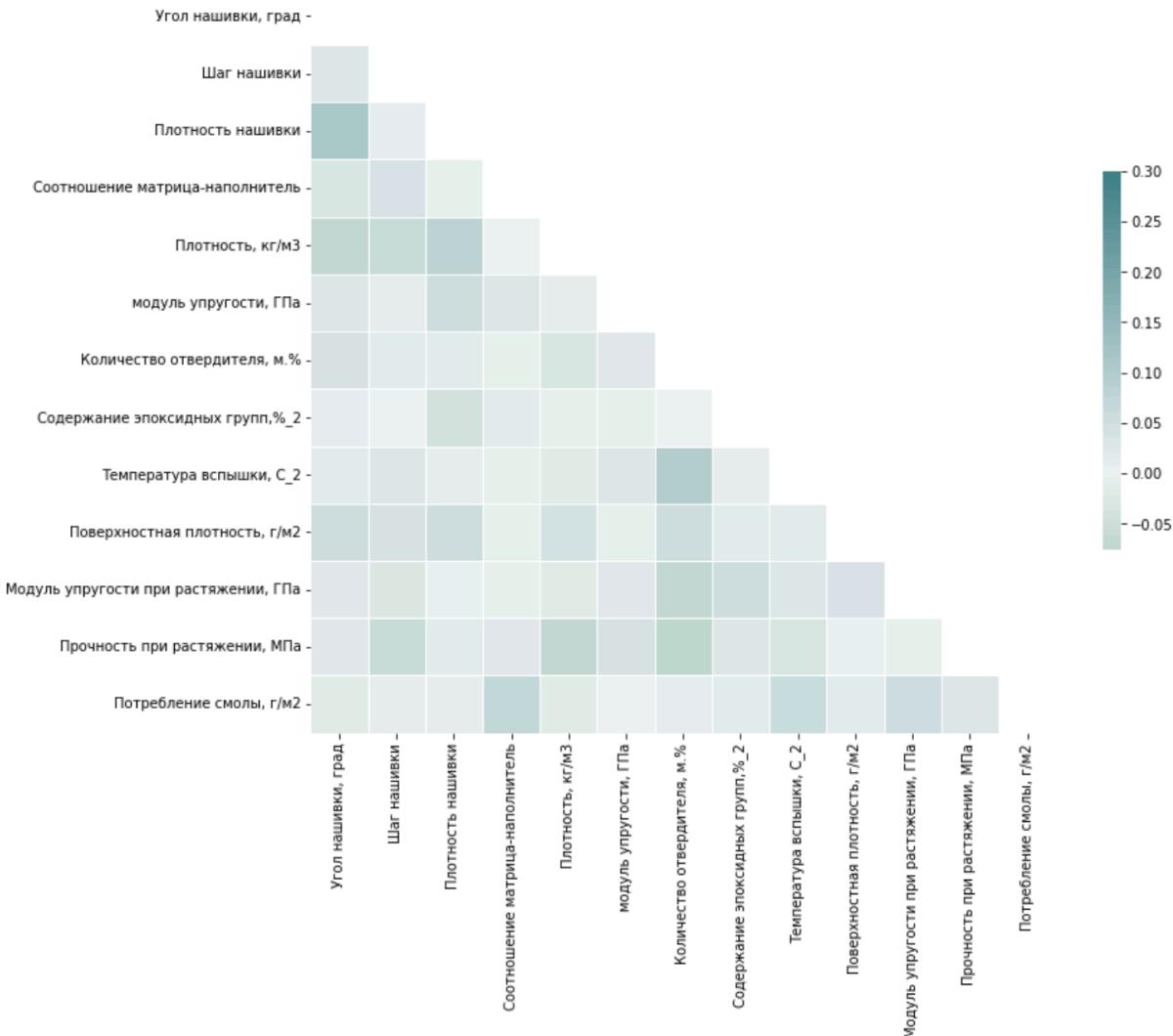
# Разведочный анализ данных

## Попарные графики



# Разведочный анализ данных

## Корреляция признаков



# Предобработка данных

## Удаление

Угол нашивки, град	0	Угол нашивки, град	0
Шаг нашивки	0	Шаг нашивки	0
Плотность нашивки	0	Плотность нашивки	0
Соотношение матрица–наполнитель	17	Соотношение матрица–наполнитель	0
Плотность, кг/м3	17	Плотность, кг/м3	0
модуль упругости, ГПа	17	модуль упругости, ГПа	0
Количество отвердителя, м.%	17	Количество отвердителя, м.%	0
Содержание эпоксидных групп, %_2	17	Содержание эпоксидных групп, %_2	0
Температура вспышки, С_2	17	Температура вспышки, С_2	0
Поверхностная плотность, г/м2	17	Поверхностная плотность, г/м2	0
Модуль упругости при растяжении, ГПа	17	Модуль упругости при растяжении, ГПа	0
Прочность при растяжении, МПа	17	Прочность при растяжении, МПа	0
Потребление смолы, г/м2	17	Потребление смолы, г/м2	0
dtype: int64		dtype: int64	

# Предобработка данных

## Нормализация

	Угол нашивки, град	Шаг нашивки	Плотность нашивки	Соотношение матрица- наполнитель	Плотность, кг/м3	модуль упругости, ГПа	Количество отвердителя, м.%	Содержание эпоксидных групп,%_2	Температура вспышки, С_2	Поверхностная плотность, г/ м2	Моду- упругос- п. растяже- ния
count	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.000000	1023.0000
mean	0.491691	0.477768	0.549616	0.488427	0.512533	0.386301	0.512273	0.426215	0.593354	0.344638	0.4978
std	0.500175	0.177519	0.118772	0.175541	0.154890	0.172978	0.156147	0.128370	0.130695	0.201092	0.1674
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
25%	0.000000	0.351790	0.478890	0.370696	0.404175	0.260652	0.412240	0.338919	0.507756	0.190296	0.3860
50%	0.000000	0.478940	0.551423	0.483912	0.516497	0.386165	0.512240	0.425487	0.593401	0.322574	0.4946
75%	1.000000	0.594597	0.624537	0.608045	0.608413	0.502528	0.618003	0.517842	0.679924	0.495105	0.6067
max	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.0000

# Предобработка данных

## Стандартизация

	Угол нашивки, град	Шаг нашивки	Плотность нашивки	Соотношение матрица- наполнитель	Плотность, кг/м3	модуль упругости, ГПа	Количество отвердителя, м.%	Содержание эпоксидных групп,%_2	Температура вспышки, С_2	Поверхностная плотность, г/ м2	Модуль упругости при растяжении, ГПа	Пр растя
0	-0.983518	-1.131530	-0.012469	-1.175779	0.736365	-0.003594	-2.848828	0.009757	-4.542215	-0.969964	-1.067720	1
1	-0.983518	-1.131530	0.230546	-1.175779	0.736365	-0.003594	-2.141666	0.626001	-0.030955	-0.969964	-1.067720	1
2	-0.983518	-1.131530	1.040595	-1.175779	0.736365	-0.003594	-2.145202	4.471955	-0.030955	-0.969964	-1.067720	1
3	-0.983518	-0.741243	-0.822518	-1.175779	0.736365	-0.003594	0.651622	-0.413446	0.344984	-0.969964	-1.067720	1
4	-0.983518	-0.741243	-0.012469	-0.174232	0.736365	0.039618	0.045585	0.009757	-0.030955	-0.969964	-1.067720	1
...	...	...	...	...	...	...	...	...	...	...	...	...
1018	1.016758	0.849718	-0.820917	-0.721995	-0.320884	0.523926	-0.833694	-0.881925	0.950375	-0.972814	-0.076219	-0
1019	1.016758	1.430947	-0.275671	0.562741	1.008971	-0.894327	1.252072	-1.099577	-0.773808	-0.469707	-0.130794	-0
1020	1.016758	-1.068634	0.848588	0.383707	-0.045622	-0.978843	-0.001319	0.712276	-0.915350	0.915476	0.450935	0
1021	1.016758	-0.228717	0.089684	0.849043	1.235730	0.004703	1.089990	-1.246274	-0.246860	0.564542	0.229077	-0
1022	1.016758	-0.320161	1.642823	0.961522	-1.157793	-0.977389	0.658107	2.174679	0.368264	0.981645	0.314722	0

1023 rows × 13 columns

## Разработка и обучение модели

Для прогноза модуля упругости при растяжении и прочности при растяжении данные были разделены в соотношении:

- 70% обучение модели;
- 30% тестирование модели.

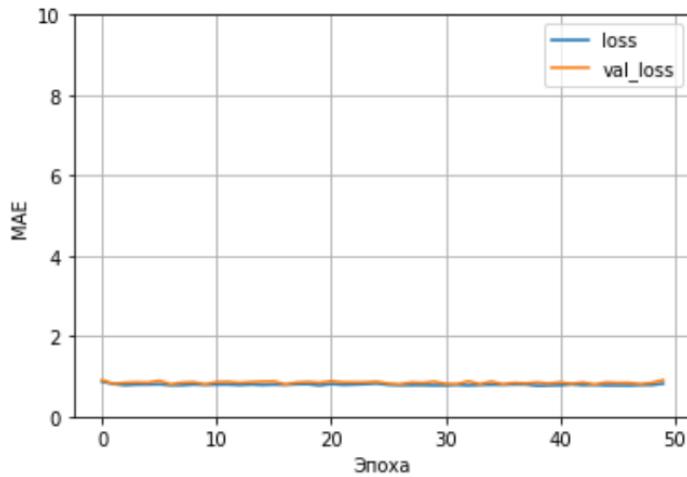
Были использованы следующие модели:

- Модель 1 - линейная регрессия leaner\_model;
- Модель 2- многослойный персепtron dnn\_model;
- Модель 3 – проведены эксперименты с многослойным персептроном dnn\_model2.

# Разработка и обучение модели

```
plot_loss(history)
linear_model.evaluate(X_test, y_test, verbose=0)
```

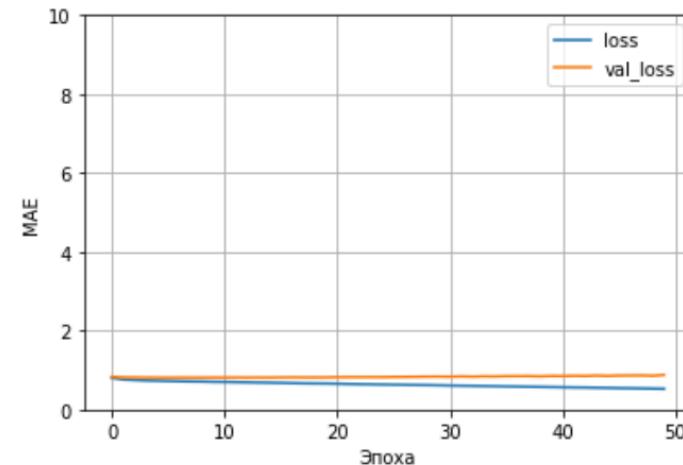
0.941592812538147



```
%%time
history = dnn_model.fit(
    X_train,
    y_train,
    validation_split=0.2,
    verbose=0, epochs=50)
plot_loss(history)
dnn_model.evaluate(X_test, y_test, verbose=0)
```

CPU times: user 3.98 s, sys: 696 ms, total: 4.67 s  
Wall time: 4.33 s

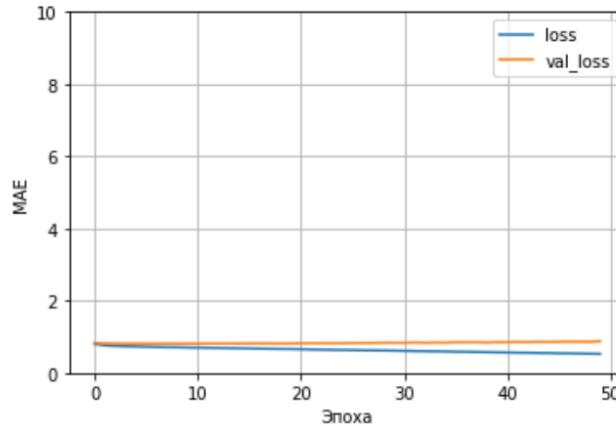
0.9466024041175842



# Разработка и обучение модели

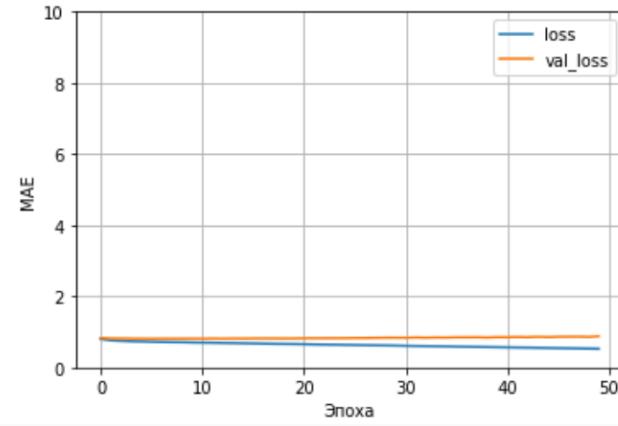
```
%%time
history2 = dnn_model2.fit(
    X_train2,
    y_train2,
    validation_split = 0.1,#изменили на 0.1
    verbose=1, epochs=70)#добавили 20 эпох
plot_loss(history)
dnn_model2.evaluate(X_test2, y_test2, verbose=0)

Wall time: 9.88 s
0.15084879100322723
```



```
%%time
history2 = dnn_model2.fit(
    X_train2,
    y_train2,
    validation_split = 0.3,
    verbose=1, epochs=50)
plot_loss(history)
dnn_model2.evaluate(X_test2, y_test2, verbose=0)

Wall time: 7.94 s
0.15041205286979675
```



## Тестирование модели

```
: pd.DataFrame(test_results, index=['Mean absolute error']).T
```

Mean absolute error	
linear_model	0.941593
dnn_model	0.946602
dnn_model2	0.150412

Для модели dnn\_model2 выявлена наименьшая абсолютная средняя ошибка.

## Решение

Для прогнозирования соотношения матрица – наполнитель была выбрана модель dnn\_model2. Активационная функция sigmoid, оптимизатор RMSprop, функция потерь MAE, 3 скрытых слоя определены экспериментальным путем, на выходе 1 нейрон.

## Заключение

В результате исследования было построено несколько моделей для прогнозирования конечных свойств композитных материалов с использованием методов, изученных на курсе «Data Science».

Разработано приложение.

Создан удаленный репозиторий и на него загружены результаты работы.

---

# Спасибо за внимание