



ЦЕНТР  
ДОПОЛНИТЕЛЬНОГО  
ОБРАЗОВАНИЯ  
МГТУ им. Н.Э. Баумана

# ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

по курсу

«Data Science»

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

Докладчик: Иванова Дарья Александровна



# Постановка задачи

1

Разведочный анализ данных

2

Описание используемых методов

3

Обучение ИНС для целевого признака "Модуль упругости при растяжении"

4

Обучение ИНС для целевого признака "Модуль упругости при растяжении"

5

Разработка веб-приложения

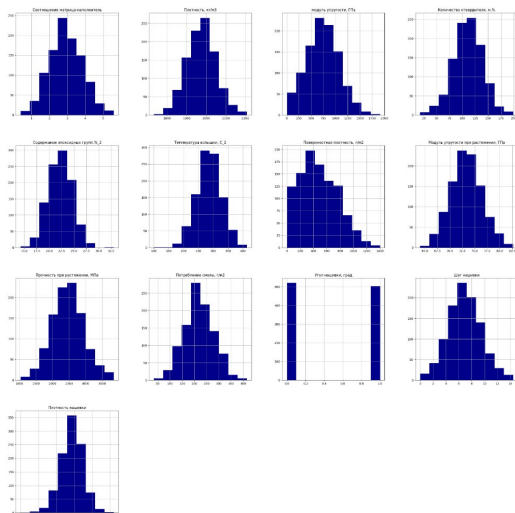
## Дополнительная информация

*Для данного исследования был создан удаленный репозиторий на GitHub, который находится по адресу <https://github.com/Darialvanova007>.*

*На него были загружены результаты работы: исследовательский notebook, код приложения.*



## Разведочный анализ данных



Гистограммы распределения  
переменных

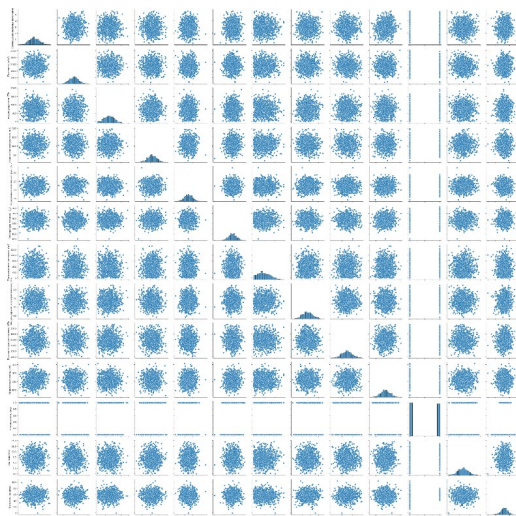
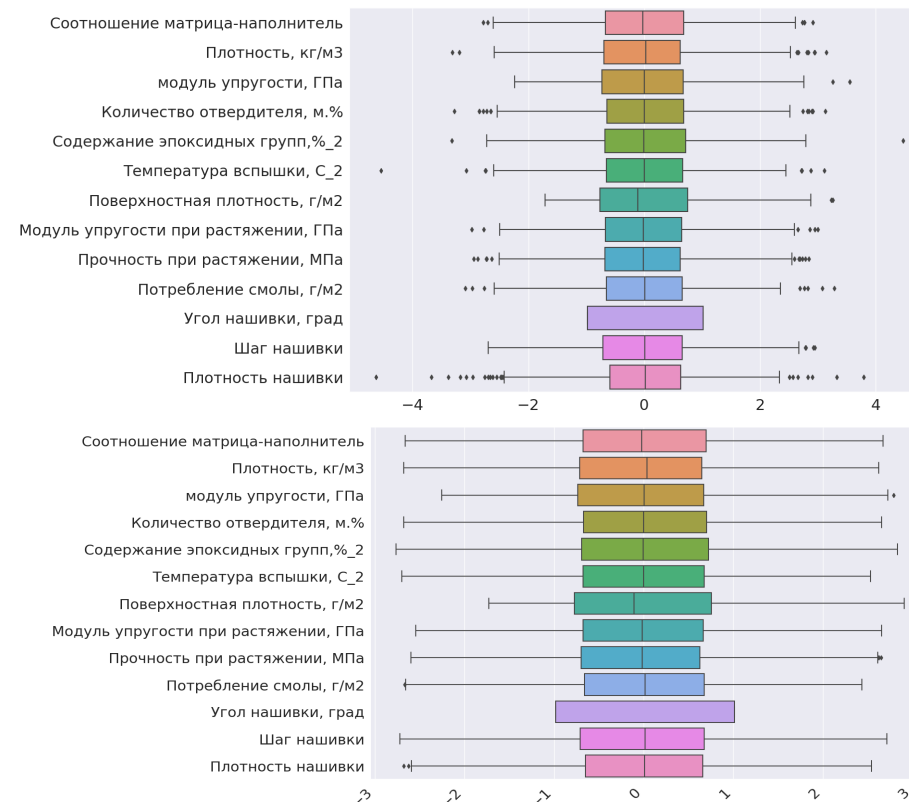


График попарного рассеяния точек

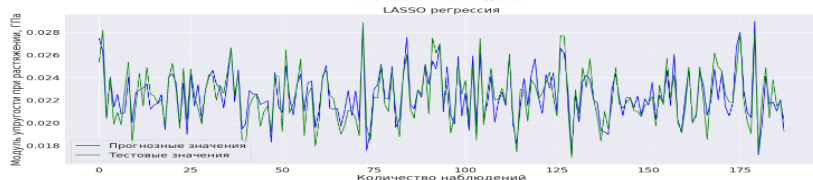
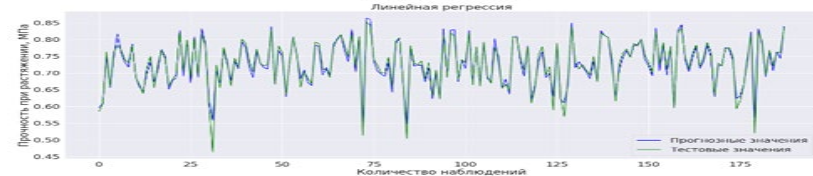


Применив методы 3-х сигм и межквартильных расстояний в датасете было найдено:

- методом 3-х сигм — 24 выброса;
- методом межквартильных расстояний — 93 выброса.



## Описание используемых методов



← Линейная регрессия

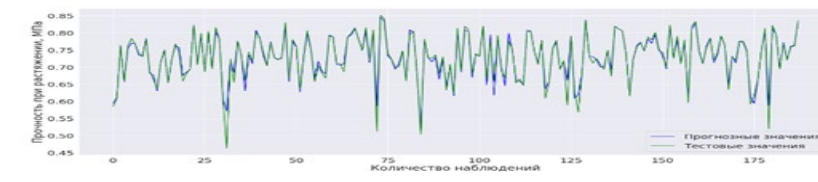
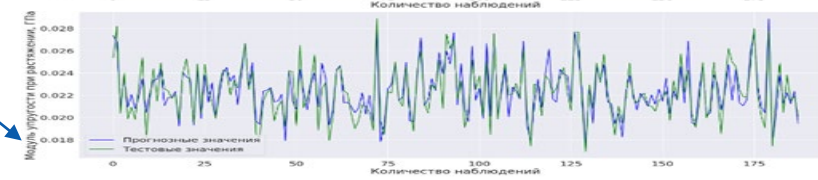
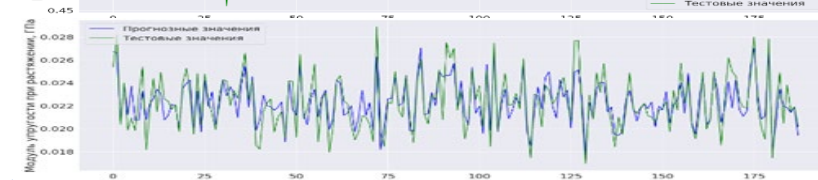
Лассо (LASSO) и гребневая

(Ridge) регрессия

ElasticNet регрессия

Метод k-ближайших соседей

Случайный лес







## Обучение ИНС для целевого признака "Модуль упругости при растяжении"

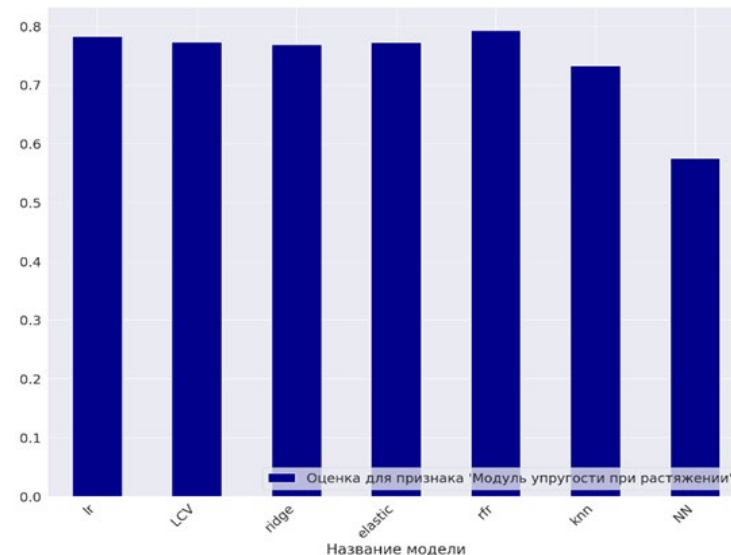
```
[101] #Оценка обученной модели на тестовом наборе для первого целевого признака
result = model.evaluate(X_test, y_test.iloc[:, 0])
print(f"loss: {result[0]} | mae: {result[1]}")

6/6 [-----] - 2s 5ms/step - loss: 0.0042 - mae: 0.0045
loss: 0.00416568061336875 | mae: 0.00450531631708145

#Обучаем модель для первого признака
history_first = model.fit(X_train,
                          y_train.iloc[:, 0],
                          epochs=100,
                          validation_split=0.1,
                          verbose=1)

22/22 [-----] - 0s 5ms/step - loss: 4.9300e-06 - mae: 0.0018 - val_loss: 2.3139e-06 - val_mae: 0.0011
Epoch 4/100
22/22 [-----] - 0s 4ms/step - loss: 3.7194e-06 - mae: 0.0015 - val_loss: 1.9124e-06 - val_mae: 0.0010
Epoch 5/100
22/22 [-----] - 0s 4ms/step - loss: 3.3739e-06 - mae: 0.0015 - val_loss: 1.6809e-06 - val_mae: 9.6592e-04
Epoch 6/100
22/22 [-----] - 0s 4ms/step - loss: 2.7986e-06 - mae: 0.0013 - val_loss: 1.6119e-06 - val_mae: 0.0010
Epoch 7/100
22/22 [-----] - 0s 5ms/step - loss: 2.3925e-06 - mae: 0.0012 - val_loss: 2.3179e-06 - val_mae: 0.0013
Epoch 8/100
22/22 [-----] - 0s 5ms/step - loss: 2.5503e-06 - mae: 0.0013 - val_loss: 1.6482e-06 - val_mae: 0.0010
Epoch 9/100
22/22 [-----] - 0s 5ms/step - loss: 2.0731e-06 - mae: 0.0012 - val_loss: 1.3862e-06 - val_mae: 9.3862e-04
Epoch 10/100
22/22 [-----] - 0s 4ms/step - loss: 1.9099e-06 - mae: 0.0011 - val_loss: 1.3975e-06 - val_mae: 9.3401e-04
Epoch 11/100
22/22 [-----] - 0s 4ms/step - loss: 1.8009e-06 - mae: 0.0011 - val_loss: 1.3501e-06 - val_mae: 9.2131e-04
Epoch 12/100
22/22 [-----] - 0s 4ms/step - loss: 1.7909e-06 - mae: 0.0011 - val_loss: 1.3473e-06 - val_mae: 9.2712e-04
Epoch 13/100
22/22 [-----] - 0s 4ms/step - loss: 1.6822e-06 - mae: 0.0010 - val_loss: 1.5746e-06 - val_mae: 9.8915e-04
Epoch 14/100
22/22 [-----] - 0s 4ms/step - loss: 1.6822e-06 - mae: 0.0010 - val_loss: 1.5746e-06 - val_mae: 9.8915e-04
0 сек. выполнено в 00:34
```

Обучение модели для первого признака



Оценка тестирования моделей для первого признака

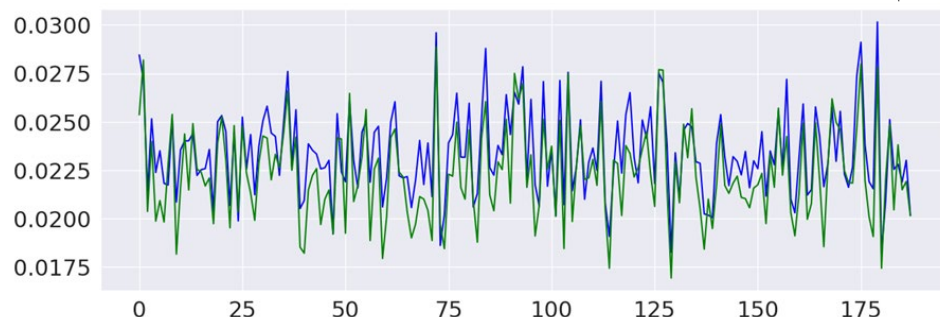


График потерь модели при обучении для первого признака

feature\_1\_with\_NN

Оценка для признака 'Модуль упругости при растяжении'

lr	0.782675
LCV	0.773547
ridge	0.768955
elastic	0.772475
rfr	0.792946
knn	0.732667
NN	0.575539

#Посмотрим информацию по данному датафрейму  
feature\_1\_with\_NN.describe()

#Среднее значение оценки равно 0.769

Оценка для признака 'Модуль упругости при растяжении'

count	7.000000
mean	0.742686
std	0.076041
min	0.575539
25%	0.750811
50%	0.772475
75%	0.778111
max	0.792946

После визуализации оценки по каждой модели для первого признака, лидером по-прежнему остается модель случайного леса.

Качество обобщения для ИНС для первого признака



## Обучение ИНС для целевого признака "Модуль упругости при растяжении"

```
[116] # Оценка обученной модели на тестовом наборе для второго целевого признака
result = model2.evaluate(X_test, y_test.iloc[:, 1])
print(f"loss: {result[0]} | mae: {result[1]}")

6/6 [-----] - 0s 2ms/step - loss: 0.5640 - mae: 0.7479
loss: 0.5639968514442444 | mae: 0.7478866577148438

# Обучаем модель для второго признака
history_second = model2.fit(X_train,
                             y_train.iloc[:, 1],
                             epochs=100,
                             validation_split=0.1)

Epoch 72/100
22/22 [-----] - 0s 4ms/step - loss: 9.9747e-06 - mae: 0.0022 - val_loss: 2.3415e-05 - val_mae: 0.0036
Epoch 73/100
22/22 [-----] - 0s 5ms/step - loss: 9.0989e-06 - mae: 0.0021 - val_loss: 1.5951e-05 - val_mae: 0.0026
Epoch 74/100
22/22 [-----] - 0s 4ms/step - loss: 1.0086e-05 - mae: 0.0023 - val_loss: 6.8806e-05 - val_mae: 0.0072
Epoch 75/100
22/22 [-----] - 0s 4ms/step - loss: 1.7588e-05 - mae: 0.0031 - val_loss: 2.0822e-05 - val_mae: 0.0027
Epoch 76/100
22/22 [-----] - 0s 4ms/step - loss: 8.7216e-06 - mae: 0.0021 - val_loss: 1.6766e-05 - val_mae: 0.0025
Epoch 77/100
22/22 [-----] - 0s 4ms/step - loss: 9.8869e-06 - mae: 0.0023 - val_loss: 2.7550e-05 - val_mae: 0.0037
Epoch 78/100
22/22 [-----] - 0s 4ms/step - loss: 1.6670e-05 - mae: 0.0033 - val_loss: 1.8211e-05 - val_mae: 0.0024
Epoch 79/100
22/22 [-----] - 0s 4ms/step - loss: 8.9504e-06 - mae: 0.0022 - val_loss: 1.4810e-05 - val_mae: 0.0022
Epoch 80/100
22/22 [-----] - 0s 4ms/step - loss: 9.6363e-06 - mae: 0.0023 - val_loss: 1.7141e-05 - val_mae: 0.0030
Epoch 81/100
22/22 [-----] - 0s 4ms/step - loss: 9.7163e-06 - mae: 0.0023 - val_loss: 1.4099e-05 - val_mae: 0.0021
Epoch 82/100
22/22 [-----] - 0s 4ms/step - loss: 8.8231e-06 - mae: 0.0022 - val_loss: 2.0670e-05 - val_mae: 0.0029
Epoch 83/100
22/22 [-----] - 0s 4ms/step - loss: 8.8623e-06 - mae: 0.0022 - val_loss: 1.5203e-05 - val_mae: 0.0028
Epoch 84/100
22/22 [-----] - 0s 4ms/step - loss: 8.1634e-06 - mae: 0.0021 - val_loss: 2.2470e-05 - val_mae: 0.0030
Epoch 85/100
0 сек. выполнено в 00:34
```

Обучение модели для второго признака

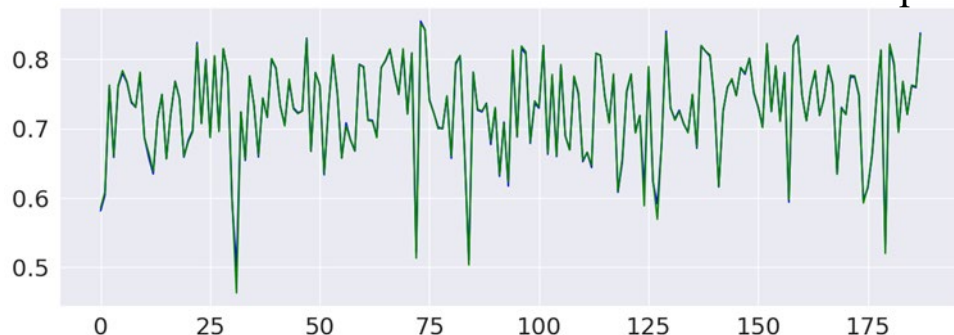
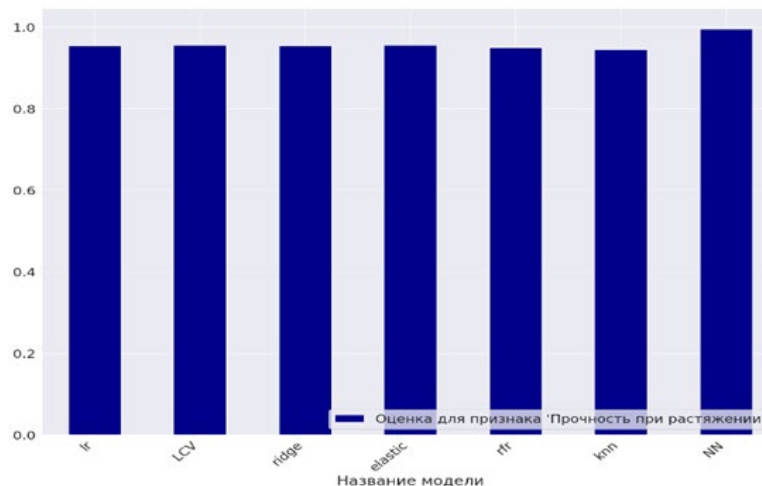


График потерь модели при обучении для второго признака



Оценка тестирования моделей для второго признака

После визуализации оценки по каждой модели для первого признака, лидером по-прежнему остается модель случайного леса.

```
[125] feature_2_with_NN
```

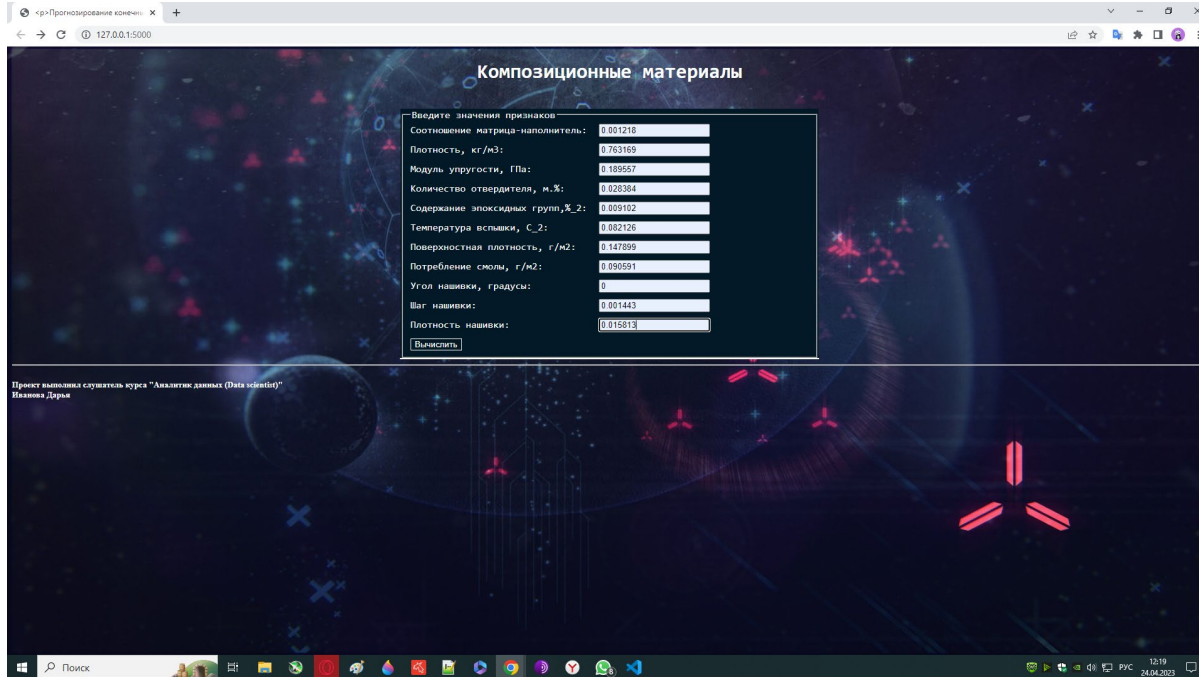
Оценка для признака 'Прочность при растяжении'	
lr	0.955251
LCV	0.956355
ridge	0.954764
elastic	0.956330
rfr	0.950822
knn	0.945850
NN	0.996219

```
#Посмотрим информацию по данному датафрейму
feature_2_with_NN.describe()
```

#Среднее значение оценки равно 0.955

Оценка для признака 'Прочность при растяжении'	
count	7.000000
mean	0.959370
std	0.016684
min	0.945850
25%	0.952793
50%	0.955251
75%	0.956342
max	0.996219

Качество обобщения для ИНС для второго признака



В веб-приложении реализованы следующие функции:

- выбор целевой переменной для предсказания;
- ввод входных параметров;
- проверка введенных параметров;
- загрузка сохраненной модели, получение и отображение прогноза выходных параметров.

Веб-приложение разработано с помощью языка Python, фреймворка Flask и VSCode.





ЦЕНТР  
ДОПОЛНИТЕЛЬНОГО  
ОБРАЗОВАНИЯ  
МГТУ им. Н.Э. Баумана



[do.bmstu.ru](https://do.bmstu.ru)

Спасибо за внимание!