

## Задание №1 (лабораторная работа №1)

Задание выполняется по вариантам на основе файла housing.csv

Рекомендуется выполнить расчеты с использованием ноутбука .ipynb либо в Google Colab

(всё нужное прикреплено в архиве)

### Цель работы:

Предложить среднеквадратическую аппроксимацию табличной функции многих переменных, проанализировать чувствительность точного решения к ошибкам округления, проверить сходимость расчетных и исходных данных

### Исходные данные:

После загрузки csv файла в среду для проведения расчетов нужно поменять параметры test\_size и random\_state (см. таблицу с вариантами):

```
, include_bias = False, interaction_only = True)

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)
poly.fit_transform(X_train)
```

И всё, исходные данные для лабораторной работы готовы!

### Содержание работы:

1. Составьте в матричном виде систему линейных алгебраических уравнений для поиска коэффициентов среднеквадратического приближения  $g$ , вычислите число обусловленности матрицы. Система составляется на основе выборки train
2. Найдите решение системы при помощи обратной матрицы, проверьте полученное приближение с использованием данных из обучающей и из тестовой выборки отдельно (рассчитайте метрику mean\_absolute\_error, сравните расчетные и исходные значения  $y$  ('MEDV') на графиках)

Повторите расчеты для матриц X\_train, X\_test, дополненных с помощью конструктора PolynomialFeatures; выполните аналогичные расчеты при активированном скейлере:

```
#скейлер сейчас "молчит"
scale = StandardScaler(with_mean = False, with_std = False)
#собираем линейные модели
```

В результате выполнения лабораторной работы будет получена таблица:

№ эксперимента	StandardScaler	PolynomialFeatures	mean_absolute_error (train)	mean_absolute_error (test)	np.linalg.cond(X_train.T @ X_train)
1	False	1	...	...	...
2	True	1	...	...	...
3	False	2	...	...	...
...					
...					
...	False	3	...	...	...

### Результат работы:

Определите связь числа обусловленности матрицы системы линейных алгебраических уравнений и средней абсолютной ошибки на обучающей (train) и тестовой (test) выборках отдельно, а результаты представьте графически

**Отчет нужно подготовить в формате pdf, в конце обязательно приведите код!**

## Варианты заданий:

<b>Вариант</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
test_size	0,3	0,28	0,22	0,27	0,2
random_state	41	48	46	40	28
<b>Вариант</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
test_size	0,26	0,24	0,21	0,27	0,29
random_state	45	50	24	38	25
<b>Вариант</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>
test_size	0,22	0,3	0,21	0,24	0,21
random_state	50	28	18	41	49
<b>Вариант</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
test_size	0,2	0,23	0,3	0,21	0,24
random_state	17	46	33	49	46
<b>Вариант</b>	<b>21</b>	<b>22</b>	<b>23</b>	<b>24</b>	<b>25</b>
test_size	0,21	0,22	0,27	0,24	0,26
random_state	35	44	36	37	24
<b>Вариант</b>	<b>26</b>	<b>27</b>	<b>28</b>	<b>29</b>	<b>30</b>
test_size	0,26	0,24	0,29	0,3	0,24
random_state	39	50	27	12	43
<b>Вариант</b>	<b>31</b>	<b>32</b>	<b>33</b>	<b>34</b>	<b>35</b>
test_size	0,22	0,3	0,26	0,22	0,21
random_state	40	15	32	47	45
<b>Вариант</b>	<b>36</b>	<b>37</b>	<b>38</b>	<b>39</b>	<b>40</b>
test_size	0,24	0,29	0,27	0,22	0,23
random_state	41	42	43	44	45
<b>Вариант</b>	<b>41</b>	<b>42</b>	<b>43</b>	<b>44</b>	<b>45</b>
test_size	0,28	0,27	0,22	0,26	0,2
random_state	46	47	48	49	50
<b>Вариант</b>	<b>46</b>	<b>47</b>	<b>48</b>	<b>49</b>	<b>50</b>
test_size	0,22	0,2	0,21	0,29	0,23
random_state	29	46	20	31	43