

# Предсказание появления ребер в темпоральных графах

Работу выполнили:

Менкеев Александр Саналович

Смолина Алина Андреевна

Штоколова Елизавета Викторовна

# Использованные структуры данных

---

- `List[List]` – список списков. Использовался для представления статического графа в виде списка смежности.
- $\{(u, v): [t_1, t_2, \dots, t_n]\}$  – словарь. Использовался для представления темпорального графа в виде списка ребер. Каждому ребру ставится в соответствие список временных меток.
- Класс `Graph` из библиотеки `networkx`

# Общие статистические свойства для тестовых графов

	Vertices	Edges	Density	WCCs	Largest WCC	Diameter	Radius	90-percentile	A.c.c.	D.a.
socfb-Middlebury45	3075	124610	0.026365	4	3069	7	4			
Testgraph_1	9	13	0.3611	1	9	3	2	3.0	0.4	-0.2037
Testgraph_2	34	78	0.13903	1	34	5	3	4.0	0.5706	-0.4756
Testgraph_3	77	254	0.0868	1	77	5	3	4.0	0.57313	-0.1652
Testgraph_4	62	159	0.08408	1	62	8	5	5.0	0.2589	-0.04359

	Vertexes	Edges	Density	WCCs	Largest W CC	Diameter	Radius	90- percentile	A.c.c.	D.a.
Testgraph_5	198	2742	0.14059	1	198	6	4	3.0	0.61745	0.20237
Testgraph_6	19428	96662								
Testgraph_7	325729	1103836	2.0807630 99336715 8e-05	1	352729					
socfb- Reed98						6	4	3.0		
Team_3	1055	55713	0.1002	1	1055	3	2	2.0	0.1	0.00257

# Общие статические свойства графов

	Vertexes	Edges	Density	WCCs	Largest WCC	Diameter	Radius	90-percentile	A.c.c.	D.a.
Rado	167	3250	0.2344	1	167	5	3	3	0.5918	-0.2945
UC	1899	13838	0.0076	4	1893	8	4	4	0.1097	-0.1877
bitA	3783	14124	0.0019	5	3775	8	6	5	0.1770	-0.1685
D-rep	30398	85733	0.0001	373	29652	7	4	4	0.0244	0.0083
SX-MO	24818	193979	0.0006	104	24668	4	2	2	0.6771	-0.1979

# Статические признаки тестовых графов

---

	testgraph _1	testgraph _2	testgraph _3	testgraph _4	testgraph _5	testgraph _6	testgraph _7	socfb- Middlebu ry45	socfb- Reed98	team_3
CN	2	7	0	0	0	0	584	2	19	15
AA	1.26	4.24	0	0	0	0	177.15	0.26	3.27	2.22
JC	0.28	0.38	0	0	0	0	0.07	0.01	0.18	0.06
PA	18	144	10	48	69	9	4718637	8262	3504	14336

# Постановка задачи

---

1. Разделяем исходный темпоральный граф на две части:

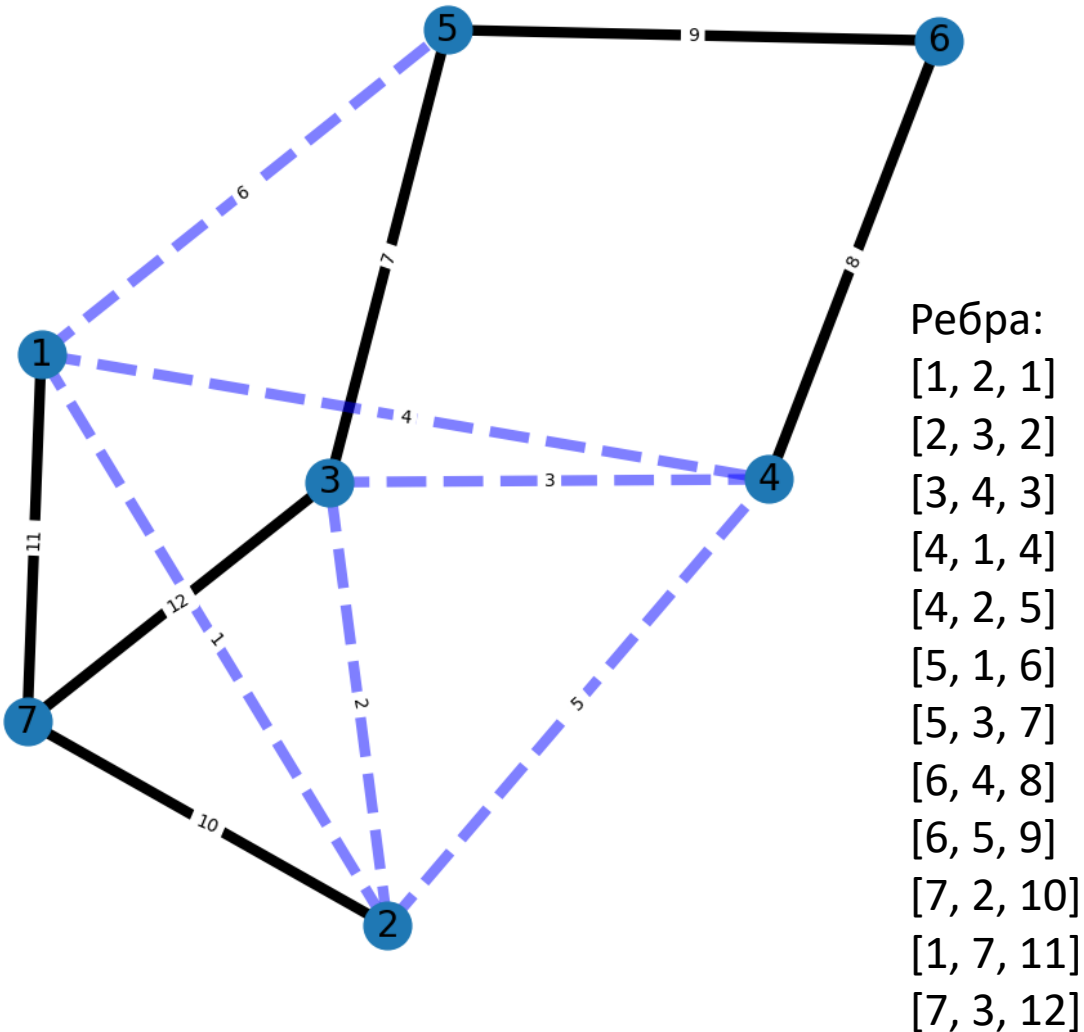
$$\{(u, v, t): t_{min} \leq t \leq qs\} ,$$

$$\{(u, v, t): qs < t \leq t_{max}\} ,$$

где  $qs$  – это  $s$ -й перцентиль массива временных меток

2. Нужно предсказать, появятся ли не существующие в данный момент (то есть до  $qs$ ) ребра в будущем (после  $qs$ )

# Выборка ребер для регрессии



t\_min: 1

t\_max: 12

qs: 6.5

Ребра до qs: [(1, 2), (1, 4), (1, 5), (2, 3), (2, 4), (3, 4)]

Ребра после qs: [(1, 7), (2, 7), (3, 5), (3, 7), (4, 6), (5, 6)]

Пары на расстоянии два: [(1, 3), (2, 5), (4, 5)]

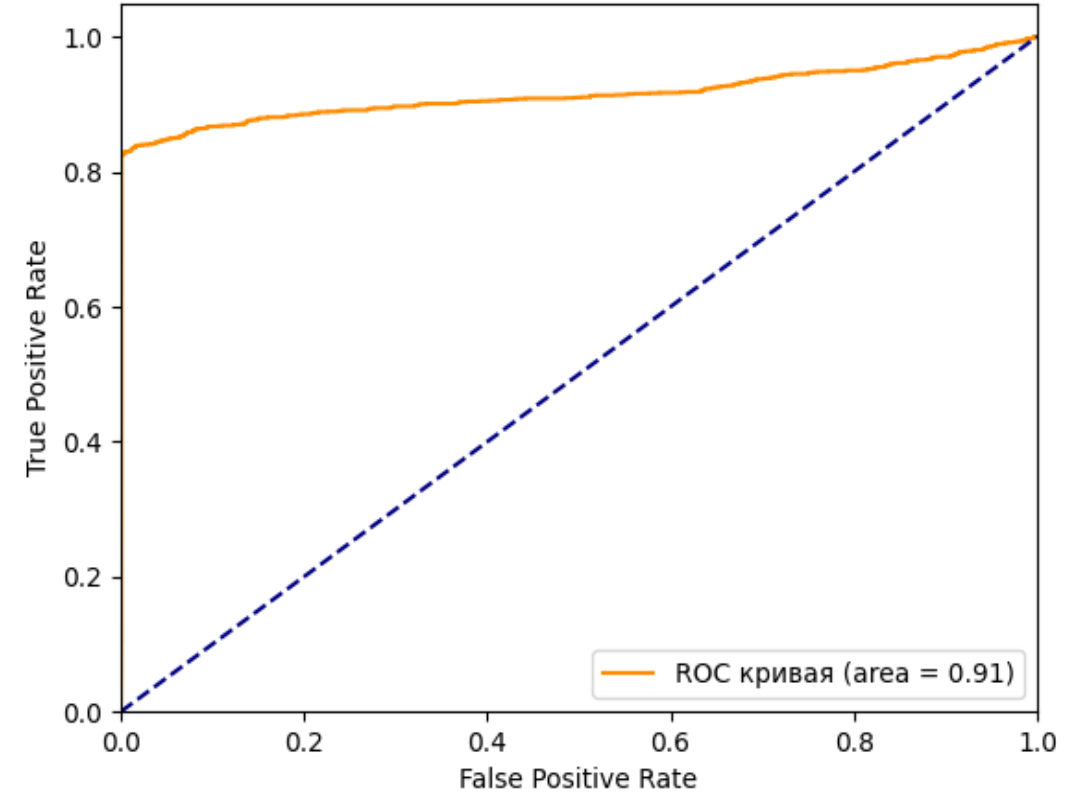
Итоговая выборка = (Пары на расстоянии два)  $\cup$   
(Ребра после qs)

Итоговая выборка: [(1, 3), (2, 5), (4, 5), (1, 7), (2, 7), (3, 5),  
(3, 7), (4, 6), (5, 6)]



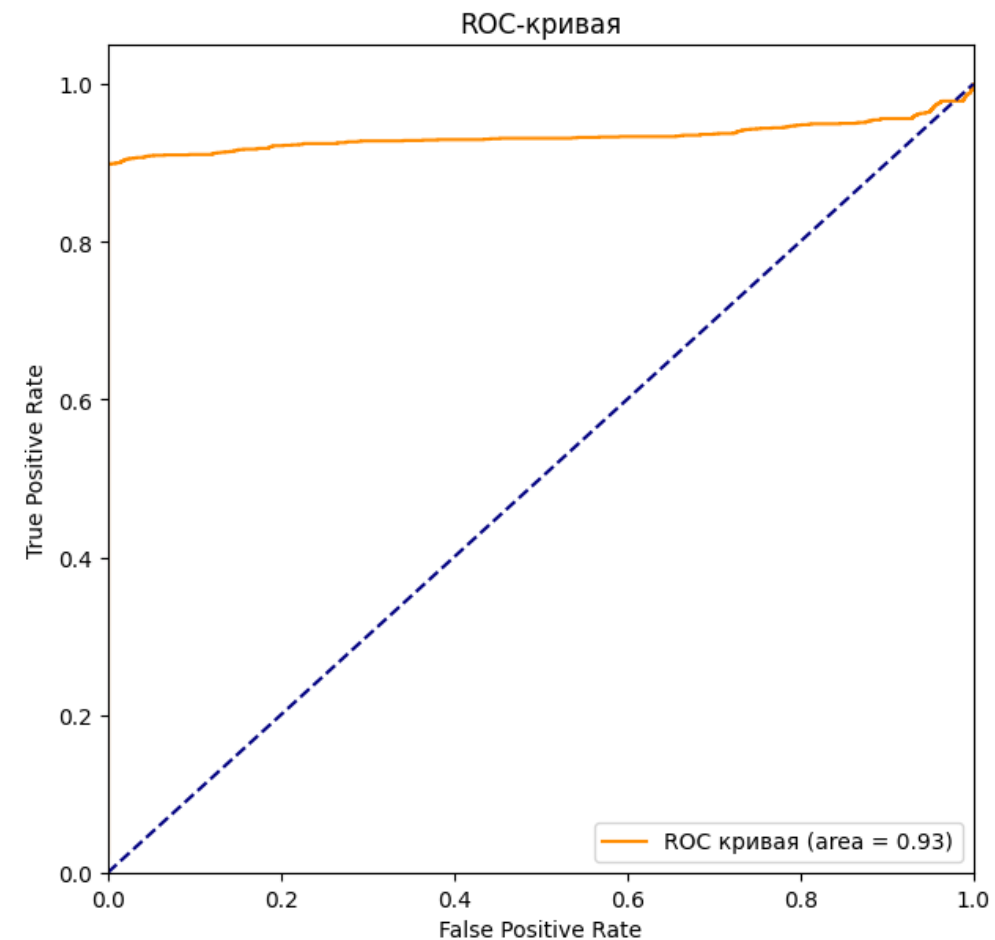
# Dataset UC. Признаки I

	precision	recall	f1-score	support
0	0.99	1.00	1.00	40662
1	0.97	0.82	0.89	1543
accuracy			0.99	42205
macro avg	0.98	0.91	0.94	42205
weighted avg	0.99	0.99	0.99	42205



# Dataset BitA. Признаки I

	precision	recall	f1-score	support
0	1.00	1.00	1.00	44040
1	0.97	0.90	0.93	1801
accuracy			1.00	45841
macro avg	0.99	0.95	0.97	45841
weighted avg	1.00	1.00	0.99	45841



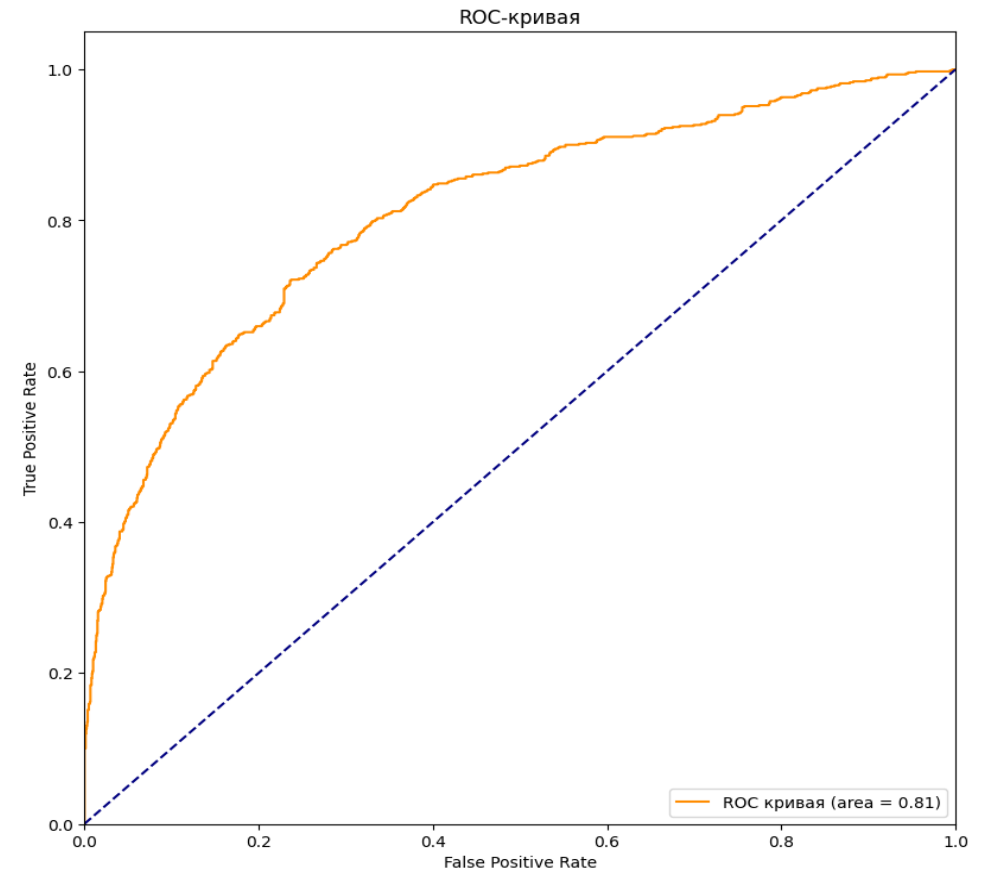
# Темпоральные признаки. Признаки II-A

# Dataset Rado. Признаки II-A

	precision	recall	f1-score	support
0	0.79	0.93	0.85	1610
1	0.75	0.47	0.58	761
accuracy			0.78	2371
macro avg	0.77	0.71	0.72	2371
weighted avg	0.78	0.78	0.76	2371

Count(1): 2964

Count(0): 6520

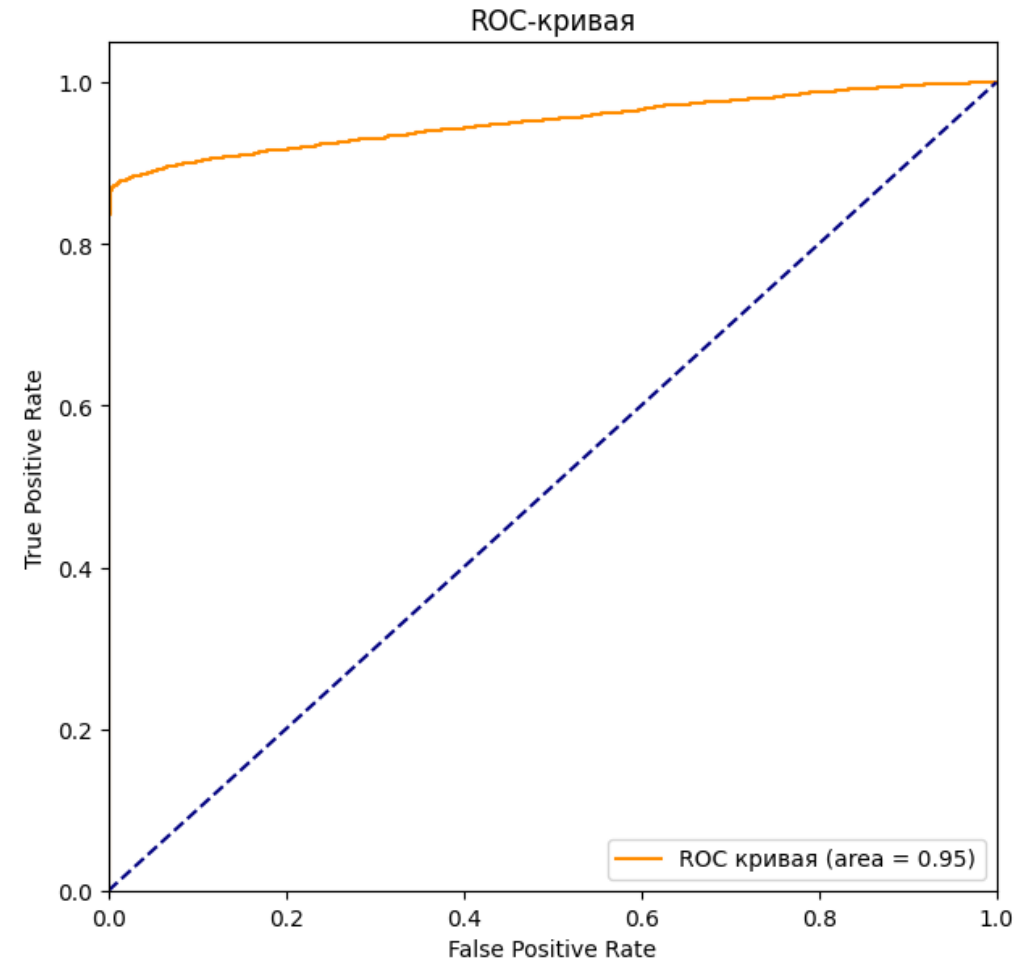


# Dataset UC. Признаки II-A

	precision	recall	f1-score	Support
0	0.98	0.99	0.98	13742
1	0.94	0.87	0.91	2741
accuracy			0.97	16483
macro avg	0.96	0.93	0.95	16483
weighted avg	0.97	0.97	0.97	16483

Count(1): 11201

Count(0): 54728

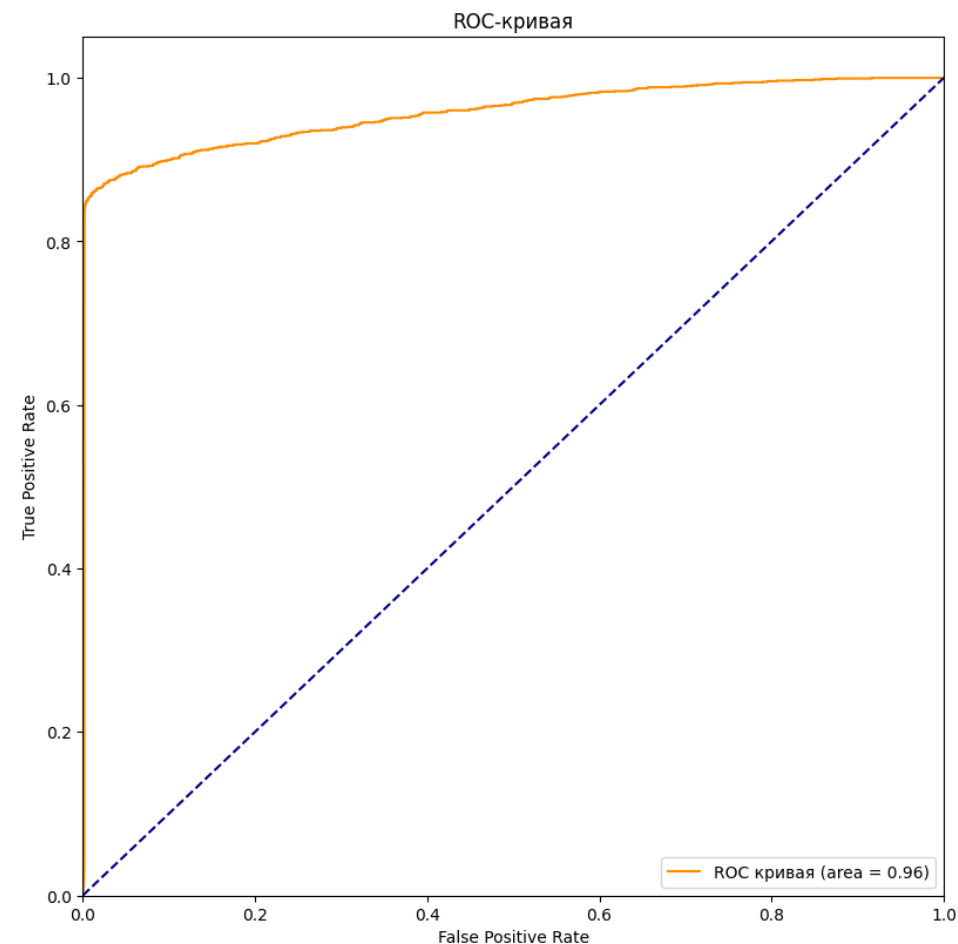


# Dataset BitA. Признаки II-A

	precision	recall	f1-score	support
0	0.99	1.00	1.00	19730
1	0.97	0.95	0.96	2452
accuracy			0.99	22182
macro avg	0.98	0.97	0.98	22182
weighted avg	0.99	0.99	0.99	22182

Count(1): 10155

Count(0): 78571



# Dataset D-rep. Признаки II-A

	precision	recall	f1-score	support
0	0.99	0.99	0.99	38808
1	0.98	0.99	0.98	17240
accuracy			0.99	56048
macro avg	0.99	0.99	0.99	56048
weighted avg	0.99	0.99	0.99	56048

Count(1): 69136

Count(0): 155056

