

Предсказание появления ребер в темпоральных графах

Работу выполнили:

Менкеев Александр Саналович

Смолина Алина Андреевна

Штоколова Елизавета Викторовна

Использованные структуры данных

- $\{u: [v_1, v_2, \dots, v_n]\}$ – словарь. 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
- Set использовался для хранения уникальных значений
- Класс Graph из библиотеки networkx

Общие статистические свойства для тестовых графов

	Vertexes	Edges	Density	WCCs	Largest WCC	Diameter	Radius	90-percentile	A.c.c.	D.a.
socfb-Middlebury45	3075	124610	0.026365	4	3069	7	4	3.0	0.2821	0.078
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

Общие статистические свойства для тестовых графов

	Vertexes	Edges	Density	WCCs	Largest WCC	Diameter	Radius	90-percentile	A.c.c.	D.a.
Testgraph_5	198	2742	0.14059	1	198	6	4	3.0	0.6174	0.20237
Testgraph_6	19428	96662	0.0005122	23	19365	10	6	6.0	0.0	-0.191557
Testgraph_7	325729	1103836	2.0807630 99336715 8e-05	1	352729	4	2	2	0.269	-0.0526
socfb-Reed98	962	18812	0.0406	1	962	6	4	3.0	0.3183	0.0234
Team_3	1055	55713	0.1002	1	1055	3	2	2.0	0.1	0.00257

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

Общие статистические 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

Постановка задачи

1. Разделяем исходный темпоральный граф на две части:

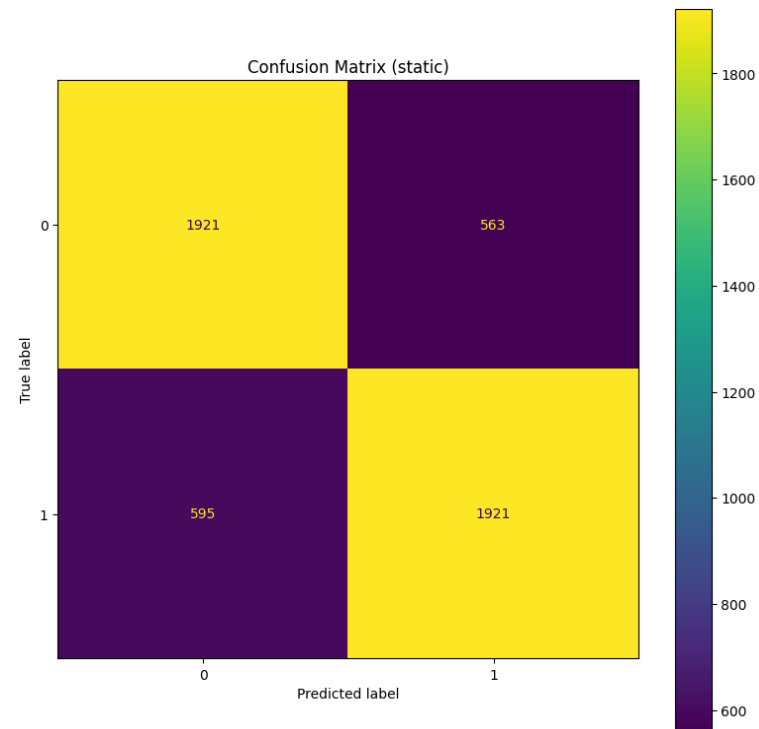
$$\{(u, v, t): t_{min} \leq t \leq qs\} ,$$

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

где qs – это s -й перцентиль массива временных меток

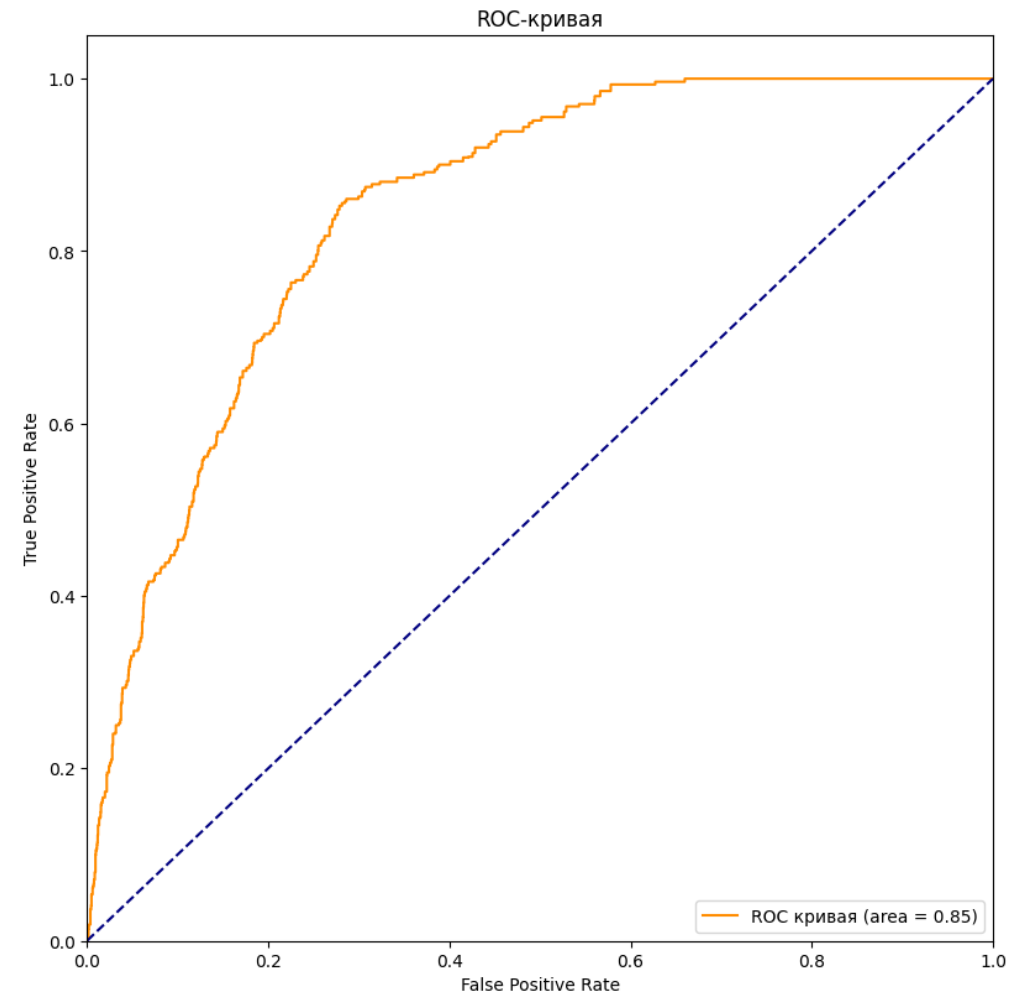
2. Нужно предсказать, появятся ли не существующие в данный момент (то есть до qs) ребра в будущем (после qs)

Rado. Признаки I



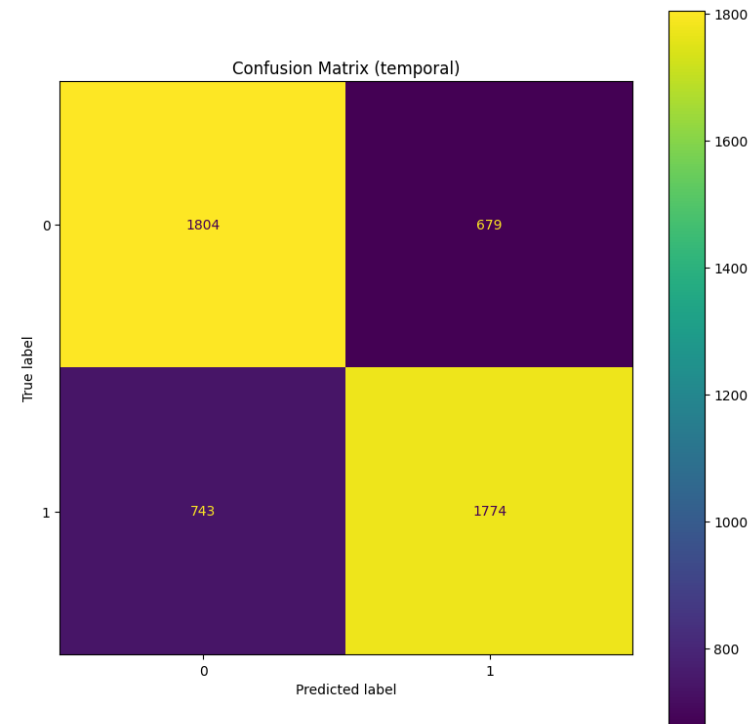
	precision	recall	f1-score	support
0	0.76	0.77	0.77	2484
1	0.77	0.76	0.77	2516
accuracy			0.77	5000
macro avg	0.77	0.77	0.77	5000
weighted avg	0.77	0.77	0.77	5000

LogisticRegression: ROC AUC=0.847



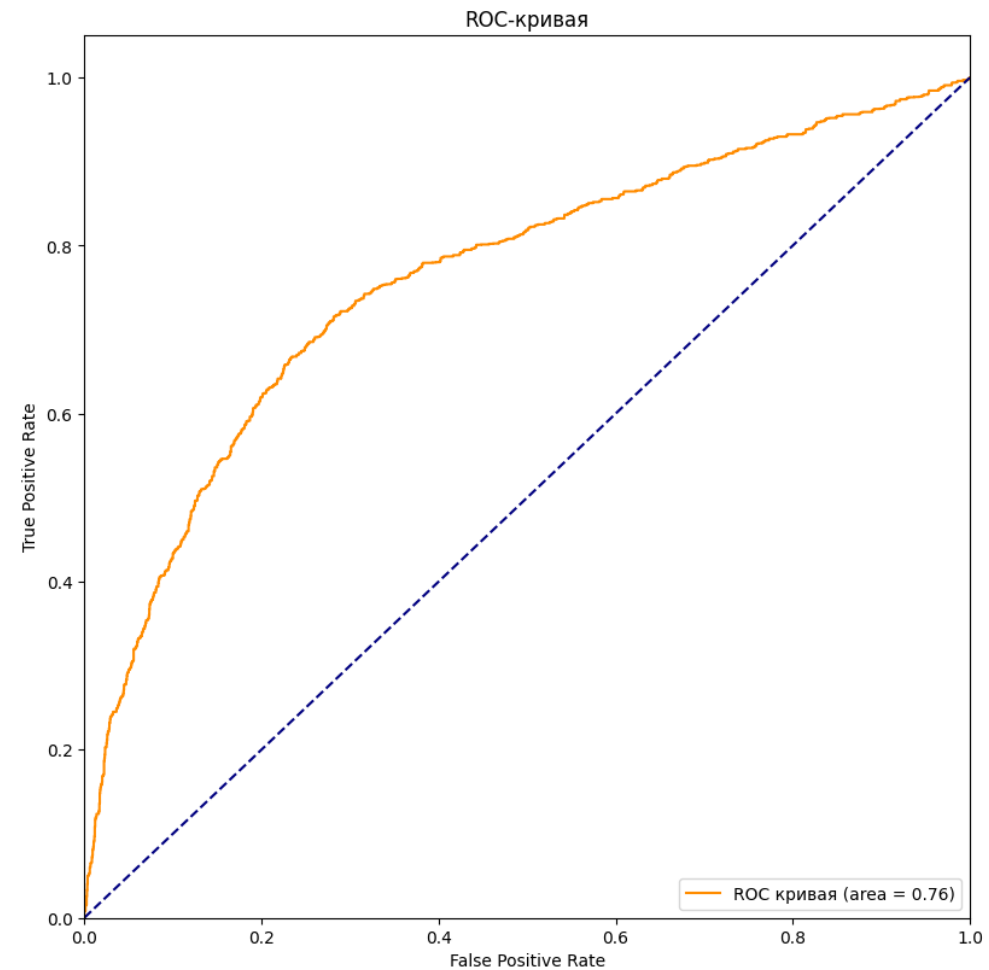
В статье: AUC=0.864

Rado. Признаки II-A



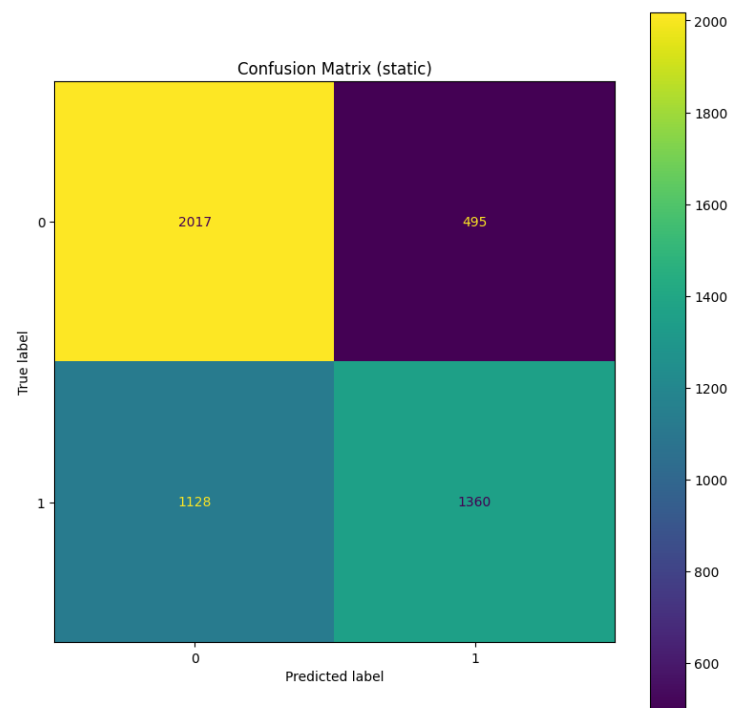
	precision	recall	f1-score	support
0	0.71	0.73	0.72	2483
1	0.72	0.70	0.71	2517
accuracy			0.72	5000
macro avg	0.72	0.72	0.72	5000
weighted avg	0.72	0.72	0.72	5000

LogisticRegression: ROC AUC=0.760



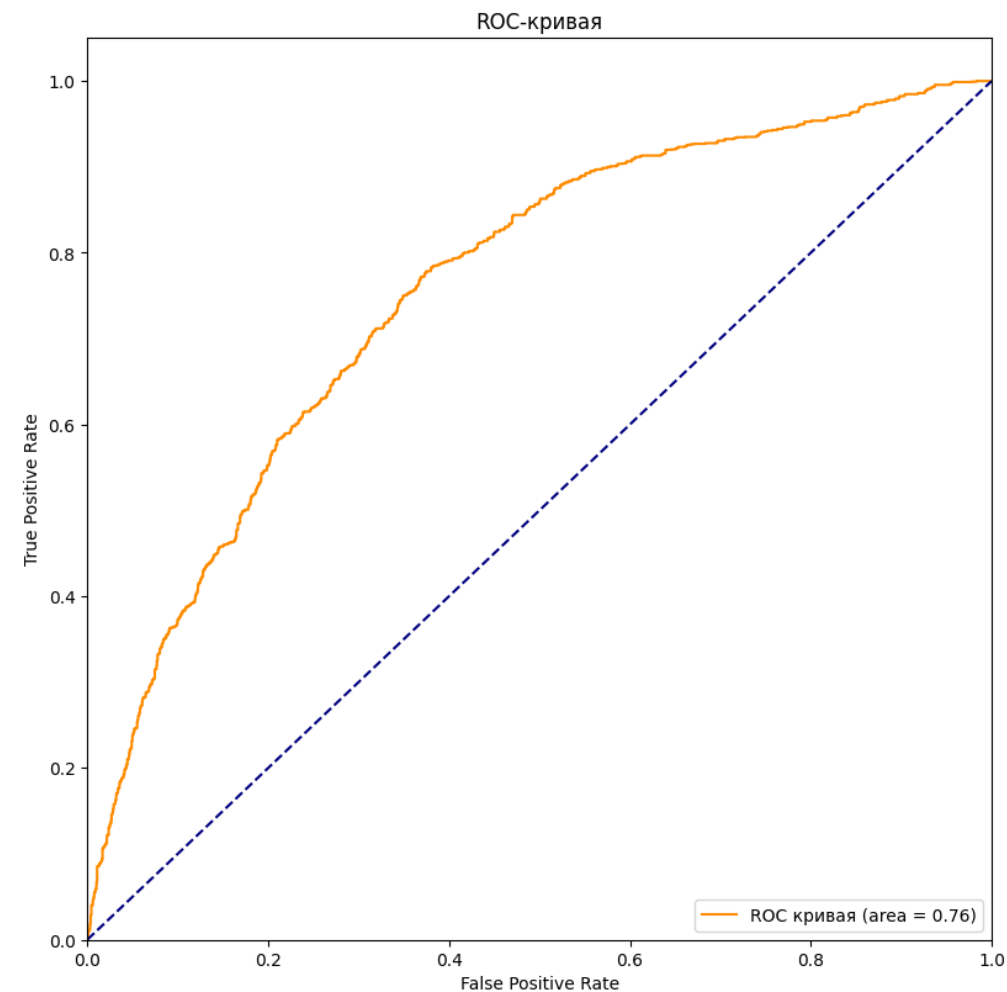
В статье: AUC=0.921

УС. Признаки I



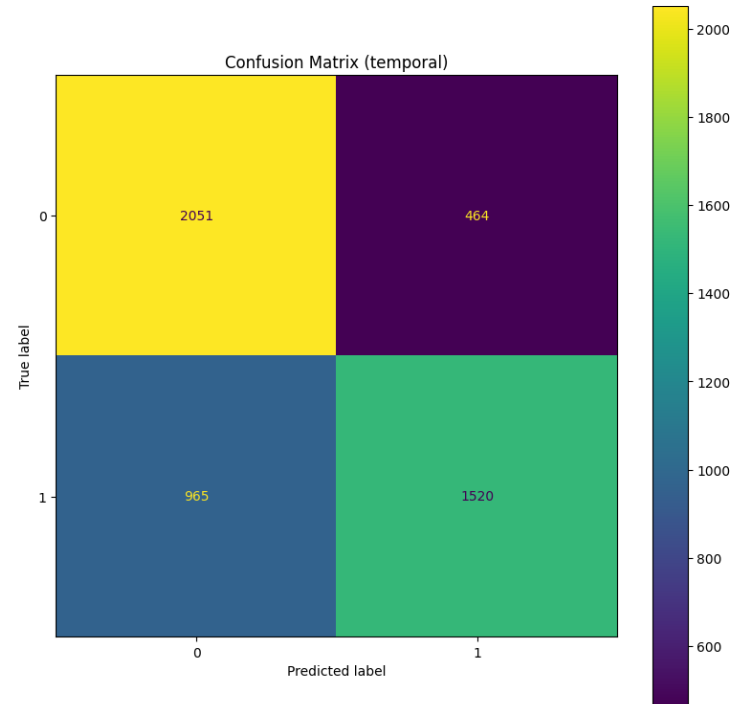
	precision	recall	f1-score	support
0	0.64	0.80	0.71	2512
1	0.73	0.55	0.63	2488
accuracy			0.68	5000
macro avg	0.69	0.67	0.67	5000
weighted avg	0.69	0.68	0.67	5000

LogisticRegression: ROC AUC=0.758



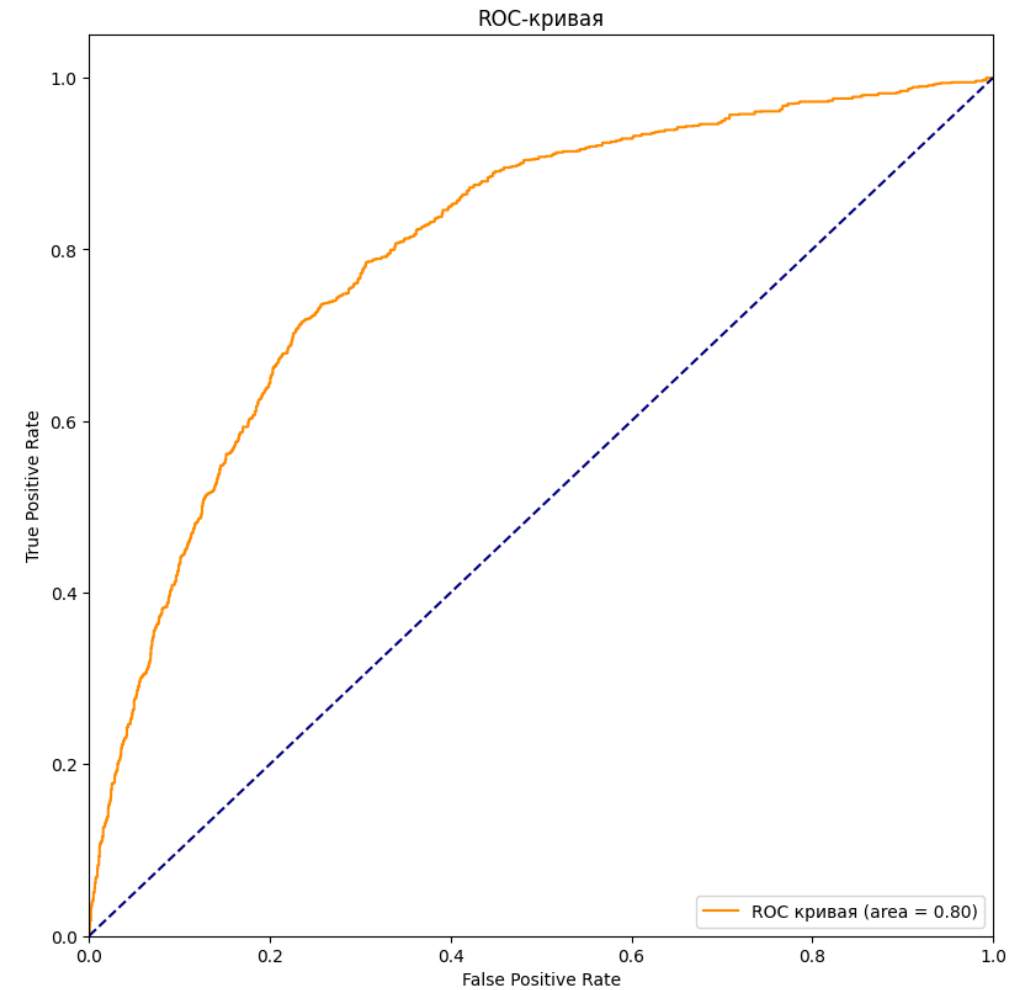
В статье: AUC=0.731

УС. Признаки II-A



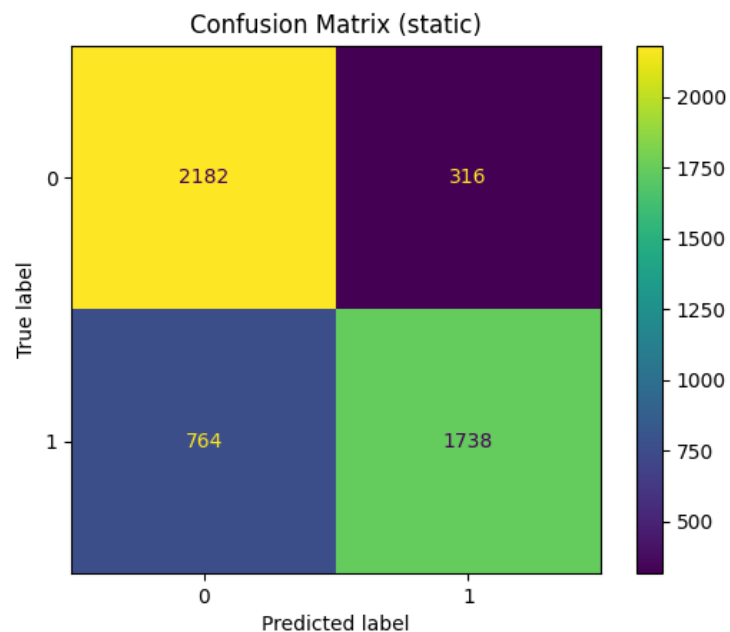
	precision	recall	f1-score	support
0	0.68	0.82	0.74	2515
1	0.77	0.61	0.68	2485
accuracy			0.71	5000
macro avg	0.72	0.71	0.71	5000
weighted avg	0.72	0.71	0.71	5000

LogisticRegression: ROC AUC=0.801

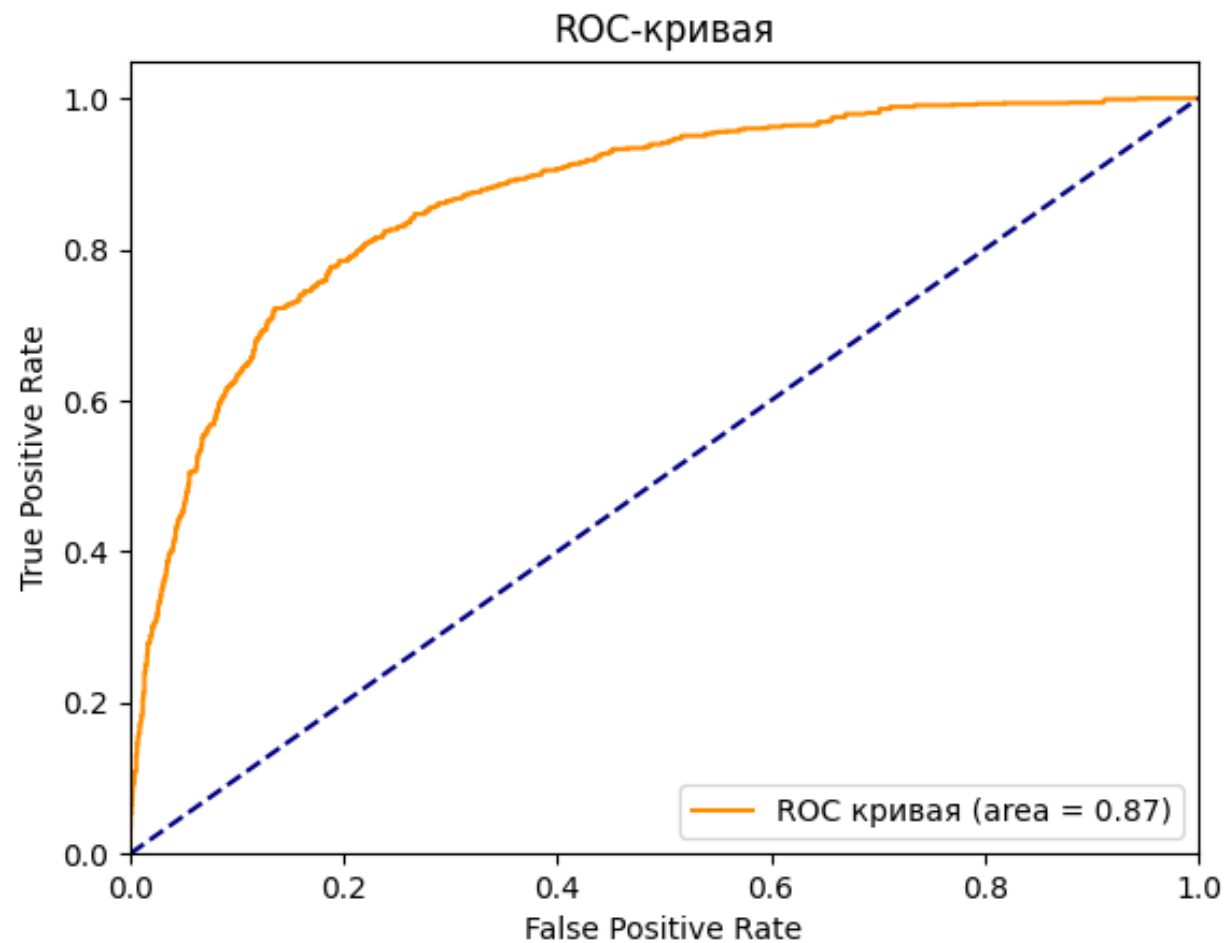


В статье: AUC=0.893

BitA. Признаки I

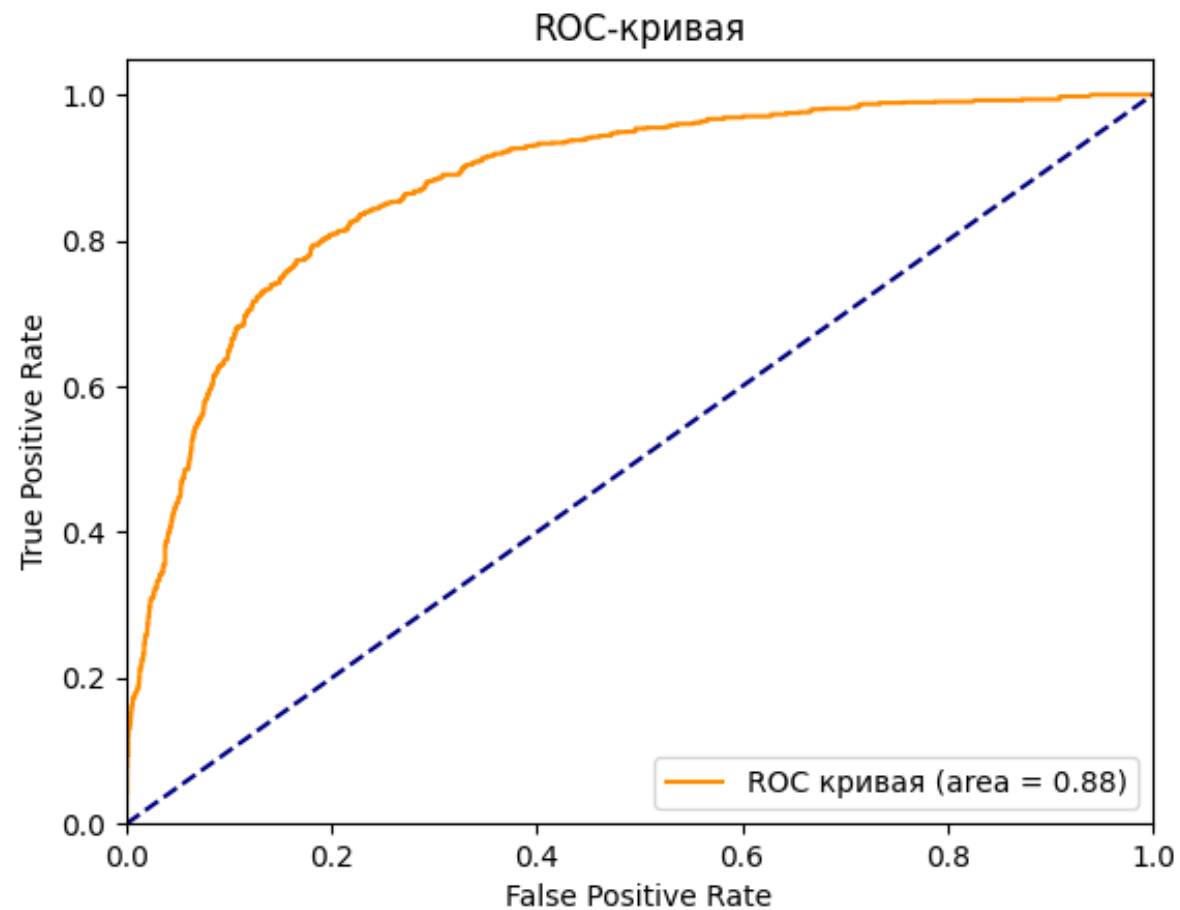
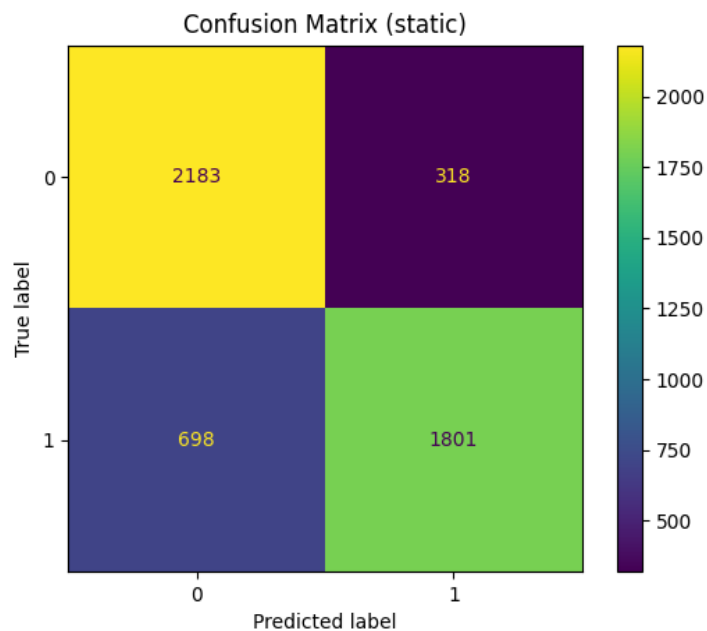


	precision	recall	f1-score	support
0	0.74	0.87	0.80	2498
1	0.85	0.69	0.76	2502
accuracy			0.78	5000
macro avg	0.79	0.78	0.78	5000
weighted avg	0.79	0.78	0.78	5000



В статье: AUC=0.868

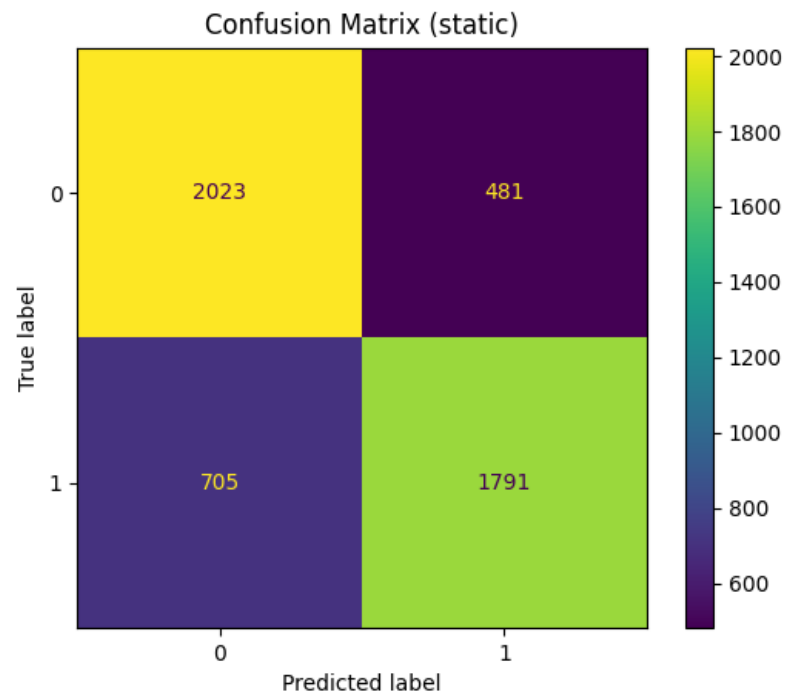
BitA. Признаки II-A



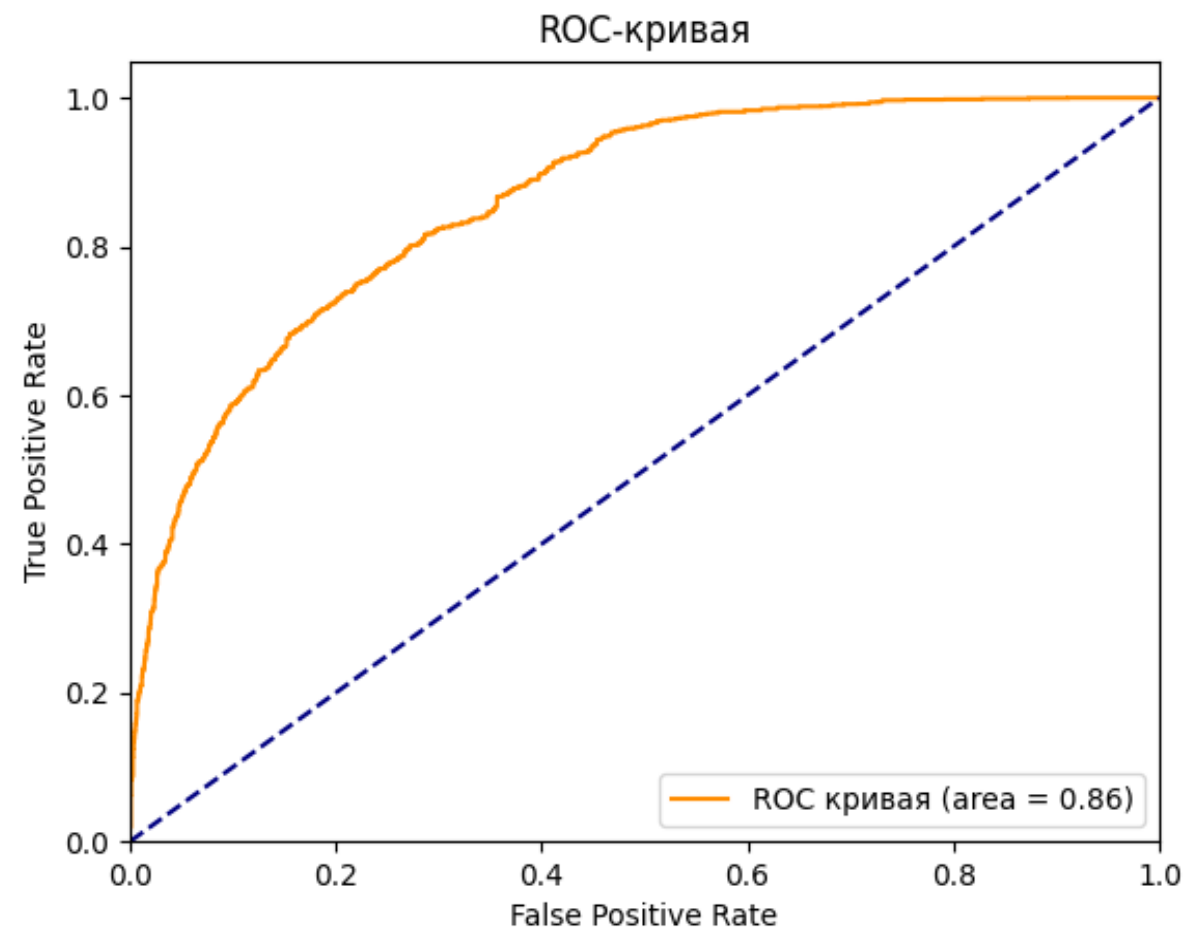
	precision	recall	f1-score	support
0	0.76	0.87	0.81	2501
1	0.85	0.72	0.78	2499
accuracy			0.80	5000
macro avg	0.80	0.80	0.80	5000
weighted avg	0.80	0.80	0.80	5000

В статье: AUC=0.945

BitOT. Признаки I

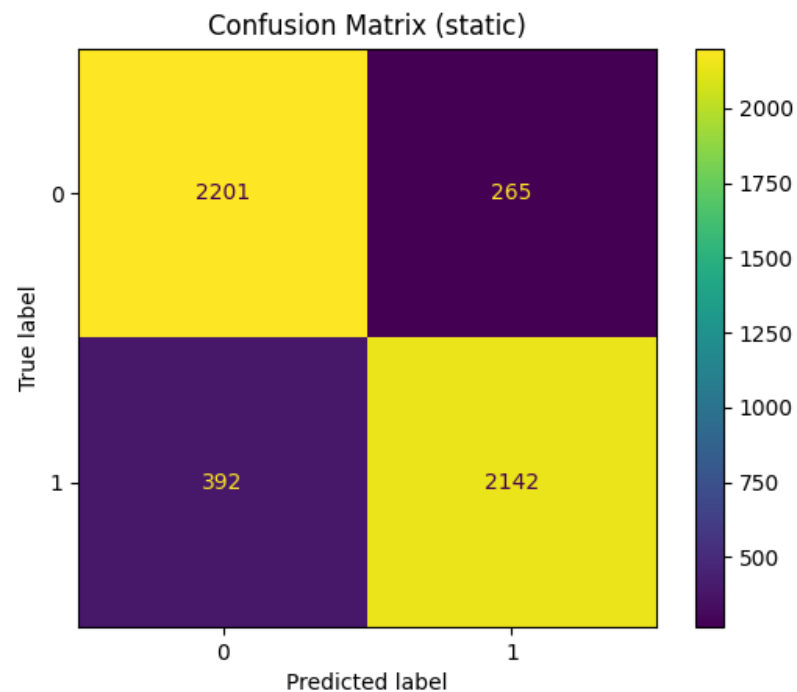


	precision	recall	f1-score	support
0	0.74	0.81	0.77	2504
1	0.79	0.72	0.75	2496
accuracy			0.76	5000
macro avg	0.76	0.76	0.76	5000
weighted avg	0.76	0.76	0.76	5000

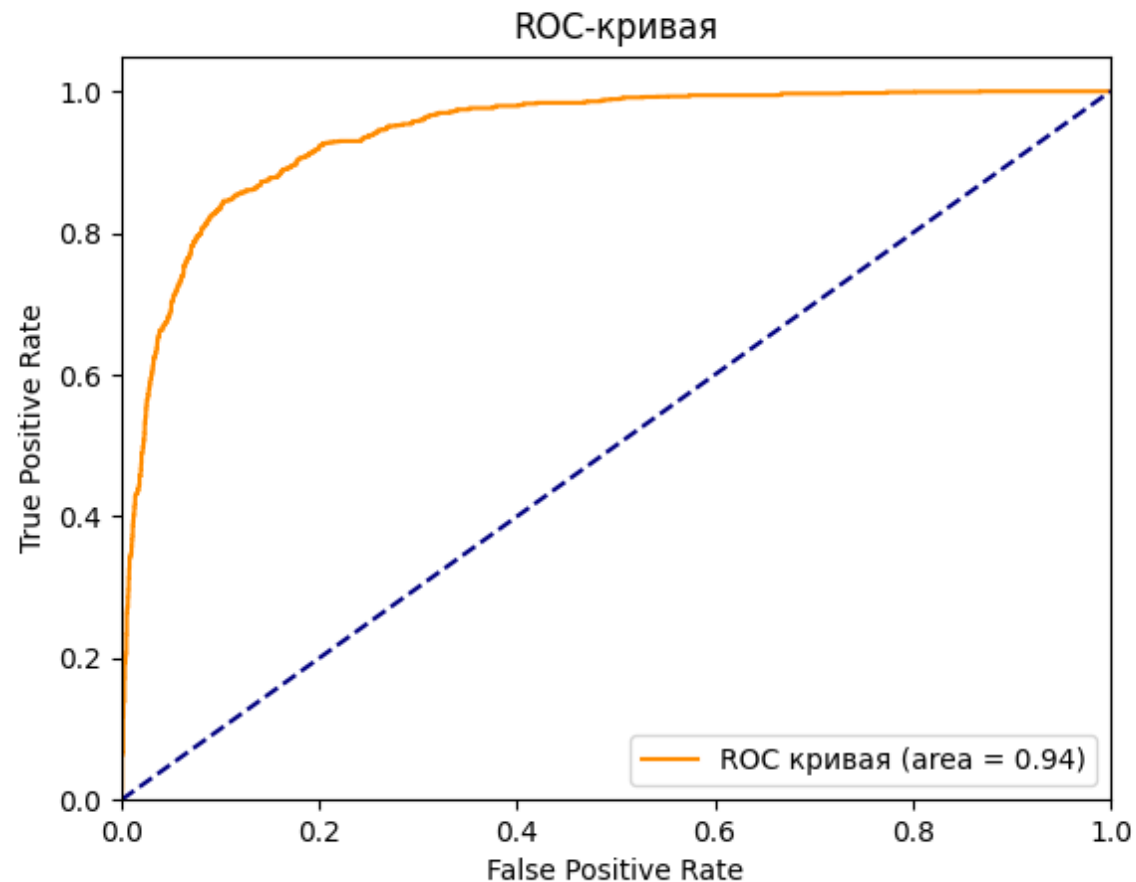


В статье: AUC=0.821

BitOT. Признаки II-A

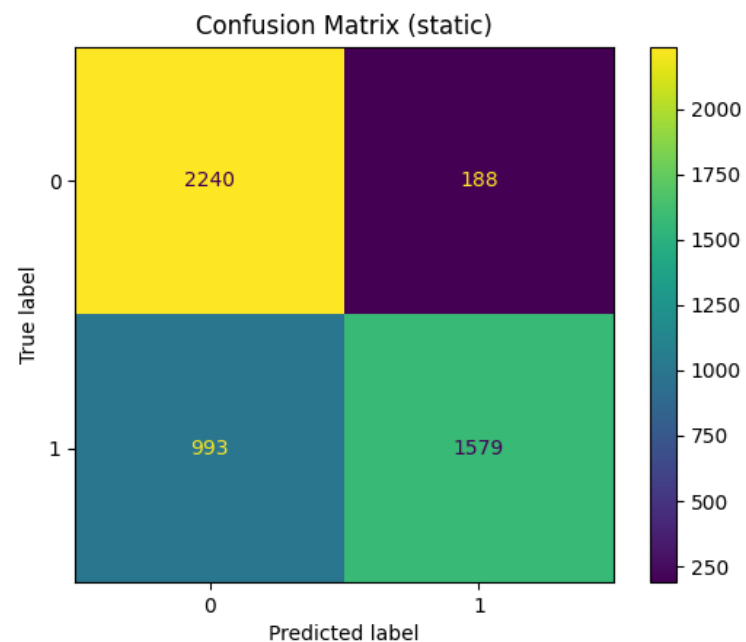


	precision	recall	f1-score	support
0	0.85	0.89	0.87	2466
1	0.89	0.85	0.87	2534
accuracy			0.87	5000
macro avg	0.87	0.87	0.87	5000
weighted avg	0.87	0.87	0.87	5000

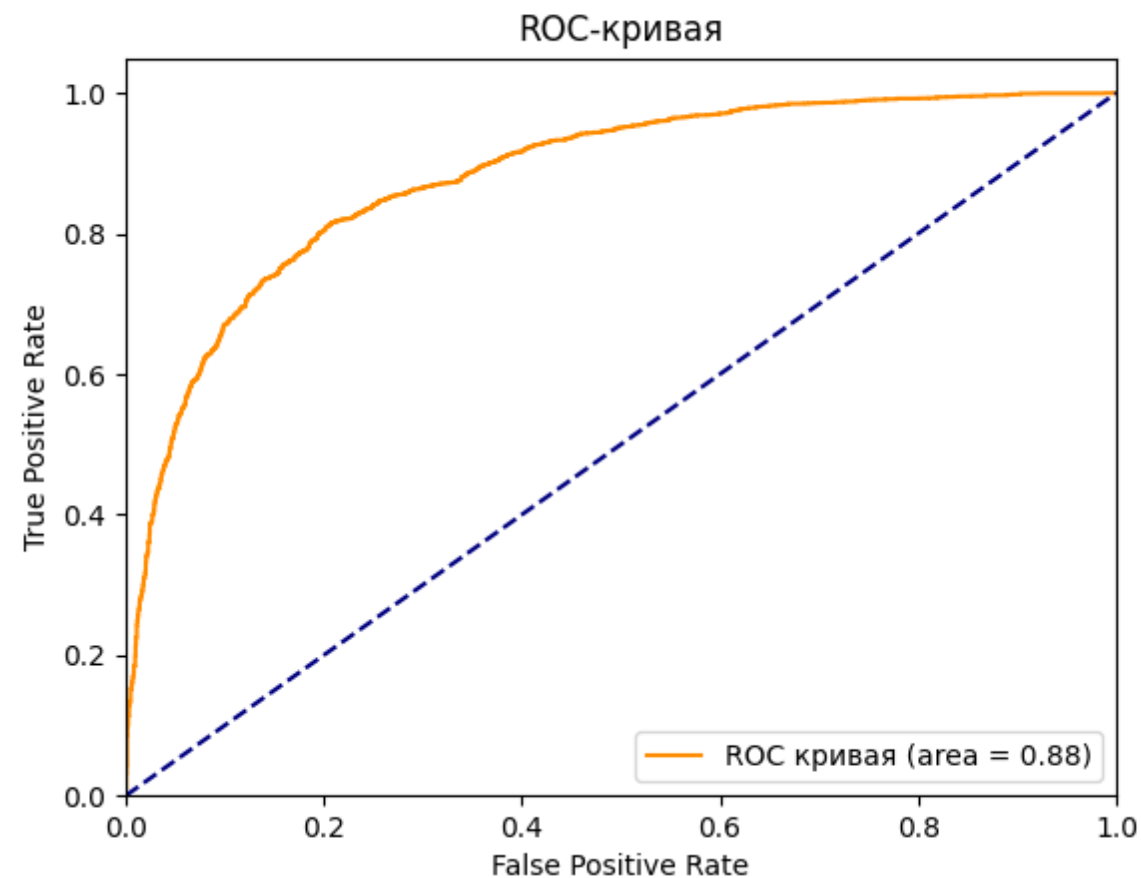


В статье: AUC=0.947

SX-МО. Признаки I

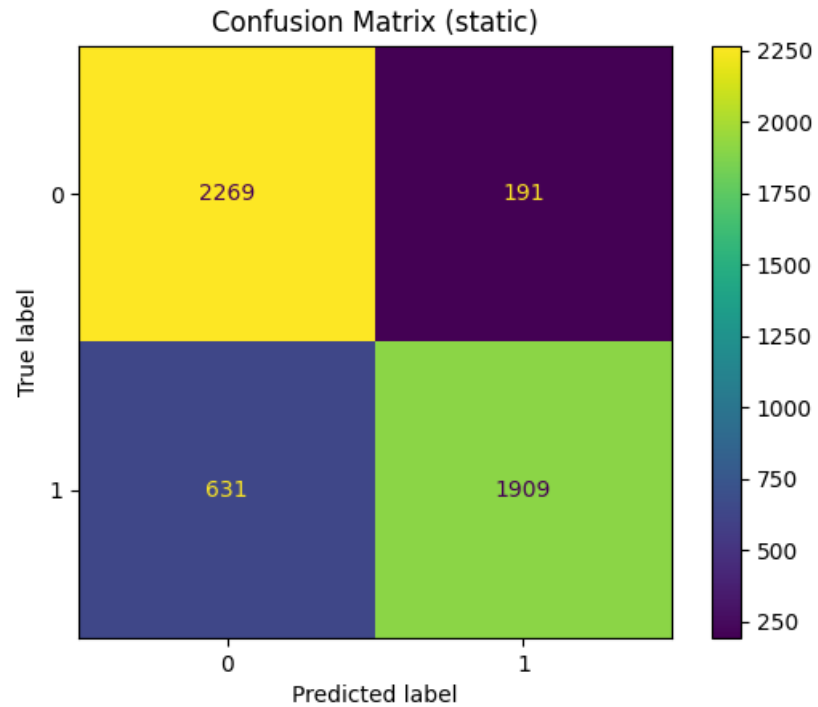


	precision	recall	f1-score	support
0	0.69	0.92	0.79	2428
1	0.89	0.61	0.73	2572
accuracy			0.76	5000
macro avg	0.79	0.77	0.76	5000
weighted avg	0.80	0.76	0.76	5000

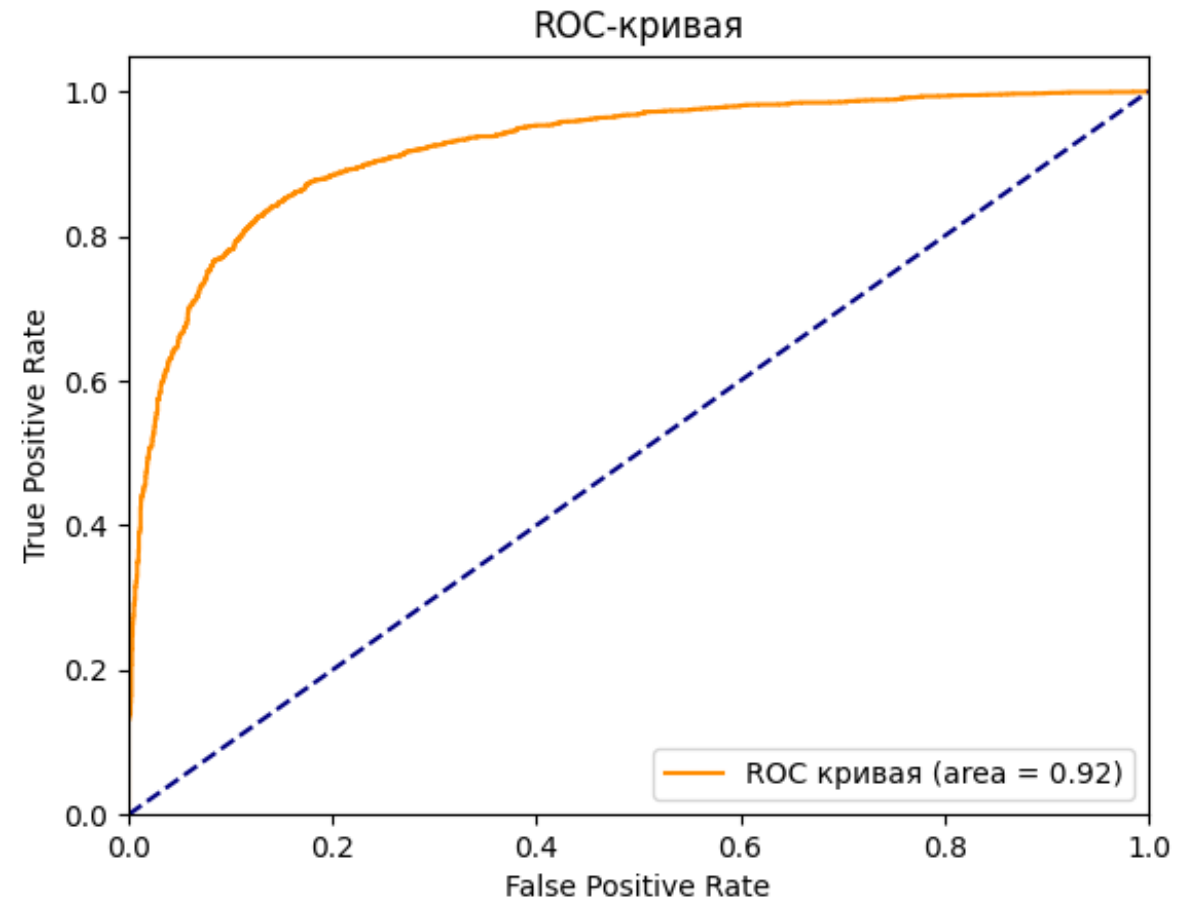


В статье: AUC=0.859

SX-МО. Признаки II-A



	precision	recall	f1-score	support	
0	0.78	0.92	0.85	2460	
1	0.91	0.75	0.82	2540	
accuracy				0.84	5000
macro avg	0.85	0.84	0.83	5000	
weighted avg	0.85	0.84	0.83	5000	



В статье: AUC=0.944