TABLE IV. Comparison of Average Rank on $N = 5$. Bold indicates the best performance. Asterisks in the row below STSMOTE denote statistical significance levels compared to all other methods (Holm-corrected Wilcoxon signed-rank test): $^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

Method	Web				XSS				DDoS				SQL				Avg.
	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	
No Aug.	7.96	7.76	8.92	7.50	7.90	8.12	8.18	7.32	8.64	8.40	8.84	5.44	8.74	8.62	8.92	8.06	8.08
RO	2.98	6.04	5.30	4.16	4.28	6.38	5.50	5.52	5.68	6.98	6.76	5.48	3.42	5.64	5.90	5.30	5.33
SMOTE	5.14	5.52	5.16	5.10	5.06	5.52	5.96	5.00	4.14	3.84	4.52	4.10	4.72	4.44	4.18	4.26	4.79
BSMOTE	4.74	5.28	5.80	3.96	4.18	4.80	4.84	3.52	4.48	3.28	4.30	3.84	4.66	4.58	5.26	4.06	4.47
ADASYN	3.30	2.88	4.80	3.46	3.16	3.82	6.04	3.14	5.00	4.96	5.78	3.52	3.42	3.74	4.26	2.64	4.00
CTGAN	8.06	6.22	4.26	7.24	7.86	5.66	3.30	7.40	5.24	5.00	3.72	7.50	7.86	6.66	4.80	7.40	6.14
TVAE	7.58	6.20	4.70	7.44	7.40	4.26	3.82	6.90	2.72	3.28	2.24	7.32	6.88	5.88	4.96	7.72	5.58
TabSyn	4.02	4.02	4.26	4.34	3.20	5.24	4.78	4.20	6.54	7.16	5.60	4.08	3.42	4.24	4.78	3.68	4.60
STSMOTE	**1.22**	**1.08**	**1.80**	**1.80**	**1.96**	**1.20**	**2.58**	**2.00**	**2.56**	**2.10**	3.24	3.72	**1.88**	**1.20**	**1.94**	**1.88**	**2.01**
(Sig.)	***	***	**	*	***	***	-	*	-	**	-	-	**	***	***	-	

TABLE V. Raw p Value on $N = 5$.

Method	Web				XSS				DDoS				SQL			
	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM
No Aug.	1.2e-5	6.0e-8	6.0e-8	6.0e-8	2.1e-5	1.2e-5	6.0e-8	3.1e-5	6.0e-8	6.0e-8	1.2e-5	1.1e-2	6.0e-8	1.2e-5	1.2e-5	6.0e-8
RO	2.9e-4	6.0e-8	4.5e-5	3.8e-4	5.6e-5	1.2e-5	1.8e-4	2.7e-5	3.0e-7	1.8e-5	3.0e-7	8.8e-5	1.0e-3	6.0e-8	3.6e-5	3.0e-5
SMOTE	1.8e-5	6.0e-8	1.0e-4	3.2e-5	3.4e-5	1.2e-5	3.5e-3	4.6e-5	2.7e-3	2.5e-4	1.6e-3	9.2e-1	3.6e-5	4.2e-7	1.4e-4	1.8e-4
BSMOTE	6.0e-8	6.0e-8	3.0e-7	1.1e-2	1.4e-4	1.4e-5	1.0e-3	4.2e-3	5.5e-4	3.8e-3	1.8e-3	4.7e-1	8.7e-5	1.8e-7	1.5e-5	4.3e-5
ADASYN	6.0e-8	2.0e-6	3.3e-6	2.5e-4	2.0e-4	8.1e-5	3.2e-5	1.2e-2	3.3e-4	5.4e-5	1.8e-5	8.8e-1	3.3e-4	1.8e-5	2.2e-5	8.3e-2
CTGAN	1.2e-5	6.0e-8	9.1e-4	6.0e-8	2.7e-5	6.0e-8	2.6e-1	2.7e-5	1.5e-5	6.6e-6	3.9e-2	8.2e-6	6.0e-8	6.0e-8	2.0e-6	6.0e-8
TVAE	6.0e-8	6.0e-8	6.3e-4	6.0e-8	2.1e-5	1.2e-5	3.6e-2	2.7e-5	1.4e-1	3.4e-3	4.3e-1	2.2e-5	6.0e-8	6.0e-8	2.4e-5	1.2e-7
TabSyn	8.2e-6	6.0e-8	8.8e-5	1.6e-4	4.3e-4	1.6e-5	8.1e-3	9.8e-4	4.2e-6	1.2e-7	3.2e-5	9.3e-1	7.1e-4	1.8e-5	4.2e-6	2.0e-2

TABLE VI. Comparison of Average Rank on $N = 2$. Bold indicates the best performance. Asterisks in the row below STSMOTE denote statistical significance levels compared to all other methods (Holm-corrected Wilcoxon signed-rank test): $^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. † Statistical significance is calculated based on pairwise F1-score distributions. In this specific case, STSMOTE significantly outperforms other methods in F1-score despite a lower average rank caused by relative performance fluctuations among baselines.

Method	Web				XSS				DDoS				SQL				Avg.
	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	
No Aug.	8.12	7.74	8.38	7.88	7.04	6.26	8.26	7.26	8.40	8.00	8.92	6.54	8.34	7.98	8.28	8.00	7.84
RO	2.58	3.94	5.10	4.50	**3.18**	6.16	4.68	4.98	4.52	5.18	5.60	4.14	2.94	4.08	5.58	4.70	4.49
SMOTE	4.58	4.74	4.14	3.22	3.96	4.98	4.32	3.48	4.62	5.54	4.82	4.38	3.72	4.70	3.74	3.04	4.25
BSMOTE	7.52	7.94	7.74	7.62	6.80	7.32	6.90	7.12	4.60	5.10	4.90	4.10	8.10	7.62	7.88	7.56	6.80
ADASYN	4.26	3.96	5.26	3.54	4.28	3.84	6.04	3.58	5.10	4.90	5.36	4.90	4.28	4.78	4.32	3.04	4.47
CTGAN	7.16	5.72	3.74	6.50	6.46	5.64	3.88	6.02	5.44	5.46	3.96	7.12	7.14	6.04	4.20	6.32	5.68
TVAE	5.74	4.76	3.34	5.98	5.78	4.62	3.48	5.16	**2.16**	3.88	**2.48**	6.50	5.64	4.26	3.74	6.68	4.64
TabSyn	3.16	4.40	4.82	4.14	4.26	5.08	5.02	4.72	6.72	5.30	5.32	3.96	3.36	4.36	5.68	4.22	4.66
STSMOTE	**1.88**	**1.80**	**2.48**	**1.62**	3.24†	**1.10**	**2.42**	**2.68**	3.44	**1.64**	3.64	3.36	**1.48**	**1.18**	**1.58**	**1.44**	**2.19**
(Sig.)	*	***	-	*	*	***	-	-	-	**	-	-	**	***	***	**	

TABLE VII. Raw p Value on $N = 2$.

Method	Web				XSS				DDoS				SQL			
	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM	Cat	XGB	LGB	TabM
No Aug.	1.2e-5	2.7e-5	2.7e-5	6.0e-8	1.1e-4	1.1e-5	1.2e-5	2.9e-4	1.5e-6	1.2e-7	6.0e-8	2.6e-4	6.0e-8	6.0e-8	1.2e-5	6.0e-8
RO	3.7e-2	2.0e-4	4.7e-4	4.9e-4	4.2e-2	1.1e-5	1.2e-3	4.4e-4	7.5e-2	6.0e-7	1.6e-2	8.5e-2	2.1e-3	3.0e-7	1.8e-5	1.8e-7
SMOTE	3.1e-5	2.3e-4	8.2e-4	5.6e-3	7.8e-4	1.2e-5	2.2e-3	1.3e-1	8.5e-2	8.3e-7	2.0e-1	8.5e-2	4.0e-5	1.8e-5	1.3e-4	4.2e-3
BSMOTE	2.1e-5	6.0e-7	4.6e-5	6.0e-8	2.0e-4	1.2e-5	8.8e-5	2.9e-4	8.0e-2	2.6e-6	2.0e-1	1.5e-1	6.0e-8	6.0e-8	1.2e-5	6.0e-8
ADASYN	1.8e-4	1.0e-4	4.0e-5	1.6e-2	5.1e-3	1.1e-5	6.9e-5	4.7e-2	6.7e-2	1.5e-6	7.1e-2	4.0e-2	4.0e-5	1.8e-5	1.3e-4	4.9e-4
CTGAN	1.0e-4	6.0e-5	1.8e-1	1.8e-5	1.7e-4	1.8e-5	5.3e-2	4.4e-4	1.6e-2	1.5e-5	7.1e-1	1.0e-5	1.2e-5	6.0e-8	4.6e-5	6.0e-8
TVAE	1.3e-4	1.0e-4	1.2e-1	2.7e-5	4.7e-4	1.8e-5	1.1e-1	4.4e-4	9.8e-1	3.1e-3	2.6e-1	3.8e-5	1.8e-5	2.6e-6	2.1e-4	6.0e-8
TabSyn	1.9e-2	3.8e-4	3.9e-3	1.4e-3	3.6e-4	1.1e-5	1.2e-3	2.9e-4	4.6e-3	8.3e-7	4.8e-2	2.0e-1	5.2e-5	1.8e-5	2.7e-5	6.0e-8