Appendix E. Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 1 attacks from the perspective of evaluation metrics for the Nemenyi test.

In general, XGBoost (eXtreme Gradient-Boosting) was the algorithm that presented the best overall results in all metrics. In the Precision and FAR metrics, RF (Random Forest) achieved similar performances to XGBoost, getting the second best overall performance. The kNN (k-Nearest Neighbors) showed better results than the LightGBM (Light Gradient-Boosting Model) in detecting the specific attacks of Reconnaissance (UNSW-NB15) and Portscan (CIC-IDS2017). LightGBM was the algorithm that presented the worst results. Next, the classification algorithms' detailed results are presented from each analyzed metric's perspective.

Precision (Table E.1): the XGBoost and RF algorithms presented significant statistical differences when compared to the LightGBM and kNN algorithms, presenting the same distance values concerning the critical limit margin, with 33% and 26% above the value of the critical distance, for the first and second largest distances, respectively. In the other comparisons, all values were within the critical limit margin, varying between 0% and 7% of the critical distance value.

Critical Distance: 1.35						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
RF	1.62	XGB	1.62	0.00	0%	Equal
RF	1.62	KNN	3.33	1.71	126%	Different
RF	1.62	LGBM	3.42	1.80	133%	Different
XGB	1.62	KNN	3.33	1.71	126%	Different
XGB	1.62	LGBM	3.42	1.80	133%	Different
KNN	3.33	LGBM	3.42	0.09	7%	Equal

Table E.1: Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 1 attacks from the perspective of **Precision metric** for the Nemenyi test.

Recall (Table E.2): the XGBoost algorithm presented a significant statistical difference when compared to the LightGBM algorithm, with the first greater distance, 35% above the critical limit margin. The kNN with the second largest distance showed a significant difference when compared to LightGBM, with 29% above the critical distance value. The third largest distance was with RF compared to LightGBM, which showed a significant difference with only 5% above the critical distance value. In the other comparisons, all values were within the critical limit margin, varying between 6% and 30% of the critical distance value.

Critical Distance: 1.35						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
XGB	1.92	KNN	2.00	0.08	6%	Equal
XGB	1.92	RF	2.33	0.41	30%	Equal
XGB	1.92	LGBM	3.75	1.83	135%	Different
KNN	2.00	RF	2.33	0.33	24%	Equal
KNN	2.00	LGBM	3.75	1.75	129%	Different
RF	2.33	LGBM	3.75	1.42	105%	Different

Table E.2: Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 1 attacks from the perspective of **Recall metric** for the Nemenyi test.

F1-Score (Table E.3): the XGBoost algorithm presented a significant statistical difference compared to the LightGBM algorithm, with the first greater distance, with 60% above the critical limit margin. The second and third largest distances were within the critical limit margin. They did not present significant differences, with 92% of the critical distance value, when comparing XGBoost with kNN and RF with LightGBM, respectively. In the other comparisons, all values were within the critical limit margin, varying between 25% and 67% of the critical distance value.

Critical Distance: 1.35						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
XGB	1.42	RF	2.33	0.91	0,67	Equal
XGB	1.42	KNN	2.67	1.25	0,92	Equal
XGB	1.42	LGBM	3.58	2.16	1,60	Different
RF	2.33	KNN	2.67	0.34	0,25	Equal
RF	2.33	LGBM	3.58	1.25	0,92	Equal
KNN	2.67	LGBM	3.58	0.91	0,67	Equal

Table E.3: Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 1 attacks from the perspective of **F1-Score metric** for the Nemenyi test.

FAR (Table E.4): the XGBoost and RF algorithms showed significant statistical differences when compared to the LightGBM and kNN algorithms, showing the same distances concerning the critical limit margin, with 32% and 14% above the critical distance value, for the first and second largest distances, respectively. In the other comparisons, all values were within the critical limit margin, varying between 0% and 18% of the critical distance value.

Critical Distance: 1.35						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
XGB	1.54	RF	1.79	0.25	18%	Equal
XGB	1.54	LGBM	3.33	1.79	132%	Different
XGB	1.54	KNN	3.33	1.79	132%	Different
RF	1.79	LGBM	3.33	1.54	114%	Different
RF	1.79	KNN	3.33	1.54	114%	Different
LGBM	3.33	KNN	3.33	0.00	0%	Equal

Table E.4: Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 1 attacks from the perspective of **FAR metric** for the Nemenyi test.