Appendix F. Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 2 attacks from the perspective of evaluation metrics for the Nemenyi test.

In general, XGBoost (eXtreme Gradient-Boosting) was the algorithm that presented the best results, obtaining the first three most considerable distances from the perspective of the Recall, F1-Score, and FAR metrics. RF (Random Forest) had the second most significant distance in overall performance, standing out in the Precision metric. kNN (k-Nearest Neighbors) performed better than LightGBM (Light Gradient-Boosting Model) in detecting specific DoS (UNSW-NB15) and DoS/DDoS (CIC-IDS2017) attacks. 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 F.1): the RF algorithm showed a significant statistical difference when compared to the LightGBM algorithm, with the first greater distance, 41% above the critical limit margin. The XGBoost, with the second largest distance, showed a difference compared to LightGBM, with 26% above the critical distance value. The third largest distance is kNN compared to LightGBM, which presented a distance within the critical limit margin, with 96% of the critical distance value. In the other comparisons, all values were within the critical limit margin, varying between 15% and 45% of the critical distance value.

Critical Distance: 1.10						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
RF	1.94	XGB	2.11	0.17	15%	Equal
RF	1.94	KNN	2.44	0.50	45%	Equal
RF	1.94	LGBM	3.50	1.56	141%	Different
XGB	2.11	KNN	2.44	0.33	30%	Equal
XGB	2.11	LGBM	3.50	1.39	126%	Different
KNN	2.44	LGBM	3.50	1.06	96%	Equal

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

Recall (Table F.2): the XGBoost algorithm presented a significant statistical difference when compared to the LightGBM algorithm, with the first greater distance, with 46% above the critical limit margin. The kNN with the second largest distance showed a difference compared to LightGBM, with 41% above the critical distance value. The third largest distance is with the RF, compared to the LGBM, which presented a distance within the critical limit margin, with 96% of the critical distance value. In the other comparisons, all values were within the critical limit margin, varying between 5% and 51% of the critical distance value.

Critical Distance: 1.10						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
XGB	1.94	KNN	2.00	0.06	5%	Equal
XGB	1.94	RF	2.50	0.56	51%	Equal
XGB	1.94	LGBM	3.56	1.62	146%	Different
KNN	2.00	RF	2.50	0.50	45%	Equal
KNN	2.00	LGBM	3.56	1.56	141%	Different
RF	2.50	LGBM	3.56	1.06	96%	Equal

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

F1-Score (Table F.3): the XGBoost algorithm presented a significant statistical difference compared to the LightGBM algorithm, with the first greater distance and 51% above the critical limit margin. The kNN with the second largest distance showed a difference compared to LightGBM, with 40% above the critical distance value. The third largest distance is with the RF compared to LightGBM, which showed a statistically significant difference, with 10% above the value of the critical distance. In the other comparisons, all values were within the critical limit margin, varying between 11% and 41% of the critical distance value.

Critical Distance: 1.10						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
XGB	1.94	KNN	2.06	0.12	11%	Equal
XGB	1.94	RF	2.39	0.45	41%	Equal
XGB	1.94	LGBM	3.61	1.67	151%	Different
KNN	2.06	RF	2.39	0.33	30%	Equal
KNN	2.06	LGBM	3.61	1.55	140%	Different
RF	2.39	LGBM	3.61	1.22	110%	Different

Table F.3: Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 2 attacks from the perspective of **F1-Score metric** for the Nemenyi test.

FAR (Table F.4): the XGBoost algorithm presented the most significant statistical differences compared to the LightGBM algorithm, having the first, second, and third largest distances, with 86%, 51%, and 26% above the critical limit margin, respectively. In the other comparisons, all values were within the critical limit margin, varying between 25% and 61% of the critical distance value.

Critical Distance: 1.10						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
XGB	1.22	RF	2.61	1.39	126%	Different
XGB	1.22	KNN	2.89	1.67	151%	Different
XGB	1.22	LGBM	3.28	2.06	186%	Different
RF	2.61	KNN	2.89	0.28	25%	Equal
RF	2.61	LGBM	3.28	0.67	61%	Equal
KNN	2.89	LGBM	3.28	0.39	35%	Equal

Table F.4: Comparison of classification algorithms using data preprocessing techniques in multiclass models of specific Group 2 attacks from the perspective of **FAR metric** for the Nemenyi test.