Appendix A. Comparison of data preprocessing techniques in binary class models from the perspective of evaluation metrics for the Nemenyi test.

The results indicate the LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques with the best performance compared to the other techniques. LE_SS is the only one that presented a significant statistical difference in the perspective of two analyzed metrics, Precision and FAR. While in the F1-Score metric, the LE_SS showed values very close to the critical distance limit margin. And in the Recall metric, the LE_SS remained within the critical distance limit margin with values close to the other techniques, with no significant statistical difference. Below, the results are presented from the perspective of each analyzed metric.

Precision (Table A.1): the LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques were the only ones that presented a significant statistical difference when compared to the GD_SS (Data Cleaning, Get Dummies and Standard Scaler) techniques, with the first greater distance, with 8% above the critical margin limit. All other comparisons are within the critical limit. The second largest distance was very close to the margin, with the LE_SS techniques compared to the GD (Data Cleaning and Get Dummies) techniques, with 87% of the critical distance value. At the third largest distance, the LE_SS and LE_MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques were aligned when compared to the LE (Data Cleaning and Label Encoder) and GD_SS (Data Cleaning, Get Dummies and Standard Scaler) techniques, respectively, both with 75% of the critical distance value. Except for the GD_MM techniques (Data Cleaning, Get Dummies, and MinMax Scaler) with 73% of the critical distance value, when compared to the GD_SS techniques, all other techniques, when compared, presented percentages that varied from 2% to 54% of the value of the critical distance.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
LE_SS	2.00	LE_MM	2.88	0.88	33%	Equal
LE_SS	2.00	GD_MM	2.94	0.94	35%	Equal
LE_SS	2.00	LE	4.00	2.00	75%	Equal
LE_SS	2.00	GD	4.31	2.31	87%	Equal
LE_SS	2.00	GD_SS	4.88	2.88	108%	Different
LE_MM	2.88	GD_MM	2.94	0.06	2%	Equal
LE_MM	2.88	LE	4.00	1.12	42%	Equal
LE_MM	2.88	GD	4.31	1.43	54%	Equal
LE_MM	2.88	GD_SS	4.88	2.00	75%	Equal
GD_MM	2.94	LE	4.00	1.06	40%	Equal
GD_MM	2.94	GD	4.31	1.37	51%	Equal
GD_MM	2.94	GD_SS	4.88	1.94	73%	Equal
LE	4.00	GD	4.31	0.31	12%	Equal
LE	4.00	GD_SS	4.88	0.88	33%	Equal
GD	4.31	GD_SS	4.88	0.57	21%	Equal

Table A.1: Comparison of data preprocessing techniques in binary class models from the perspective of the **Precision metric** for the Nemenyi test.

Recall (Table A.2): all techniques, when compared, showed a percentage value within the critical limit margin. The LE_MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques presented the first greater distance, with 40% of the critical distance value, compared to the GD (Data Cleaning and Get Dummies) techniques. The second largest distance was also with the LE_MM techniques compared to the LE (Data Cleaning and Label Encoder) techniques, with 38%

of the critical distance value. The third largest distance was for LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques compared to GD (Data Cleaning and Get Dummies) techniques, with 33% of the critical distance value. When compared, all the other techniques presented percentages ranging from 0% to 31% of the critical distance value.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
LE_MM	2.94	LE_SS	3.12	0.18	7%	Equal
LE_MM	2.94	GD_SS	3.50	0.56	21%	Equal
LE_MM	2.94	GD_MM	3.50	0.56	21%	Equal
LE_MM	2.94	LE	3.94	1.00	38%	Equal
LE_MM	2.94	GD	4.00	1.06	40%	Equal
LE_SS	3.12	GD_SS	3.50	0.38	14%	Equal
LE_SS	3.12	GD_MM	3.50	0.38	14%	Equal
LE_SS	3.12	LE	3.94	0.82	31%	Equal
LE_SS	3.12	GD	4.00	0.88	33%	Equal
GD_SS	3.50	GD_MM	3.50	0.00	0%	Equal
GD_SS	3.50	LE	3.94	0.44	17%	Equal
GD_SS	3.50	GD	4.00	0.50	19%	Equal
GD_MM	3.50	LE	3.94	0.44	17%	Equal
GD_MM	3.50	GD	4.00	0.50	19%	Equal
LE	3.94	GD	4.00	0.06	2%	Equal

Table A.2: Comparison of data preprocessing techniques in binary class models from the perspective of the **Recall metric** for the Nemenyi test.

F1-Score (Table A.3): all techniques, when compared, also showed a percentage values within the critical limit margin. The LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques presented the three most considerable distances when compared to the GD_SS (Data Cleaning, Get Dummies and Standard Scaler), LE (Data Cleaning and Label Encoder) and GD (Data Cleaning and Get Dummies) techniques), with 87%, 85% and 73% of the critical distance value, respectively. The first and second largest distances are close to the critical limit margin. Except for the LE_MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques when compared to the GD_SS (Data Cleaning, Get Dummies and Standard Scaler) and LE (Data Cleaning and Label Encoder) techniques, with 63% and 61% of the value critical distance, respectively, all the other techniques when compared, presented percentages that varied from 2% to 49% of the value of the critical distance.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
LE_SS	2.12	LE_MM	2.75	0.63	24%	Equal
LE_SS	2.12	GD_MM	3.25	1.13	42%	Equal
LE_SS	2.12	GD	4.06	1.94	73%	Equal
LE_SS	2.12	LE	4.38	2.26	85%	Equal
LE_SS	2.12	GD_SS	4.44	2.32	87%	Equal
LE_MM	2.75	GD_MM	3.25	0.50	19%	Equal
LE_MM	2.75	GD	4.06	1.31	49%	Equal
LE_MM	2.75	LE	4.38	1.63	61%	Equal
LE_MM	2.75	GD_SS	4.44	1.69	63%	Equal
GD_MM	3.25	GD	4.06	0.81	30%	Equal
GD_MM	3.25	LE	4.38	1.13	42%	Equal
GD_MM	3.25	GD_SS	4.44	1.19	45%	Equal

GD	4.06	LE	4.38	0.32	12%	Equal
GD	4.06	GD_SS	4.44	0.38	14%	Equal
LE	4.38	GD_SS	4.44	0.06	2%	Equal

Table A.3: Comparison of data preprocessing techniques in binary class models from the perspective of the **F1-Score metric** for the Nemenyi test.

FAR (Table A.4): the LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques were the only ones that showed a significant statistical difference when compared to the GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) techniques, with the first greater distance, 8% above the critical distance value. All the other techniques, when compared, presented distances within the range of the critical limit. The second largest distance was also with the LE_SS techniques compared to the GD (Data Cleaning and Get Dummies) techniques, with 78% of the value of the critical distance. The third largest distance was for the LE_MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques compared to the GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) techniques, with 75% of the critical distance value. Except for the GD_MM (Data Cleaning, Get Dummies, and MinMax Scaler) techniques with 73% of the critical distance value, when comparing the GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) techniques, all the other techniques, when compared, presented percentages that ranged from 2% to 56% of the critical distance value.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypothesis**
LE_SS	2.12	LE_MM	3.00	0.88	33%	Equal
LE_SS	2.12	GD_MM	3.06	0.94	35%	Equal
LE_SS	2.12	LE	3.62	1.50	56%	Equal
LE_SS	2.12	GD	4.19	2.07	78%	Equal
LE_SS	2.12	GD_SS	5.00	2.88	108%	Different
LE_MM	3.00	GD_MM	3.06	0.06	2%	Equal
LE_MM	3.00	LE	3.62	0.62	23%	Equal
LE_MM	3.00	GD	4.19	1.19	45%	Equal
LE_MM	3.00	GD_SS	5.00	2.00	75%	Equal
GD_MM	3.06	LE	3.62	0.56	21%	Equal
GD_MM	3.06	GD	4.19	1.13	42%	Equal
GD_MM	3.06	GD_SS	5.00	1.94	73%	Equal
LE	3.62	GD	4.19	0.57	21%	Equal
LE	3.62	GD_SS	5.00	1.38	52%	Equal
GD	4.19	GD_SS	5.00	0.81	30%	Equal

Table A.4: Comparison of data preprocessing techniques in binary class models from the perspective of the **FAR metric** for the Nemenyi test.

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

In general, no technique showed a significant statistical difference, all within the critical distance's limiting margin. The most considerable distance within the critical limit is considered low, presented by the LE_MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques, with 59% of the critical distance value from the perspective of the Recall metric. The GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) techniques showed the most significant distance from the standpoint of the Precision metric, with 33% of the critical distance value. The LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) and GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) techniques were aligned with the most significant distance in the F1-Score metric, with 47% of the critical distance value. And the LE (Data Cleaning and Label Encoder) techniques showed the most significant distance from the perspective of the FAR metric, with 26% of the critical distance value. As all distance values are close and within the limiting margin of the critical distance, it is impossible to state which techniques would be considered the best in terms of performance. Below, the results are presented from the perspective of each analyzed metric.

Precision (Table B.1): all techniques, when compared, presented a percentage value within the margin of the critical distance limit. The GD_SS (Data Cleaning, Get Dummies and Standard Scaler) techniques showed the most significant distance when compared to the GD (Data Cleaning and Get Dummies) techniques, with the first greater distance within the limit margin of the critical distance, with 33% of the value of the critical distance. In the second largest distance were the techniques (Data Cleaning, Label Encoder, and MinMax Scaler) compared to the GD techniques, with 28% of the value of the critical distance. The third largest distance is from the GD_SS techniques compared to the LE_SS techniques (Data Cleaning, Label Encoder, and Standard Scaler), with 24% of the critical distance value. When compared, all the other techniques presented percentages ranging from 0% to 19% of the critical distance value.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
GD_SS	3.06	LE_MM	3.19	0.13	5%	Equal
GD_SS	3.06	LE	3.56	0.50	19%	Equal
GD_SS	3.06	GD_MM	3.56	0.50	19%	Equal
GD_SS	3.06	LE_SS	3.69	0.63	24%	Equal
GD_SS	3.06	GD	3.94	0.88	33%	Equal
LE_MM	3.19	LE	3.56	0.37	14%	Equal
LE_MM	3.19	GD_MM	3.56	0.37	14%	Equal
LE_MM	3.19	LE_SS	3.69	0.50	19%	Equal
LE_MM	3.19	GD	3.94	0.75	28%	Equal
LE	3.56	GD_MM	3.56	0.00	0%	Equal
LE	3.56	LE_SS	3.69	0.13	5%	Equal
LE	3.56	GD	3.94	0.38	14%	Equal
GD_MM	3.56	LE_SS	3.69	0.13	5%	Equal
GD_MM	3.56	GD	3.94	0.38	14%	Equal
LE_SS	3.69	GD	3.94	0.25	9%	Equal

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

Recall (Table B.2): all techniques, when compared, showed a percentage value within the margin of the critical distance limit. The LE_MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques showed the most significant distance when compared to the GD_MM (Data Cleaning, Get Dummies, and MinMax Scaler) techniques, with the first greater distance within the limit margin of the critical distance, with 59% of the value of the critical distance. The second largest distance is also for the LE_MM group when comparing the LE (Data Cleaning and Label Encoder) and GD (Data Cleaning and Get Dummies) techniques, with 52% of the critical distance value. In the third largest distance were the LE_SS techniques (Data Cleaning, Label Encoder, and Standard Scaler) compared to the GD_MM techniques (Data Cleaning, Get Dummies, and MinMax Scaler), with a value of 49% of the critical distance value. When compared, all the other techniques presented percentages ranging from 0% to 42% of the critical distance value.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
LE_MM	2.62	LE_SS	2.88	0.26	10%	Equal
LE_MM	2.62	GD_SS	3.31	0.69	26%	Equal
LE_MM	2.62	LE	4.00	1.38	52%	Equal
LE_MM	2.62	GD	4.00	1.38	52%	Equal
LE_MM	2.62	GD_MM	4.19	1.57	59%	Equal
LE_SS	2.88	GD_SS	3.31	0.43	16%	Equal
LE_SS	2.88	LE	4.00	1.12	42%	Equal
LE_SS	2.88	GD	4.00	1.12	42%	Equal
LE_SS	2.88	GD_MM	4.19	1.31	49%	Equal
GD_SS	3.31	LE	4.00	0.69	26%	Equal
GD_SS	3.31	GD	4.00	0.69	26%	Equal
GD_SS	3.31	GD_MM	4.19	0.88	33%	Equal
LE	4.00	GD	4.00	0.00	0%	Equal
LE	4.00	GD_MM	4.19	0.19	7%	Equal
GD	4.00	GD_MM	4.19	0.19	7%	Equal

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

F1-Score (Table B.3): all techniques, when compared, presented a percentage value within the margin of the critical distance limit. The LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) and GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) techniques were aligned and presented the most significant distances within the limit margin of the critical distance when compared to the GD (Data Cleaning and Standard Scaler) techniques. Get Dummies), both at 47% of the critical distance value. In the second largest distance were the LE_MM techniques (Data Cleaning, Label Encoder, and MinMax Scaler) compared to the GD techniques, with 45% of the value of the critical distance. In the third largest distance, the LE_SS and GD_SS techniques were also aligned when compared to the GD_MM techniques (Data Cleaning, Get Dummies, and MinMax Scaler), with a value of 35% of the critical distance value. When compared, all the other techniques presented percentages ranging from 0% to 33% of the critical distance value.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
LE_SS	3.06	GD_SS	3.06	0.00	0%	Equal
LE_SS	3.06	LE_MM	3.12	0.06	2%	Equal
LE_SS	3.06	LE	3.44	0.38	14%	Equal
LE_SS	3.06	GD_MM	4.00	0.94	35%	Equal
LE_SS	3.06	GD	4.31	1.25	47%	Equal

GD_SS	3.06	LE_MM	3.12	0.06	2%	Equal
GD_SS	3.06	LE	3.44	0.38	14%	Equal
GD_SS	3.06	GD_MM	4.00	0.94	35%	Equal
GD_SS	3.06	GD	4.31	1.25	47%	Equal
LE_MM	3.12	LE	3.44	0.32	12%	Equal
LE_MM	3.12	GD_MM	4.00	0.88	33%	Equal
LE_MM	3.12	GD	4.31	1.19	45%	Equal
LE	3.44	GD_MM	4.00	0.56	21%	Equal
LE	3.44	GD	4.31	0.87	33%	Equal
GD_MM	4.00	GD	4.31	0.31	12%	Equal

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

FAR (Table B.4): all techniques, when compared, showed a percentage value within the margin of the critical distance limit. The LE (Data Cleaning and Label Encoder) techniques showed the most significant distance within the limit margin of the critical distance when compared to the GD_MM (Data Cleaning, Get Dummies, and MinMax Scaler) techniques, with 26% of the critical distance value. In the second largest distance were the LE_MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques compared to the GD_MM (Data Cleaning, Get Dummies, and MinMax Scaler) techniques, with a value of 24% of the critical distance value. The LE techniques were also in the third largest distance compared to the GD (Data Cleaning and Get Dummies) techniques, with 23% of the value of the critical distance. When compared, all the other techniques presented percentages ranging from 2% to 21% of the critical distance value.

Critical Distance: 2.66						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
LE	3.81	LE_MM	3.75	0.06	2%	Equal
LE	3.81	GD_SS	3.69	0.12	5%	Equal
LE	3.81	LE_SS	3.44	0.37	14%	Equal
LE	3.81	GD	3.19	0.62	23%	Equal
LE	3.81	GD_MM	3.12	0.69	26%	Equal
LE_MM	3.75	GD_SS	3.69	0.06	2%	Equal
LE_MM	3.75	LE_SS	3.44	0.31	12%	Equal
LE_MM	3.75	GD	3.19	0.56	21%	Equal
LE_MM	3.75	GD_MM	3.12	0.63	24%	Equal
GD_SS	3.69	LE_SS	3.44	0.25	9%	Equal
GD_SS	3.69	GD	3.19	0.50	19%	Equal
GD_SS	3.69	GD_MM	3.12	0.57	21%	Equal
LE_SS	3.44	GD	3.19	0.25	9%	Equal
LE_SS	3.44	GD_MM	3.12	0.32	12%	Equal
GD	3.19	GD_MM	3.12	0.07	3%	Equal

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

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

In general, no technique showed a significant statistical difference, with all techniques within the limiting margin of the critical distance. The most considerable distance within the critical limit is considered low, presented by the GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) techniques, with 59% of the critical distance value from the perspective of the Recall metric. The GD_SS techniques also obtained the highest from the perspective of the Precision metric, with a value of 38% of the critical distance value. The LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques had the most significant distance in the F1-Score and FAR metrics, with 38% and 50% of the critical distance value, respectively. As all the distance differences between the techniques are close and within the critical distance limit, it is impossible to state which techniques would be considered the best in terms of performance. Below, the results are presented from the perspective of each analyzed metric.

Precision (Table C.1): all techniques, when compared, presented a percentage value within the margin of the critical distance limit. The GD_MM (Data Cleaning, Get Dummies, and MinMax Scaler) techniques showed the most significant distance compared to the LE (Data Cleaning and Label Encoder) techniques, with the first greater distance within the limit of the critical distance, with 38% of the distance value criticism. The second largest distance was obtained with the LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques compared to the LE (Data Cleaning and Label Encoder) techniques, with a 34% of critical distance value. In the third largest distance were the GD_SS techniques (Data Cleaning, Get Dummies, and Standard Scaler) compared to the LE techniques, with 27% of the value of the critical distance. When compared, all the other techniques presented percentages ranging from 4% to 23% of the critical distance value.

Critical Distance: 2.17						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
GD_MM	3.17	LE_SS	3.25	0.08	4%	Equal
GD_MM	3.17	GD_SS	3.42	0.25	11%	Equal
GD_MM	3.17	GD	3.50	0.33	15%	Equal
GD_MM	3.17	LE_MM	3.67	0.50	23%	Equal
GD_MM	3.17	LE	4.00	0.83	38%	Equal
LE_SS	3.25	GD_SS	3.42	0.17	8%	Equal
LE_SS	3.25	GD	3.50	0.25	11%	Equal
LE_SS	3.25	LE_MM	3.67	0.42	19%	Equal
LE_SS	3.25	LE	4.00	0.75	34%	Equal
GD_SS	3.42	GD	3.50	0.08	4%	Equal
GD_SS	3.42	LE_MM	3.67	0.25	11%	Equal
GD_SS	3.42	LE	4.00	0.58	27%	Equal
GD	3.50	LE_MM	3.67	0.17	8%	Equal
GD	3.50	LE	4.00	0.50	23%	Equal
LE_MM	3.67	LE	4.00	0.33	15%	Equal

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

Recall (Table C.2): all techniques, when compared, showed a percentage value within the margin of the critical distance limit. The GD_MM (Data Cleaning, Get Dummies, and MinMax Scaler) techniques showed the most significant distance compared to the LE (Data Cleaning and Label Encoder) techniques, with the first greater distance within the critical distance limit, with 59% of the distance value criticism. The second most significant distance was obtained with the GD (Data

Cleaning and Get Dummies) techniques compared to the LE (Data Cleaning and Label Encoder) techniques, with 57% of the critical distance value. In the third largest distance were the GD_SS techniques (Data Cleaning, Get Dummies, and Standard Scaler) compared to the LE techniques, with 51% of the value of the critical distance. When compared, all the other techniques presented percentages ranging from 2% to 40% of the critical distance value.

Critical Distance: 2.17						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
GD_MM	3.04	GD	3.08	0.04	2%	Equal
GD_MM	3.04	GD_SS	3.21	0.17	8%	Equal
GD_MM	3.04	LE_SS	3.46	0.42	19%	Equal
GD_MM	3.04	LE_MM	3.88	0.84	39%	Equal
GD_MM	3.04	LE	4.33	1.29	59%	Equal
GD	3.08	GD_SS	3.21	0.13	6%	Equal
GD	3.08	LE_SS	3.46	0.38	17%	Equal
GD	3.08	LE_MM	3.88	0.80	37%	Equal
GD	3.08	LE	4.33	1.25	57%	Equal
GD_SS	3.21	LE_SS	3.46	0.25	11%	Equal
GD_SS	3.21	LE_MM	3.88	0.67	31%	Equal
GD_SS	3.21	LE	4.33	1.12	51%	Equal
LE_SS	3.46	LE_MM	3.88	0.42	19%	Equal
LE_SS	3.46	LE	4.33	0.87	40%	Equal
LE_MM	3.88	LE	4.33	0.45	21%	Equal

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

F1-Score (Table C.3): all techniques, when compared, presented a percentage value within the margin of the critical distance limit. The LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques showed the most significant distance compared to the LE (Data Cleaning and Label Encoder) techniques, with the first greater distance within the limit of the critical distance, with 38% of the distance value criticism. In the second largest distance, the GD_SS (Data Cleaning, Get Dummies, and Standard Scaler) and GD MM (Data Cleaning, Get Dummies, and MinMax Scaler) techniques were aligned when compared to the LE techniques, both with 27% of the critical distance value. In the third largest distance, the GD (Data Cleaning and Get Dummies) and LE MM (Data Cleaning, Label Encoder, and MinMax Scaler) techniques were aligned when compared to the LE techniques, both with 23% of the critical distance value. When compared, all the other techniques presented percentages ranging from 0% to 15% of the critical distance value.

Critical Distance: 2.17						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
LE_SS	3.17	GD_SS	3.42	0.25	11%	Equal
LE_SS	3.17	GD_MM	3.42	0.25	11%	Equal
LE_SS	3.17	GD	3.50	0.33	15%	Equal
LE_SS	3.17	LE_MM	3.50	0.33	15%	Equal
LE_SS	3.17	LE	4.00	0.83	38%	Equal
GD_SS	3.42	GD_MM	3.42	0.00	0%	Equal
GD_SS	3.42	GD	3.50	0.08	4%	Equal
GD_SS	3.42	LE_MM	3.50	0.08	4%	Equal
GD_SS	3.42	LE	4.00	0.58	27%	Equal
GD_MM	3.42	GD	3.50	0.08	4%	Equal
GD_MM	3.42	LE_MM	3.50	0.08	4%	Equal

GD_MM	3.42	LE	4.00	0.58	27%	Equal
GD	3.50	LE_MM	3.50	0.00	0%	Equal
GD	3.50	LE	4.00	0.50	23%	Equal
LE_MM	3.50	LE	4.00	0.50	23%	Equal

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

FAR (Table C.4): all techniques, when compared, showed a percentage value within the margin of the critical distance limit. The LE_SS (Data Cleaning, Label Encoder, and Standard Scaler) techniques showed the most significant distance compared to the GD (Data Cleaning and Get Dummies) techniques, with the first greater distance within the critical distance limit, with 50% of the distance value criticism. The second largest distance is also from the LE_SS techniques compared to the GD_SS techniques (Data Cleaning, Get Dummies, and Standard Scaler), with 42% of the critical distance value. In the third largest distance were the LE (Data Cleaning and Label Encoder) techniques compared to the GD techniques, with 34% of the value of the critical distance. When compared, all the other techniques presented percentages ranging from 0% to 27% of the critical distance value.

Critical Distance: 2.17						
Group 1	**Ranking**	**Group 2**	**Ranking**	**Distance**	**(%)**	**Hypotheses**
LE_SS	2.92	LE	3.25	0.33	15%	Equal
LE_SS	2.92	LE_MM	3.50	0.58	27%	Equal
LE_SS	2.92	GD_MM	3.50	0.58	27%	Equal
LE_SS	2.92	GD_SS	3.83	0.91	42%	Equal
LE_SS	2.92	GD	4.00	1.08	50%	Equal
LE	3.25	LE_MM	3.50	0.25	11%	Equal
LE	3.25	GD_MM	3.50	0.25	11%	Equal
LE	3.25	GD_SS	3.83	0.58	27%	Equal
LE	3.25	GD	4.00	0.75	34%	Equal
LE_MM	3.50	GD_MM	3.50	0.00	0%	Equal
LE_MM	3.50	GD_SS	3.83	0.33	15%	Equal
LE_MM	3.50	GD	4.00	0.50	23%	Equal
GD_MM	3.50	GD_SS	3.83	0.33	15%	Equal
GD_MM	3.50	GD	4.00	0.50	23%	Equal
GD_SS	3.83	GD	4.00	0.17	8%	Equal

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