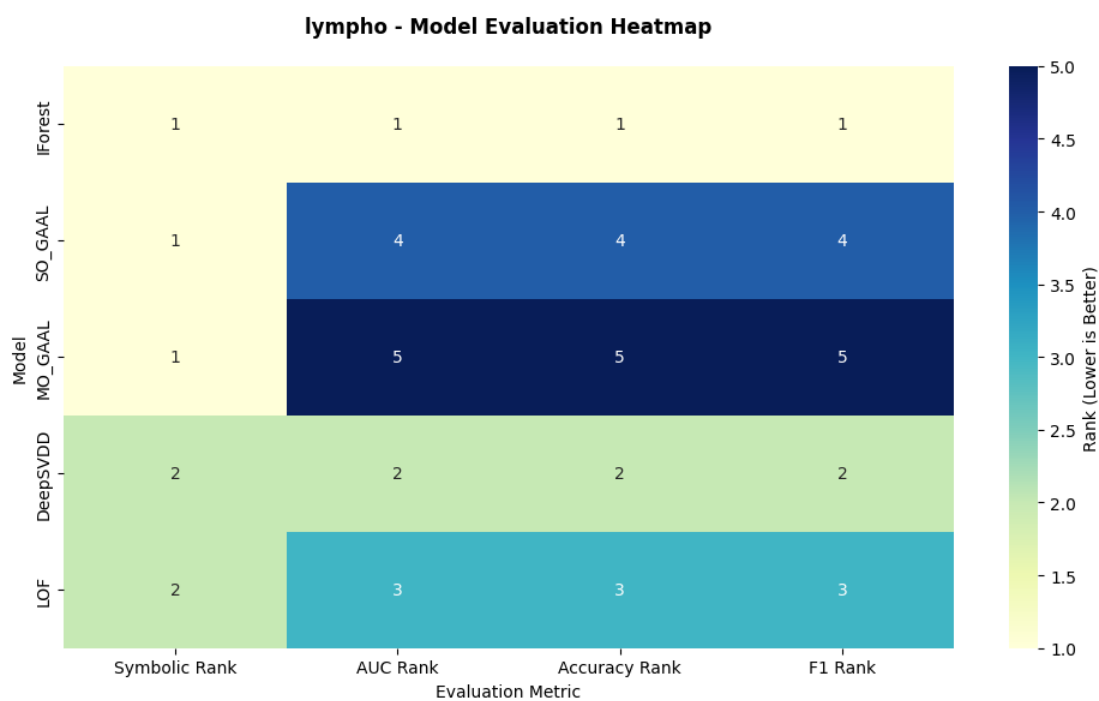


Model Evaluation Summary Table for lympho

Model	Symbolic Rank	Symbolic Score	ROC AUC	Accuracy	Average Precision	F1 (Minority)	Precision (Minority)	Recall (Minority)	AUC Rank	Accuracy Rank	F1 Rank
IForest	1	2.4	1.0	0.9392	1.0	0.5714	0.4	1.0	1	1	1
SO_GAAL	1	2.4	0.5833	0.8851	0.2834	0.1905	0.1333	0.3333	4	4	4
MO_GAAL	1	2.4	0.4354	0.8716	0.0588	0.0952	0.0667	0.1667	5	5	5
LOF	2	2.0	0.9683	0.9122	0.7128	0.381	0.2667	0.6667	3	3	3
DeepSVDD	2	2.0	0.973	0.9257	0.7302	0.4762	0.3333	0.8333	2	2	2

Model Evaluation Heatmap for lympho



LLM Analysis and Recommendations for lympho

LLM Analysis & Model Recommendation for lympho Dataset

Overview Summary for lympho

The lympho dataset consists of 148 samples with 18 features each. The dataset is highly imbalanced, with an anomaly ratio of 0.040541, and contains no missing values. The data is medium-dimensional and relatively clean. The skewness and kurtosis of the dataset are relatively high, with 16.67% of the features having a skewness greater than 2.0 and the same percentage having a kurtosis greater than 7.0.

Comparison of Symbolic vs. Actual Scores

The symbolic scores, which are calculated based on the dataset's characteristics and the strengths and weaknesses of each model, were compared with the actual evaluation scores. The IForest model, which was ranked first symbolically, also performed best in the actual evaluations, with a ROC AUC of 1.0, an average precision of 1.0, and an accuracy of 0.9392.

The SOGAAL and MOGAAL models, despite having the same symbolic rank as IForest, performed significantly worse in the actual evaluations. The DeepSVDD and LOF models, which had a lower symbolic rank, performed better than expected, with DeepSVDD ranking second in the actual evaluations.

Model Suitability Summary Table for lympho

Model	Symbolic Rank	Symbolic Score	ROC AUC	Average Precision	Accuracy	F1 (Minority)	Precision (Minority)	Recall (Minority)	Eval Rank (ROC AUC)	Eval Rank (F1)	Eval Rank (Accuracy)	Eval Rank (AP)
IForest	1	2.4	1.0	1.0	0.9392	0.5714	0.4	1.0	1	1	1	1
SOGAAL	1	2.4	0.5833	0.2834	0.8851	0.1905	0.1333	0.3333	4	4	4	4
MOGAAL	1	2.4	0.4354	0.0588	0.8716	0.0952	0.0667	0.1667	5	5	5	5
LOF	2	2.0	0.9683	0.7128	0.9122	0.381	0.2667	0.6667	3	3	3	3
DeepSVDD	2	2.0	0.973	0.7302	0.9257	0.4762	0.3333	0.8333	2	2	2	2

0.1905 / 0.1333 / 0.3333 / 4 / 4 / 4 / 4 // MOGAAL | 1 | 2.4 | 0.4354 | 0.0588 | 0.8716 | 0.0952 | 0.0667 | 0.1667 | 5 | 5 | 5 | 5 | | LOF | 2 | 2.0 | 0.9683 | 0.7128 | 0.9122 | 0.381 | 0.2667 | 0.6667 | 3 | 3 | 3 | 3 | | DeepSVDD | 2 | 2.0 | 0.973 | 0.7302 | 0.9257 | 0.4762 | 0.3333 | 0.8333 | 2 | 2 | 2 | 2 |

Model Evaluation Heatmap for lympho

The heatmap visualization shows that the IForest model outperforms all other models across all evaluation metrics. The DeepSVDD model also performs well, particularly in terms of ROC AUC, average precision, and accuracy. The SOGAAL and MOGAAL models, despite their high symbolic scores, perform poorly across all metrics.

Analysis of Summary Table and Heatmap

The IForest model is the clear winner in terms of both symbolic and actual evaluation scores. Despite its weakness in handling small sample sizes, it performs exceptionally well on the lympho dataset, which is small and highly imbalanced. This suggests that the model's strengths in handling structured, clean, and highly imbalanced data outweigh its weakness in handling small sample sizes.

The DeepSVDD model also performs well, ranking second in the actual evaluations. Despite its lower symbolic score, it demonstrates strong performance on the lympho dataset, particularly in terms of ROC AUC, average precision, and accuracy.

LLM Recommendations and Justification for Model Selection

Based on the symbolic and actual evaluation scores, the IForest model is recommended for deployment. It performs exceptionally well on the lympho dataset, achieving a ROC AUC, average precision, and accuracy of 1.0, 1.0, and 0.9392, respectively. Its strengths in handling structured, clean, and highly imbalanced data make it well-suited to the lympho dataset.

The DeepSVDD model can also be considered as a second option. Despite its lower symbolic score, it performs well in the actual evaluations and could potentially be improved further with hyperparameter tuning.

Suggested Preprocessing Steps for lympho

Given that the lympho dataset is already clean and structured, no additional preprocessing steps are necessary. However, feature scaling could be applied to normalize the data and potentially improve model performance.

Hyperparameter Tuning Recommendations

For the IForest model, hyperparameters such as the number of estimators and the contamination parameter could be tuned to optimize performance. For the DeepSVDD model, the nu parameter, which controls the trade-off between the proportion of outliers and the size of the decision boundary, could be tuned.

Final Professional Recommendation

In conclusion, the IForest model is recommended for deployment on the lympho dataset. This recommendation is based on the model's exceptional performance in both symbolic and actual evaluations, as well as its suitability for the dataset's characteristics. The DeepSVDD model can also be considered as a second option. Hyperparameter tuning could potentially improve the performance of both models.