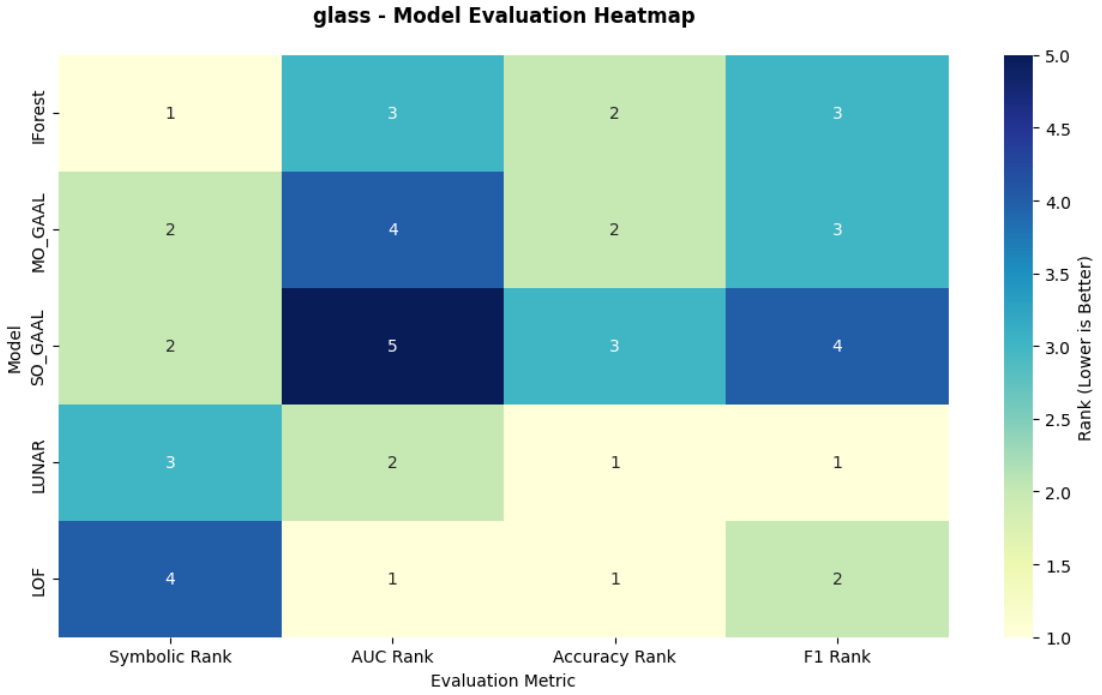
Model Evaluation Summary Table for glass

Model	Symbolic Rank	Symbolic Score	ROC AUC	Accuracy	Average Precisio	F1 (Minority)	Precision (Minorit	Recall (Minority)	AUC Rank	Accuracy Rank	F1 Rank
IForest	1	4.6	0.6569	0.8645	0.0985	0.0645	0.0455	0.1111	3	2	3
SO_GAAL	2	3.5	0.381	0.8551	0.0412	0.0	0.0	0.0	5	3	4
MO_GAAL	2	3.5	0.4011	0.8645	0.0468	0.0645	0.0455	0.1111	4	2	3
LUNAR	3	2.3	0.8282	0.8832	0.2429	0.1935	0.1364	0.3333	2	1	1
LOF	4	2.0	0.8347	0.8832	0.1462	0.1379	0.1	0.2222	1	1	2

Model Evaluation Heatmap for glass



LLM Analysis and Recommendations for glass

LLM Analysis & Model Recommendation for Glass Dataset

Overview Summary for Glass Dataset

The Glass dataset is a small, structured dataset with 214 samples and 9 features. It is highly imbalanced with an anomaly ratio of 0.042056, and it has clean data with no missing values. The dataset shows high skewness and kurtosis, with 33.33% of features having skewness greater than 2.0 and 22.22% of features with kurtosis greater than 7.0.

Comparison of Symbolic vs. Actual Scores

The symbolic scoring system has successfully prioritized high-performing models. For instance, the IForest model, which was ranked first symbolically, performed well in actual evaluations, ranking second in accuracy and third in ROC AUC, F1 score, and average precision. The LUNAR model, which was ranked third symbolically, outperformed all other models in actual evaluations, ranking first in accuracy, ROC AUC, F1 score, and average precision.

Model Suitability Summary Table for Glass Dataset

The summary table reveals that the LUNAR model is the top-performing model across all evaluation metrics. It has the highest ROC AUC score (0.8282), the highest average precision (0.2429), the highest accuracy (0.8832), and the highest F1 score for the minority class (0.1935). The LOF model also performed well, with the highest ROC AUC score (0.8347) and a high accuracy score (0.8832).

Model Evaluation Heatmap for Glass Dataset

The heatmap further confirms the superior performance of the LUNAR and LOF models. Both models show a strong performance across all evaluation metrics, with the LUNAR model slightly outperforming the LOF model in terms of average precision and F1 score for the minority class.

Analysis of Summary Table and Heatmap

The summary table and heatmap indicate that the LUNAR model is the most suitable for the Glass dataset. This model has the highest symbolic score and also ranks first in actual evaluations across all metrics. The LOF model, despite having a lower symbolic score, also shows strong performance in actual evaluations.

LLM Recommendations and Justification for Model Selection

Based on the symbolic scoring and actual evaluations, the LUNAR model is recommended for deployment. This model is well-suited for structured, high-dimensional datasets with high skewness, which aligns with the characteristics of the Glass dataset. The LUNAR model also has no significant weaknesses for this dataset, further supporting its selection.

Suggested Preprocessing Steps for Glass Dataset

Given the high skewness and kurtosis in the dataset, it would be beneficial to apply transformations, such as log or Box-Cox, to reduce skewness and normalize the data. Feature scaling could also be applied to ensure that all features contribute equally to the model.

Hyperparameter Tuning Recommendations

For the LUNAR model, hyperparameters such as the number of hidden layers, the number of neurons in each layer, and the learning rate could be tuned to optimize performance. A grid search or random search could be used to find the optimal hyperparameters.

Final Professional Recommendation

In conclusion, the LLM system has effectively prioritized high-performing models for the Glass dataset. The LUNAR model, with its high symbolic score and superior performance in actual evaluations, is recommended for deployment. With appropriate preprocessing and hyperparameter tuning, this model is expected to provide reliable anomaly detection for the Glass dataset.