IRBAI AUDIT - DATA (AEM-T1)

The International Regulatory Body for AI (IRBAI) has developed a framework to ensure transparency, accountability, and the responsible use of data in AI systems.

Performing the Audit:

This document outlines the specifications for conducting an audit on evaluation metrics of the models and the performance of the models implemented. The purpose of this audit is to evaluate the accuracy of the output data.

Questionnaire

The person responsible for the training models must complete a questionnaire to provide essential information for the audit. The questionnaire should include the following indicatives and corresponding answer options:

	Indicatives	Answer Options
Accuracy	How accurately does the model classify instances?	[percentage]
Precision	How many of the instances predicted as positive are actually positive?	[Number]

Recall (Sensitivity)	How many of the positive instances are correctly identified by the model?	[Number]
Specificity	How many of the negative instances are correctly identified by the model?	[Number]
F1 Score	The harmonic mean of precision and recall.	[mean]
Area Under ROC Curve	How well does the model distinguish between positive and negative instances across different probability thresholds?	[delta p]
Confusion Matrix	A matrix showing the counts of true positives, true negatives, false positives, and false negatives.	[mtx v]
Mean Absolute Error	The average absolute difference between the predicted and actual values.	[delta]
Mean Squared Error	The average squared difference between the predicted and actual values.	[sqd]
Root Mean Squared Error	The square root of the average squared difference between the predicted and actual values.	[sq]

R-squared	The proportion of the variance in the target variable explained by the model.	[Proportion]
Mean Average Precision	The average precision calculated at different recall levels.	[Number]
Normalized Entropy	The measure of information content in the predicted class probabilities.	[Probabilistic values]
Diversity	How diverse are the predictions made by the model?	[Diversity level]
Robustness	How well does the model perform under varying conditions or perturbations?	[Risk Level]
Scalability	How well does the model handle larger datasets or higher dimensional feature spaces?	[Level]
Training Time	How much time does it take to train the model?	[Time]
Inference Time	How much time does it take to make predictions for new instances?	[Time]

Computational Resources	What computational resources are required to train or use the model?	[List]
Documentation	Is the model's documentation comprehensive and well-explained?	[Well-documented, Partially documented, Not documented]
Reproducibility	Can the model and its results be reproduced?	[Reproducible, Partially reproducible, Not reproducible]
Bias and Fairness	Are there any biases or fairness concerns in the models predictions?	[Evaluated, Not evaluated]
Interpretability	How interpretable are the models predictions?	[Level]
Privacy and Security	Does the model handle data privacy and security concerns appropriately?	[Addressed, Not addressed]
Model Performance in Real-world Scenarios	How well does the model perform when applied to real-world scenarios or unseen data?	[High performance, Moderate performance, Low performance]

Evaluation

Once the data, training models and optimization models are accessible by the IRBAI Audit system, they will undergo an independent evaluation. The evaluation process will be conducted in an objective and impartial manner.

Audit Report and Publication

After completing the audit, a comprehensive report will be generated, detailing the findings, recommendations, and any identified issues related to the AI system. The report will be published to the IRBAI platform and will present the transparency and if needed advise on improvements in optimization models and AI system performance. The report will include anonymized examples and statistics to support the conclusions drawn during the audit.

Audit Duration

The length of the audit will vary depending on the size and complexity of the data and models being audited. A timeline for the audit process will be established based on the specific requirements of each audit, ensuring sufficient time for thorough analysis and evaluation.