

FAIRLY

Controls Library

Name	Description
ACCURACY_SCORE	The ratio of correct predictions to total predictions
ADJUSTED_MUTUAL_INFO_SCORE	MI score adjusted for chance. Returns 1 when the two partitions are identical.
ADJUSTED_RAND_SCORE	Rand score adjusted for chance. A symmetric measure, defined by the ratio of the number of agreements and all pairs of samples, used to measure similarity between two data clustering.
ADJUSTED_R_SQUARED	Adjusted R-squared is R-squared with a penalty for the number of variables
AGGREGATE_RISK_RATING	The result of the fuzzy logic risk measurement system.
AIC_SCORE	Method of scoring and selecting a model. Models with lower AIC scores are preferred. Penalizes more complex models less.
AUC_PR_SCORE	A measure of the area under a precision-recall curve.
AUC_SCORE	A measure of the area under an roc curve
AVERAGE_PRECISION	The average value of precision as a function of recall
BIAS_INSPECTOR_DATASET_AOD_CHART	The output dataset results aod chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_DATASET_CORRELATION_METRICS_CHART	The output dataset results correlation chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_DATASET_DISTRIBUTION_CHART	The output dataset results distribution chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_DATASET_DI_CHART	The output dataset results di chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_DATASET_PERFORMANCE_TABLE	The output dataset performance metrics table from the AI based Bias Inspection feature.
BIAS_INSPECTOR_DATASET_PRECISION_CHART	The output dataset results precision chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_DATASET_RECALL_CHART	The output dataset results recall chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_DATASET_ROCAUC_CHART	The output dataset results rocauc chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_MODEL_AOD_CHART	The output model results aod chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_MODEL_DI_CHART	The output model results di chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_MODEL_PERFORMANCE_TABLE	The output model results performance metrics table from the AI based Bias Inspection feature.
BIAS_INSPECTOR_MODEL_PRECISION_CHART	The output model results precision chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_MODEL_RECALL_CHART	The output model results recall chart from the AI based Bias Inspection feature.
BIAS_INSPECTOR_MODEL_ROCAUC_CHART	The output model results rocauc graph from the AI based Bias Inspection feature.
BIN_WIDTH	The partitions of a continuous variable which convert it into a categorical variable
CLASSIFICATION_ACCURACY	Ratio of numbers of correct prediction to total numbers of prediction
COMPLETENESS_SCORE	Used to check if all members of a given class are assigned to the same cluster. Returns 1 when it is perfectly complete labeling.
CONFUSION_MATRIX	A table which displays TP, TN, FP and FN in a 2x2 format
CORRELATION_MATRIX	A table of the correlations and autocorrelations between the features
D2_TWEEDIE_SCORE	A generalization of the R ² score where the squared error is replaced by the Tweedie deviance. The best score is 1.0 and it can be negative.
DATADRIFT_CATEGORICAL_CHART	A chart which shows the difference in average of categorical columns in the dataset.
DATADRIFT_CONTINUOUS_CHART	A chart which shows the difference in average of continuous columns in the dataset.
DISPARATE_IMPACT	The difference in receiving certain performance score by protected groups.
EQUALIZED_ODDS	A classifier satisfies this definition if the subjects in the protected and unprotected groups have equal true positive rate and equal false positive rate
EXPLAINED_VARIATION_SCORE	The proportion to which a mathematical model accounts for the variation (dispersion) of a given data set.
F1_SCORE	F1 Score is the weighted average of Precision and Recall.
FALSE_NEGATIVES	The amount of false negatives.

FALSE_POSITIVES	The amount of false positives.
FALSE_POSITIVE_RATE	A false positive ratio (also known as fall-out or false alarm ratio) is the probability of falsely rejecting the null hypothesis for a particular test. The false positive rate is calculated as the ratio between the number of negative events wrongly categorized as positive (false positives) and the total number of actual negative events (regardless of classification).
FBETA_SCORE	Beta is used to weigh the balance between precision and recall when evaluating using this method
FUZZY_LOGIC_CHART	The output chart from the fuzzy logic controller.
GAIN_LIFT_CHART	The accumulative gain chart (lift chart) checks the effectiveness of a classification model
GINI	the Gini index (or coefficient) is a synthetic indicator that captures the level of inequality for a given variable and population. It varies between 0 (perfect equality) and 1 (extreme inequality). Between 0 and 1, the higher the Gini index, the greater the inequality.
GINI_COEFF	Ratio between the the area of the ROC curve and the diagonal line and the area of the above triangle. Calculated using $2 * AUC - 1$
HAMMING_LOSS	The fraction of incorrectly predicted labels
HINGE_LOSS	The hinge loss is used for "maximum-margin" classification, most notably for support vector machines.
HOMOGENEITY_SCORE	Used to check if each cluster contains only members of a single class. Returns 1 when it is perfectly homogeneous labeling.
HYPERPARAMETER_LIST	A list of model hyperparameters.
JACCARD_INDEX	Ratio of overlap section to union section of two sets. The higher the percentage, the more similar the two sets/populations.
K-FOLD_CROSS_VALIDATION	Divide the population into k samples, validate on one sample while train models on other k-1 samples to improve overfitting.
KULLBACK_LEIBLER_CONVERGENCE	A measure of how one probability distribution is different from a second, reference probability distribution.
LIME	LIME: Local Interpretable Model-agnostic Explanations is a technique that approximates any black box machine learning model with a local, interpretable model to explain each individual prediction. the lime analysis bar chart which shows the feature sensitivity for a particular observation. A large positive or negative value for a particular feature indicates that it was the dominant contributor to the outcome. If a protected category attribute is the dominant feature, then there is potential for bias.
LOG_LOSS	The log_loss metric represents the penalty for falsely classifying datapoints. The exact penalty is equal to the sum of the logs of the probability.
MALLOWS_CP	Used to asses the fit of a regression model that has been estimated using ordinary least squares
MARKET_VOLATILITY	The square of the standard deviation of market returns.
MATTHEWS_CORRELATION_COEFFICIENT	Correlation coefficient between observed and predicted primary classifications.
MAX_ERROR	Absolute value of the most significant difference between a predicted value and its real value.
MEAN_ABSOLUTE_ERROR	The absolute difference between the estimated values and the actual value.
MEAN_AVERAGE_PRECISION	Mean of the average precision scores for each query
MEAN_SQUARED_ERROR	The average squared difference between the estimated values and the actual value.
MEDIAN_ABSOLUTE_ERROR	Median absolute error regression loss. Best value is 0.
MLFLOW_ARTIFACT	A file generated from the ML Flow process.
MLFLOW_MODEL	The model data used for ML Flow.
MLFLOW_TAG	Tags generated from the ML Flow process.
MUTUAL_INFO_SCORE	The measure of the similarity between two labels of the same data.
NEG_MEAN_GAMMA_DEVIANCE	Mean Gamma deviance regression loss.
NEG_MEAN_POISSON_DEVIANCE	Mean Poisson deviance regression loss.
NORMALIZED_MUTUAL_INFO_SCORE	Normalization of the MI score. Returns 1 when there is perfect correlation.

PINBALL_LOSS	Used to evaluate the predictive performance of quantile regression models.
PRECISION_SCORE	Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.
PRE_MODEL_ANALYSIS	Pre-model analysis has been renamed as Data Validation .
RAG_STATUS	Red = High risk, Amber = Medium risk, Green = Low risk
RAND_SCORE	The ratio of the number of agreements and all pairs of samples, used to measure similarity between two data clustering.
RECALL_SCORE	Recall is the ratio of correctly predicted positive observations to the all observations in actual class.
ROC_CURVE	A chart which displays the accuracy of the model under different contingencies
ROOT_MEAN_SQUARE_ERROR	Assuming errors are unbiased and follows normal distribution, it can show the difference from predicted values and actual values.
ROOT_MEAN_SQUARE_LOG_ERROR	The relative error between and the predicted and the actual value
R_SQUARED	R-squared (R2) is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model.
SENSITIVITY_SCORE	Coverage of actual positive sample.
SHAP_BAR	The SHAP bar chart shows mean SHAP value for each feature
SHAP_BEESWARM	The beeswarm chart shows the global feature sensitivity by showing the distribution of the SHAP values. Any feature which presents a skewed distribution shows strong negative or positive feature sensitivity.
SHARPE_RATIO	A skewed distribution shows strong negative or positive feature sensitivity.
SPECIFICITY_SCORE	Coverage of actual negative sample. Calculated using $TN/(TN + FP)$
SUPPORT	Number of samples of the true response that lie in each class of target values
TRUE_POSITIVE_RATE	The ratio of correctly predicted positive observations to the total predicted positive observations.
V_MEASURE_SCORE	The harmonic mean between homogeneity and completeness and it is symmetric.
ZERO_ONE_LOSS	The indicator function that returns the fraction or number of misclassifications. Lower scores are preferred (0 is the best performance).