**CONFUSION MATRIX:**

**WHAT IS CONFUSION MATRIX?**

* It is a table that is used in classification problems to assess where errors in the model were made.
* The rows represent the actual classes the outcomes should have been. While the columns represent the predictions we have made. Using this table it is easy to see which predictions are wrong.

A confusion matrix is a performance evaluation tool used in machine learning and statistical analysis. It is a table that summarizes the performance of a classification algorithm by comparing predicted and actual classes across different levels of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

In a binary classification problem, the confusion matrix has two classes, positive and negative. A positive outcome is the outcome that the model is trying to predict, and a negative outcome is the opposite of the positive outcome.

**The confusion matrix is organized into four cells, as follows:**

True Positive (TP): The model correctly predicted a positive outcome.

False Positive (FP): The model incorrectly predicted a positive outcome when the actual outcome was negative.

True Negative (TN): The model correctly predicted a negative outcome.

False Negative (FN): The model incorrectly predicted a negative outcome when the actual outcome was positive.

The following is an example of a confusion matrix:

|  |  |  |
| --- | --- | --- |
|  | **ACTUAL POSITIVE** | **ACTUAL NEGATIVE** |
| **PREDICTED POSITIVE** | TP | FP |
| **PREDICTED NEGATIVE** | FN | TN |

The values in the confusion matrix can be used to calculate various performance metrics such as accuracy, precision, recall, F1-score, and others, which can help assess the quality of the classification model.

**1. TRUE POSITIVE(TP):**

In a confusion matrix, a true positive (TP) is a term used to describe the number of cases where a positive outcome was correctly predicted by a machine learning model.

More specifically, a true positive is the number of cases where the model predicted a positive outcome, and the actual outcome was also positive. If a model was trained to detect whether an hard disk is failed or not, and it predicted that an hard disk is failed, and the actual outcome was indeed failed, then this would be considered a true positive.

In a confusion matrix, the true positive value is located in the upper-left cell, which indicates the number of true positives.

**2. FALSE POSITIVE(FP) :**

In a confusion matrix, a false positive (FP) is a term used to describe the number of cases where a negative outcome was incorrectly predicted as positive by a machine learning model.

More specifically, a false positive is the number of cases where the model predicted a positive outcome, but the actual outcome was negative. If a model was trained to detect whether hard disk is failed or not, and it predicted that an hard disk is failed, but the actual outcome was not failed, then this would be considered a false positive.

In a confusion matrix, the false positive value is located in the upper-right cell, which indicates the number of false positives.

**3.TRUE NEGATIVE(TN) :**

In a confusion matrix, a true negative (TN) is a term used to describe the number of caseswhere a negative outcome was correctly predicted by a machine learning model.

More specifically, a true negative is the number of cases where the model predicted a negative outcome, and the actual outcome was also negative. If a model was trained to detect whether hard disk is failed or not , and it predicted that an hard disk is not failed, and the actual outcome was indeed not failed, then this would be considered a true negative.

In a confusion matrix, the true negative value is located in the lower-left cell, which indicates the number of true negatives.

**4. FALSE NEGATIVE(FN) :**

In a confusion matrix, a false negative (FN) is a term used to describe the number of cases where a positive outcome was incorrectly predicted as negative by a machine learning model.

More specifically, a false negative is the number of cases where the model predicted a negative outcome, but the actual outcome was positive. If a model was trained to detect whether hard disk is failed or not, and it predicted that hard disk is not failed, but the actual outcome was indeed failed, then this would be considered a false negative.

In a confusion matrix, the false negative value is located in the lower-right cell, which indicates the number of false negatives.

**CREATED METRICS:**

The matrix provides us with many useful metrics that help us to evaluate out classification model.

The different measures include: Accuracy, Precision, Sensitivity (Recall), , and the F-score, explained below.

**1.ACCURACY:**

Accuracy measures how often the model is correct.

Accuracy = (True Positive + True Negative) / Total Predictions

**2. SENSITIVITY(RECALL):**

Of all the positive cases, what percentage are predicted positive?

Sensitivity (sometimes called Recall) measures how good the model is at predicting positives.

This means it looks at true positives and false negatives (which are positives that have been incorrectly predicted as negative).

Recall = True Positive / (True Positive + False Negative)

**3. F-SCORE:**

F-score is the "harmonic mean" of precision and sensitivity.

It considers both false positive and false negative cases and is good for imbalanced datasets.

F-score = 2 \* ((Precision \* Sensitivity) / (Precision + Sensitivity))

**4. ROC SCORE:**

The ROC score, also known as the AUC (Area Under the Curve) score, is calculated by finding the area under the ROC curve. The ROC curve is a graph that plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

The formula for calculating the ROC score is as follows:

**ROC score = AUC = ∫ TPR(FPR) dFPR**

where TPR is the true positive rate and FPR is the false positive rate. The integral represents the area under the ROC curve, which ranges from 0 to 1.

In practice, the ROC curve is usually plotted using a set of predicted probabilities for the positive class. The probabilities are thresholded at various levels to generate a series of TPR and FPR values, which are then used to plot the ROC curve. The AUC is then calculated numerically using numerical integration methods or the trapezoidal rule.