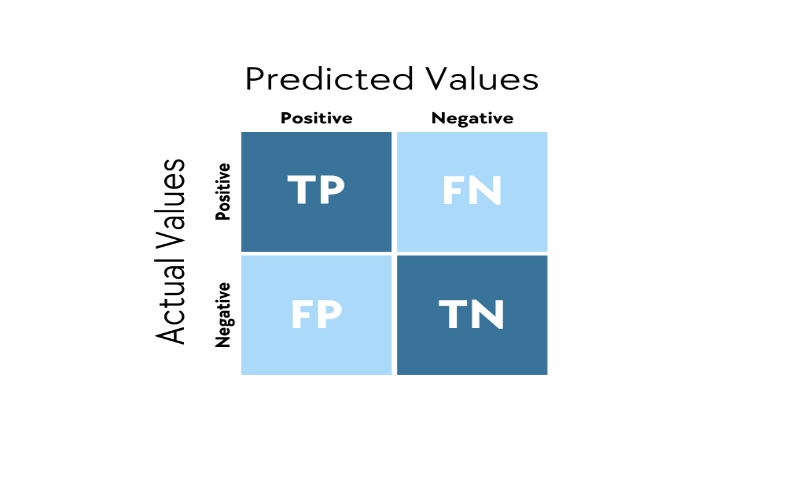
**AI/ML Intern – Bhavya**

**Day 5 Task 4**

**Learning Topics:**

* Confusion Matrix.
* Precision, Recall, F1- Score.
* When Accuracy is Misleading.

**1. Confusion Matrix:**A Confusion Matrix is a simple table in machine learning to evaluate the performance of the classification model. It shows the actual versus predicted outcomes, thereby giving you clues about those cases where the model is right or wrong.



Imagine a model predicting emails as spam or not spam. The confusion matrix would compare the actual labels of true spam or not spam versus what the model predicted. It contains four parts:

* True Positive (TP): Model correctly predicts spam (e.g., 50 spam emails correctly identified).
* True Negative (TN): Model correctly predicts not spam (e.g., 40 non-spam emails correctly identified).
* False Positive (FP): Model incorrectly predicts spam when it was not (e.g., 10 non-spam emails marked as spam).
* False Negative (FN): Model incorrectly predicts not spam when it is actually spam (e.g., 5 spam emails missed).

Predicted Spam Predicted Not Spam

Actual Spam TP (50) FN (5)

Actual Not Spam FP (10) TN (40)

**2. Precision, Recall, and F1-Score** are metrics used in machine learning to evaluate a classification model’s performance, especially when classes are imbalanced. Let’s use the spam email example from your confusion matrix.

1. **Precision**: Measures how many predicted positives are actually correct.

**Formula**:



1. **Recall:** Measures how many actual positives the model catches.

**Formula:**



1. **F1-Score:** Combines precision and recall into a single score, balancing both.

**Formula:**

.

**3**. **When Accuracy is Misleading?:**

When working with imbalanced datasets, accuracy can be misleading in assessing the performance of machine-learning models. Accuracy is essentially the ratio between the number of correct predictions (True Positives + True Negatives) to the number of total predictions. But it doesn't provide us with the whole story when one class is evident over the other.

Suppose you’re trying to build a model for extremely rare disease detection, where out of 1000 patients only 1% (10 sick versus 990 healthy) has the disease. If your model just assumes everybody is healthy, it will get 990 scenarios right (True Negatives) and 10 wrong (False Negatives). In such a case, accuracy will be calculated as (990 + 0) / 1000 = 99%: an appearance of a good model outright, yet it fails to pinpoint sick ones: rather an essential thing!

In this case, accuracy is misleading due to the fact that the dataset is imbalanced (far more healthy individuals than sick ones). The model succeeds by ignoring the rare class; however, from the standpoint of the task, this is completely useless. Precision (how many are actually predicted positives), recall (catching actual positives), and F1 score (balancing precision and recall) would be more useful metrics to use when looking at performance for the minority class. Recall for this disease case will reveal a 0% score (no sick patients detected), clearly highlighting the failure of the model.

Check for class balance and use precision, recall, or F1 score with accuracy to have an overview of the model’s performance.

**Impact on Model:**

Model evaluation using accuracy alone may lead to misleading inferences. The metric can be fooling practitioners as clazzes with higher priority-performing occur in poor manner, sort of inefficient in real-life implementations.

Impact on Model Development: Based on pure accuracy, some training adverse effects may occur:

* Overconfidence: Developers could have deployed bad models assuming that they were useful.
* Neglected classes: Minority classes such as fraud or disease are left aside, thus hurting the model.
* Bad decision-making: Decisions at work or medical level based on such models are doomed to fail and bring harm or even economic losses.

The complementary use of precision, recall, and F1-score with accuracy can produce a more balanced evaluation that will emphasize deficiencies in minority classes' predictions so that steps can be taken to strengthen these aspects in the model, therefore obtaining more reliable output.