**CONFUSION MATRIX – EVALUATION METRICS**

**Validating Parameters for Supervised Learning:**

Consider **POSITIVE** as **APPLE** and **NEGATIVE** as **ORANGE** for better understanding.

|  |  |  |  |
| --- | --- | --- | --- |
| **N=Test**  **Dataset**  **Count** | | **ACTUAL CLASS** | |
| **Positive**  **(Apple)** | **Negative**  **(Orange)** |
| **PREDICTED CLASS** | **Positive**  **(Apple)** | True Positive - **TP**  (True Apple - **TA**)  Correctly Classified as **Apple** (Positive) | False Positive - **FP**  (False Apple - **FA**)  Should Classified as Apple (Positive), But wrongly classified as **Orange** (Negative) |
| **Negative**  **(Orange)** | False Negative - **FN**  (False Orange - **FO**)  Should Classified as Orange (Negative), But wrongly classified as **Apple** (Positive) | True Negative - **TN**  (True Orange - **TO**)  Correctly Classified as **Orange** (Negative) |

**PARAMETERS:**

* Accuracy
* Recall
* Precision
* F1 Score
* Macro Average
* Weighted Average

1. **ACCURACY or OVERALL PERFORMANCE**

**Question**:

**What is the Percentage of Correct Classification of Both POSITIVE and NEGATIVE (Apple & Orange) to the total input of the test set?**

|  |  |
| --- | --- |
| **Accuracy =** | **TP + TN** |
|  |
| **TP + TN + FP + FN** |

Assumption:

|  |  |
| --- | --- |
| Accuracy = | T(Apple) + T(Orange) |
|  |
| T(Apple) + T(Orange)+ F(Apple) + F(Orange) |

1. **RECALL**

Recall talks about only the **correctly classified class**.

**I. Recall - Positive**

**Question**:

**What is the Percentage of Correct Classification of POSITIVE (Apple) to the total input of POSITIVE in the test set?**

|  |  |
| --- | --- |
| **Recall =** | **TP** |
|  |
| **TP + FP** |

Assumption:

|  |  |
| --- | --- |
| Recall = | T(Apple) |
|  |
| T(Apple) + F(Apple) |

**II. Recall - Negative**

**Question**:

**What is the Percentage of Correct Classification of NEGATIVE (Orange) to the total input of NEGATIVE in the test set?**

|  |  |
| --- | --- |
| **Recall =** | **TN** |
|  |
| **TN + FN** |

Assumption:

|  |  |
| --- | --- |
| Recall = | T(Orange) |
|  |
| T(Orange) + F(Orange) |

1. **PRECISION**

Precision talks about **correctly and wrongly classification of the class**.

**I. Precision - Positive**

**Question**:

**What is the Percentage of Correct Classification of POSITIVE (Apple) to the sum of Correctly Classified as POSITIVE and wrongly classified as POSITIVE in the test set?**

|  |  |
| --- | --- |
| **Precision =** | **TP** |
|  |
| **TP + FN** |

Assumption:

|  |  |
| --- | --- |
| Precision = | T(Apple) |
|  |
| T(Apple) + F(Orange) |

**II. Precision - Negative**

**Question**:

**What is the Percentage of Correct Classification of NEGATIVE (Orange) to the sum of Correctly Classified as NEGATIVE and wrongly classified as NEGATIVE in the test set?**

|  |  |
| --- | --- |
| **Precision =** | **TN** |
|  |
| **TN + FP** |

Assumption:

|  |  |
| --- | --- |
| Precision = | T(Orange) |
|  |
| T(Orange) + F(Apple) |

1. **F1 MEASURE**

If the **Recall** value is **high** and P**recision** is **low**. How will you validate your Model Performance?

**I. F1 Measure - Positive**

**Question**:

**What is the Overall Performance of POSITIVE (Apple)?**

|  |  |
| --- | --- |
| **F1 Measure =** | **(Recall X Precision) of Positive** |
| **2X** |
| **(Recall + Precision) of Positive** |

Assumption:

|  |  |
| --- | --- |
| F1 Measure = | (Recall X Precision) of Apple |
| 2X |
| (Recall + Precision) of Apple |

**II. F1 Measure - Negative**

**Question**:

**What is the Overall Performance of NEGATIVE (Orange)?**

|  |  |
| --- | --- |
| **F1 Measure =** | **(Recall X Precision) of Negative** |
| **2X** |
| **(Recall + Precision) of Negative** |

Assumption:

|  |  |
| --- | --- |
| F1 Measure = | (Recall X Precision) of Orange |
| 2X |
| (Recall + Precision) of Orange |

**5. MACRO AVERAGE**

**I. Macro Average - Precision**

**Question**:

**What is the Average Performance of Precision - Correctly and wrongly classified (Apple & Orange)?**

|  |  |
| --- | --- |
| **Macro Average (Precision) =** | **Precision of Positive +**  **Precision of Negative** |
|  |
| **2** |

Assumption:

|  |  |
| --- | --- |
| Macro Average  (Precision) = | Precision (Apple) +  Precision (Orange) |
|  |
| 2 |

**II. Macro Average - Recall**

**Question**:

**What is the Average Performance of Recall - Correctly classified (Apple & Orange)?**

|  |  |
| --- | --- |
| **Macro Average**  **(Recall) =** | **Recall of Positive +**  **Recall of Negative** |
|  |
| **2** |

Assumption:

|  |  |
| --- | --- |
| Macro Average (Recall) = | **Recall** (Apple) +  **Recall** (Orange) |
|  |
| 2 |

**III. Macro Average – F1 Measure**

**Question**:

**What is the Average Performance of F1 Measure (Apple & Orange)?**

|  |  |
| --- | --- |
| **Macro Average**  **(F1 Measure) =** | **F1 Measure of Positive +**  **F1 Measure of Negative** |
|  |
| **2** |

Assumption:

|  |  |
| --- | --- |
| Macro Average (F1 Measure) = | **F1** (Apple) + **F2** (Orange) |
|  |
| 2 |

**6. WEIGHTED AVERAGE**

**I. Weighted Average - Precision**

**Question**:

**What is the Sum of Product of proportion rate (weight) of Precision of each class POSITIVE and NEGATIVE (Apple & Orange)?**

|  |  |
| --- | --- |
| **Weighted Average (Precision) =** | **(Precision of Positive ) X (Total Sum of Positive Samples / Total Number of Samples in the Test Set)** |
| **+** |
| **(Precision of Negative ) X (Total Sum of Negative Samples / Total Number of Samples in the Test Set)** |

Assumption:

|  |  |
| --- | --- |
| Weighted Average (Precision) = | Precision (Apple) X (Total Sum of Apple Samples / Total Number of Samples in the Test Set) |
| + |
| Precision (Orange) X (Total Sum of Orange Samples / Total Number of Samples in the Test Set) |

**II. Weighted Average - Recall**

**Question**:

**What is the Sum of Product of proportion rate (weight) of Recall of each class POSITIVE and NEGATIVE (Apple & Orange)?**

|  |  |
| --- | --- |
| **Weighted Average (Recall) =** | **(Recall of Positive ) X (Total Sum of Positive Samples / Total Number of Samples in the Test Set)** |
| **+** |
| **(Recall of Negative ) X (Total Sum of Negative Samples / Total Number of Samples in the Test Set)** |

Assumption:

|  |  |
| --- | --- |
| Weighted Average (Recall) = | Recall (Apple) X (Total Sum of Apple Samples / Total Number of Samples in the Test Set) |
| + |
| Recall (Orange) X (Total Sum of Orange Samples / Total Number of Samples in the Test Set) |

**III. Weighted Average – F1 Measure**

**Question**:

**What is the Sum of Product of proportion rate (weight) of each F1 Measure of class POSITIVE and NEGATIVE (Apple & Orange)?**

|  |  |
| --- | --- |
| **Weighted Average**  **(F1 Measure) =** | **(F1 Measure of Positive ) X (Total Sum of Positive Samples / Total Number of Samples in the Test Set)** |
| **+** |
| **(F1 Measure of Negative ) X (Total Sum of Negative Samples / Total Number of Samples in the Test Set)** |

Assumption:

|  |  |
| --- | --- |
| Weighted Average  (F1 Measure) = | F1 Measure (Apple) X (Total Sum of Apple Samples / Total Number of Samples in the Test Set) |
| + |
| F1 Measure (Orange) X (Total Sum of Orange Samples / Total Number of Samples in the Test Set) |

**Confusion Matrix for Multiple Classes (Sample)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Apple** | **Orange** | **Mango** | **Banana** |
| **Apple** | A1 | A2 | A3 | A4 |
| **Orange** | O1 | O2 | O3 | O4 |
| **Mango** | M1 | M2 | M3 | M4 |
| **Banana** | B1 | B2 | B3 | B4 |