**Documentation of Classification**

|  |  |  |
| --- | --- | --- |
| **Recall** | Recall Value should always be High | Recall talks about Only Correctly Classified Class **Hint to Remember the Formula :** Row wise Data would come in Denominator |
| **Precision** | Precision Value should always be High | Precision Talks about the Correctly and Wrongly Classification of the Class **Hint to Remember the Formula :** Column wise Data would come in Denominator |

**Classification - Decision Tree**

|  |  |  |  |
| --- | --- | --- | --- |
| **Evaluation Metrics** | **Purpose / Explanation** | **Formula** | **Calculated Value Based on Formula** |
| **Accuracy** | What is the Percentage of Correct Classification of Both Not Purchased and Purchased to the Total Input of the Test Set ? | Accuracy = T(Not Purchased) + T(Purchased) / T(Not Purchased) + T(Purchased) + F(Not Purchased) + F(Purchased) | 0.87 |
| **Recall** | What is the Percentage of Correct Classification of (Not Purchased) to the Total Input of (Not Purchased) in the Test Set ? | Recall = T(Not Purchased) / T(Not Purchased) + F(Not Purchased) | 0.89 |
| **Recall** | What is the Percentage of Correct Classification of (Purchased) to the Total Input of (Purchased) in the Test Set ? | Recall = T(Purchased) / T(Purchased) + F(Purchased) | 0.84 |
| **Precision** | What is the Percentage of Correctly Classified as (Not Purchased) and Wrongly Classified as (Not Purchased) in the Test Set ? | Precision = T(Not Purchased) / T(Not Purchased) + F(Purchased) | 0.90 |
| **Precision** | What is the Percentage of Correctly Classified as (Not Purchased) and Wrongly Classified as (Purchased) in the Test Set ? | Precision = T(Purchased) / T(Purchased) + F(Not Purchased) | 0.82 |
| **F1 Measure** | What is the Overall Performance of Not Purchased ? | 2 x (Recall x Precision)  -------------------------- (Recall + Precision) | 0.90 |
| **F1 Measure** | What is the Overall Performance of Purchased ? | 2 x (Recall x Precision)  -------------------------- (Recall + Precision) | 0.83 |
| **Macro Average** | **What is the sum of Average Performance of each class Through Precision/Recall/F1 Measure ?** | **Macro Average refers to Average Performance of Individuals** |  |
| **Macro Average** | What is the Average Performance of Precision? (Correctly Classified as (Not Purchased) and Wrongly Classified as (Purchased) in the Test Set) | (Precision (Not Purchased) + Precision (Purchased)) / 2 | 0.86 |
| **Macro Average** | What is the Average Performance of Recall ? (Correct Classification of (Not Purchased) to the Total Input of (Not Purchased) in the Test Set) | (Recall (Not Purchased) + Recall (Purchased)) / 2 | 0.87 |
| **Macro Average** | What is the Average Performance of F1 Measure ? (What is the Overall Performance) | (F1 (Not Purchased) + F2 (Purchased)) / 2 | 0.86 |
| **Weighted Average** | **What is the sum of Product of Proportion Rate (Weight) of each class Through Precision ?** | **Weighted Average refers to Proportion of Individuals** |  |
| **Weighted Average** | Weighted Average for Precision | Precision (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  Precision (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.87 |
| **Weighted Average** | Weighted Average for Recall | Recall (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  Recall (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.87 |
| **Weighted Average** | Weighted Average for F1 Measure | F1 (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  F2 (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.87 |

**Classification – Random Forest**

|  |  |  |  |
| --- | --- | --- | --- |
| **Evaluation Metrics** | **Purpose / Explanation** | **Formula** | **Calculated Value Based on Formula** |
| **Accuracy** | What is the Percentage of Correct Classification of Both Not Purchased and Purchased to the Total Input of the Test Set ? | Accuracy = T(Not Purchased) + T(Purchased) / T(Not Purchased) + T(Purchased) + F(Not Purchased) + F(Purchased) | 0.90 |
| **Recall** | What is the Percentage of Correct Classification of (Not Purchased) to the Total Input of (Not Purchased) in the Test Set ? | Recall = T(Not Purchased) / T(Not Purchased) + F(Not Purchased) | 0.92 |
| **Recall** | What is the Percentage of Correct Classification of (Purchased) to the Total Input of (Purchased) in the Test Set ? | Recall = T(Purchased) / T(Purchased) + F(Purchased) | 0.88 |
| **Precision** | What is the Percentage of Correctly Classified as (Not Purchased) and Wrongly Classified as (Not Purchased) in the Test Set ? | Precision = T(Not Purchased) / T(Not Purchased) + F(Purchased) | 0.93 |
| **Precision** | What is the Percentage of Correctly Classified as (Not Purchased) and Wrongly Classified as (Purchased) in the Test Set ? | Precision = T(Purchased) / T(Purchased) + F(Not Purchased) | 0.86 |
| **F1 Measure** | What is the Overall Performance of Not Purchased ? | 2 x (Recall x Precision)  -------------------------- (Recall + Precision) | 0.92 |
| **F1 Measure** | What is the Overall Performance of Purchased ? | 2 x (Recall x Precision)  -------------------------- (Recall + Precision) | 0.87 |
| **Macro Average** | **What is the sum of Average Performance of each class Through Precision/Recall/F1 Measure ?** | **Macro Average refers to Average Performance of Individuals** |  |
| **Macro Average** | What is the Average Performance of Precision ? (Correctly Classified as (Not Purchased) and Wrongly Classified as (Purchased) in the Test Set) | (Precision (Not Purchased) + Precision (Purchased)) / 2 | 0.89 |
| **Macro Average** | What is the Average Performance of Recall ? (Correct Classification of (Not Purchased) to the Total Input of (Not Purchased) in the Test Set) | (Recall (Not Purchased) + Recall (Purchased)) / 2 | 0.90 |
| **Macro Average** | What is the Average Performance of F1 Measure ? (What is the Overall Performance) | (F1 (Not Purchased) + F2 (Purchased)) / 2 | 0.90 |
| **Weighted Average** | **What is the sum of Product of Proportion Rate (Weight) of each class Through Precision ?** | **Weighted Average refers to Proportion of Individuals** |  |
| **Weighted Average** | Weighted Average for Precision | Precision (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  Precision (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.90 |
| **Weighted Average** | Weighted Average for Recall | Recall (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  Recall (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.90 |
| **Weighted Average** | Weighted Average for F1 Measure | F1 (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  F2 (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.90 |

**Classification – SVM**

|  |  |  |  |
| --- | --- | --- | --- |
| **Evaluation Metrics** | **Purpose / Explanation** | **Formula** | **Calculated Value Based on Formula** |
| **Accuracy** | What is the Percentage of Correct Classification of Both Not Purchased and Purchased to the Total Input of the Test Set ? | Accuracy = T(Not Purchased) + T(Purchased) / T(Not Purchased) + T(Purchased) + F(Not Purchased) + F(Purchased) | 0.78 |
| **Recall** | What is the Percentage of Correct Classification of (Not Purchased) to the Total Input of (Not Purchased) in the Test Set ? | Recall = T(Not Purchased) / T(Not Purchased) + F(Not Purchased) | 0.96 |
| **Recall** | What is the Percentage of Correct Classification of (Purchased) to the Total Input of (Purchased) in the Test Set ? | Recall = T(Purchased) / T(Purchased) + F(Purchased) | 0.47 |
| **Precision** | What is the Percentage of Correctly Classified as (Not Purchased) and Wrongly Classified as (Not Purchased) in the Test Set ? | Precision = T(Not Purchased) / T(Not Purchased) + F(Purchased) | 0.76 |
| **Precision** | What is the Percentage of Correctly Classified as (Not Purchased) and Wrongly Classified as (Purchased) in the Test Set ? | Precision = T(Purchased) / T(Purchased) + F(Not Purchased) | 0.88 |
| **F1 Measure** | What is the Overall Performance of Not Purchased ? | 2 x (Recall x Precision)  -------------------------- (Recall + Precision) | 0.85 |
| **F1 Measure** | What is the Overall Performance of Purchased ? | 2 x (Recall x Precision)  -------------------------- (Recall + Precision) | 0.61 |
| **Macro Average** | **What is the sum of Average Performance of each class Through Precision/Recall/F1 Measure ?** | **Macro Average refers to Average Performance of Individuals** |  |
| **Macro Average** | What is the Average Performance of Precision ? (Correctly Classified as (Not Purchased) and Wrongly Classified as (Purchased) in the Test Set) | (Precision (Not Purchased) + Precision (Purchased)) / 2 | 0.82 |
| **Macro Average** | What is the Average Performance of Recall ? (Correct Classification of (Not Purchased) to the Total Input of (Not Purchased) in the Test Set) | (Recall (Not Purchased) + Recall (Purchased)) / 2 | 0.72 |
| **Macro Average** | What is the Average Performance of F1 Measure ? (What is the Overall Performance) | (F1 (Not Purchased) + F2 (Purchased)) / 2 | 0.73 |
| **Weighted Average** | **What is the sum of Product of Proportion Rate (Weight) of each class Through Precision ?** | **Weighted Average refers to Proportion of Individuals** |  |
| **Weighted Average** | Weighted Average for Precision | Precision (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  Precision (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.81 |
| **Weighted Average** | Weighted Average for Recall | Recall (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  Recall (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.78 |
| **Weighted Average** | Weighted Average for F1 Measure | F1 (Not Purchased) x (Total Input of Not Purchased / Total No. of Input of Dataset)  +  F2 (Purchased) x (Total Input of Purchased / Total No. of Input of Dataset) | 0.76 |

**Overall Result**

* The Best Model would be saved as Deployment Phase File for the One which has Highest Accuracy Value as Highlighted Below :

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Methodology** | **Evaluation Metrics** | **Calculated Value Based on Formula** |
| **1** | **Support Vector Machine** | **Accuracy** | **0.78** |
| 2 | Decision Tree | Accuracy | 0.87 |
| **3** | **Random Forest** | **Accuracy** | **0.90** |