**Section I: Dataset-IRIS,**

**URL: https://archive.ics.uci.edu/ml/datasets/Iris**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Number of instances** | **Number of attributes** | **Number of attributes used** | **Classification method** | **Parameters** | **Train/Test Ratio** | **Test Accuracy** | **Test Error** | **Precision** | **Recall** | **FP-Measure** |
| 150 | 4 and 1 class | 4 and 1 class | Decision Tree | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo | 60/40 | 93.44 | 6.55 | 0.93 | 0.9344 | 0.0357 |
| 150 | 4 and 1 class | 2 attributes and 1 class | Decision Tree | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo | 60/40 | 95.08 | 4.918 | 0.952 | 0.9504 | 0.0258 |
| 150 | 4 and 1 class | 4 and 1 class | Random Forest | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo,numTrees,Futuresubsetstartegy | 75/25 | 93.1818 | 6.818 | 0.9327 | 0.9318 | 0.0365 |
| 150 | 4 and 1 class | 2 and 1 class | Random Forest | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo,numTrees,Futuresubsetstartegy | 75/25 | 95.4545 | 4.545 | 0.9615 | 0.9545 | 0.01515 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Number of instances | Number of attributes | Number of attributes used | Classification method | Parameters | Train/Test Ratio | Test Accuracy | Test Error | Precision | Recall | FP-Measure |
| 150 | 4 and 1 class | 4 and 1 class | Decision Tree | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo | 75/25 | 97.72 | 2.27 | 0.979 | 0.977 | 0.008 |
| 150 | 4 and 1 class | 3 attributes and 1 class | Decision Tree | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo | 75/25 | 90.9 | 9.09 | 0.911 | 0.909 | 0.04 |
| 150 | 4 and 1 class | 4 and 1 class | Random Forest | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo,numTrees,Futuresubsetstartegy | 65/35 | 92.7272 | 7.2727 | 0.9272 | 0.9272 | 0.038 |
| 150 | 4 and 1 class | 2 and 1 class | Random Forest | Impurity, maxDepth, maxBins, Numberofclasses required, categoricalFeaturesInfo,numTrees,Futuresubsetstartegy | 65/35 | 92.72 | 7.27727 | 0.9324 | 0.9272 | 0.033 |

**Databricks URL for DecisionTrees including Principal Component Analysis(Dimensionality reduction)**

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/855389465014467/850335936676778/4847935780219074/latest.html>

**Databricks URL for RandomForest including Principal Component Analysis(Dimensionality reduction)**

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/855389465014467/850335936676798/4847935780219074/latest.html>

**Section II:**

Dataset : WholeSale Customers

Url: <https://archive.ics.uci.edu/ml/datasets/Wholesale+customers>

Databricks URL:

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/855389465014467/949768318295698/4847935780219074/latest.html>Output:

|  |  |  |
| --- | --- | --- |
| Number of Clusters | Number Of Iterations | Within Set Sum of Squared Errors |
| **3** | **30** | 8.154092862484204E10 |
| **2** | **20** | 1.2106080014193402E11 |
| **2** | **1000** | 1.1955543710406216E11 |
| **3** | **1000** | 8.074159128232605E10 |
| **4** | **5000** | 6.589923833859655E10 |