

Purchase Prediction Machine Learning Classification Model Create

Decision Tree Classification Method Best Model is

Best Model = {'criterion': 'gini', 'splitter': 'random'}

Accuracy = **0.85**

Precision = {Purchased: **0.77**, Not-Purchased:**0.90**}

Recall = {Purchased: **0.80**, Not-Purchased:**0.87**}

F1_Score = {Purchased: **0.79**, Not-Purchased:**0.88**}

Purchased-TestDataSet=41

Not-Purchased-TestDataSet=79

Is this a Good Model?

1. The accuracy of 85% is solid for many classification tasks.
2. Precision and recall are balanced and relatively high.
3. The model uses criterion 'gini' and a 'random' splitter, common in decision tree algorithms.

Random Forest Classification Method Best Model is

Best Model = {'criterion': 'gini', 'max_features': 'sqrt', 'n_estimators': 50, 'random_state': 0}

Accuracy = **0.91**

Precision = {Purchased: **0.86**, Not-Purchased:**0.94**}

Recall = {Purchased: **0.88**, Not-Purchased:**0.92**}

F1_Score = {Purchased: **0.87**, Not-Purchased:**0.93**}

Purchased-TestDataSet=41

Not-Purchased-TestDataSet=79

Is This Model Good?

1. An accuracy of 91% is excellent in many contexts, indicating the model makes very few errors.
2. High precision and recall values for both classes suggest reliable and balanced detection of purchased and not-purchased cases.
3. The F1 scores close to 0.9 further confirm strong performance.

Support Vector Machine Classification Method Best Model is

Best Model = {'gamma': 'scale', 'kernel': 'rbf', 'random_state': 0}

Accuracy = **0.79**

Precision = {Purchased: **0.90**, Not-Purchased:**0.77**}

Recall = {Purchased: 0.44, Not-Purchased: 0.97}
F1_Score = {Purchased: 0.59, Not-Purchased: 0.86}
Purchased-TestDataSet=41
Not-Purchased-TestDataSet=79

Is This Model Good?

1. The overall accuracy (79%) is average to somewhat below average compared to your previous models.
2. The model is strong at identifying non-purchases but weak at detecting actual purchases (low recall for purchased).
3. This could lead to many missed purchase cases, which is critical in many real-world purchase prediction tasks.

Logistic Regression Classification Method Best Model is

Best Model = {'multi_class': 'auto', 'penalty': None}
Accuracy = 0.89
Precision = {Purchased: 0.87, Not-Purchased: 0.90}
Recall = {Purchased: 0.80, Not-Purchased: 0.94}
F1_Score = {Purchased: 0.84, Not-Purchased: 0.92}
Purchased-TestDataSet=41
Not-Purchased-TestDataSet=79

Is This Model Good?

1. The 89% accuracy shows the model performs well overall.
2. Precision and recall values indicate reliable detection and classification for both classes.
3. F1 scores close to or above 0.8 reflect good balance and robustness.

The Random Forest model is the best classification model for your dataset based on accuracy, precision, recall, and F1 scores. It balances detection of both purchased and not-purchased classes well and has the highest scores across most metrics.