Model Development Phase Template

Date	17 July 2024
Team ID	SWTID1720190389
Project Title	E-Commerce Shipping Prediction
Maximum Marks	4 Marks

Initial Model Training Code, Model Validation and Evaluation Report

The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

Initial Model Training Code:

```
lj = LogisticRegression(solver="liblinear").fit(xTrain,yTrain)
gnb = GaussianNB().fit(xTrain,yTrain)
knnc = KNeighborsClassifier().fit(xTrain,yTrain)
cartc = DecisionTreeClassifier(random_state=42).fit(xTrain,yTrain)
rfc = RandomForestClassifier(random_state=42,verbose=False).fit(xTrain,yTrain)
gbmc = GradientBoostingClassifier(verbose=False).fit(xTrain,yTrain)
xgbc = XGBClassifier().fit(xTrain,yTrain)
catbc = CatBoostClassifier(verbose=False).fit(xTrain,yTrain)
```

```
modelsc = [lj,gnb,knnc,cartc,rfc,gbmc,xgbc,catbc]
```

```
from sklearn.metrics import classification_report, confusion_matrix
from sklearn.model_selection import cross_val_score
from sklearn.metrics import roc_curve
for model in modelsc:
    name = model.__class__.__name__
    predict = model.predict(xTest)
    R2CV = cross_val_score(model,xTest,yTest,cv=10,verbose=False).mean()
    error = -cross_val_score(model,xTest,yTest,cv=10,scoring="neg_mean_squared_error",verbose=False).mean()
    roc = roc_curve(yTest,predict)
    Classification_Report = classification_report(yTest,predict)
    confusion_matrics = confusion_matrix(yTest,predict)
    print(name + ": ")
    print("-" * 10)
    print("classification_report -->")
    print(Classification_Report)
    print("confusion_matrix -->")
    print(confusion_matrics)
    print("ACC-->",accuracy_score(yTest,predict))
    print("R2CV-->",R2CV)
    print("MEAN SQUARED ERROR-->",np.sqrt(error))
    print("ROC-->",roc)
    print("-" * 30)
```

Model Validation and Evaluation Report:

Model	Classification Report	Accuracy	Confusion Matrix
Logistic Regression	classification_report --> precision recall f1-score support 0 0.60 0.67 0.63 911 1 0.74 0.68 0.71 1288 accuracy 0.68 2199 macro avg 0.67 0.67 0.67 2199 weighted avg 0.68 0.68 0.68 2199	67.62	confusion_matrix --> [[607 304] [408 880]]
GaussianNB	classification_report --> precision recall f1-score support 0 0.53 1.00 0.70 911 1 1.00 0.38 0.55 1288 accuracy 0.64 2199 macro avg 0.77 0.69 0.62 2199 weighted avg 0.81 0.64 0.61 2199	63.85	confusion_matrix --> [[911 0] [795 493]]
KNN	classification_report --> precision recall f1-score support 0 0.60 0.63 0.61 911 1 0.73 0.70 0.71 1288 accuracy 0.67 2199 macro avg 0.66 0.66 0.66 2199 weighted avg 0.67 0.67 0.67 2199	67.03	confusion_matrix --> [[576 335] [390 898]]
Decision Tree	classification_report --> precision recall f1-score support 0 0.58 0.53 0.56 911 1 0.69 0.73 0.71 1288 accuracy 0.65 2199 macro avg 0.63 0.63 0.63 2199 weighted avg 0.64 0.65 0.64 2199	64.71	confusion_matrix --> [[486 425] [351 937]]
Random Forest	classification_report --> precision recall f1-score support 0 0.59 0.68 0.64 911 1 0.75 0.67 0.71 1288 accuracy 0.67 2199 macro avg 0.67 0.68 0.67 2199 weighted avg 0.68 0.67 0.68 2199	67.44	confusion_matrix --> [[623 288] [428 860]]

Gradient Boosting	classification_report --> precision recall f1-score support 0 0.59 0.88 0.71 911 1 0.87 0.57 0.69 1288 accuracy 0.70 2199 macro avg 0.73 0.73 0.70 2199 weighted avg 0.76 0.70 0.70 2199	69.99	confusion_matrix --> [[806 105] [555 733]]
XGBoost	classification_report --> precision recall f1-score support 0 0.59 0.67 0.63 911 1 0.74 0.67 0.70 1288 accuracy 0.67 2199 macro avg 0.66 0.67 0.66 2199 weighted avg 0.68 0.67 0.67 2199	66.89	confusion_matrix --> [[608 303] [425 863]]
CatBoost	classification_report --> precision recall f1-score support 0 0.59 0.75 0.66 911 1 0.78 0.64 0.70 1288 accuracy 0.68 2199 macro avg 0.69 0.69 0.68 2199 weighted avg 0.70 0.68 0.68 2199	68.26	confusion_matrix --> [[683 228] [470 818]]