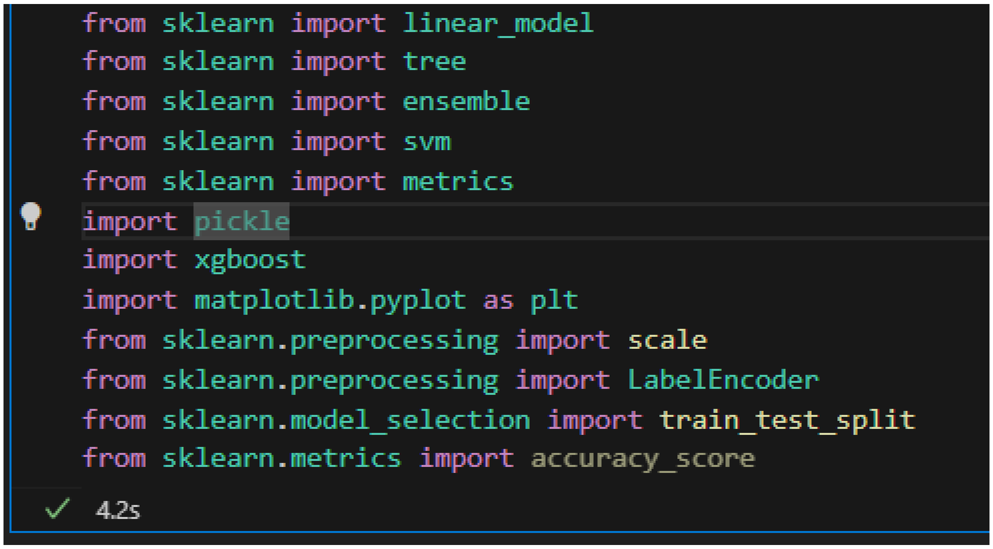
Model Development Phase Template

Date	15th October 2024
Team ID	LTVIP2024TMID24968
Project Title	Traffic Telligence - Advanced Traffic Volume Estimation with MachineLearning
Maximum Marks	4 Marks

Initial Model Training Code, Model Validation and Evaluation Report

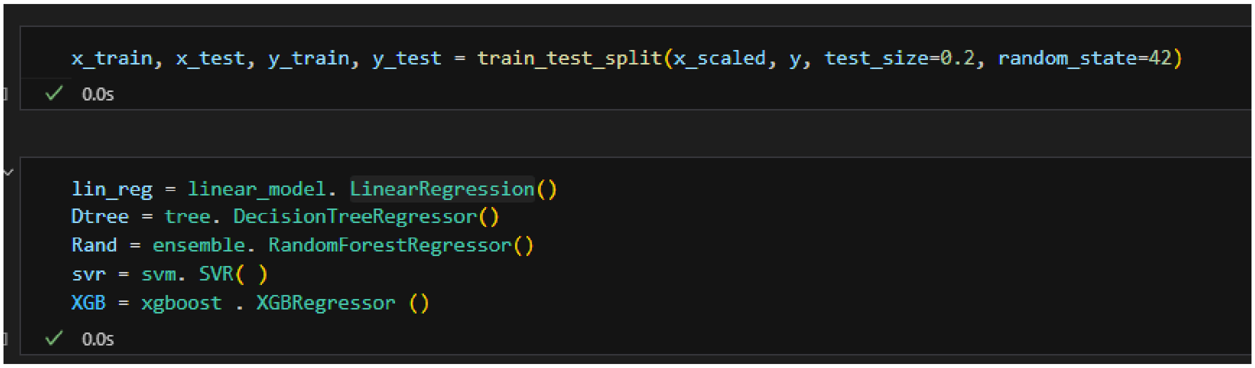
The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include accuracy, and confusion matrices for multiple models like Random Forest Regression, Support Vector Regression, XGBoost, Linear Regression, etc,., presented through respective screenshots.

Initial Model Training Code:



```
from sklearn import linear_model
from sklearn import tree
from sklearn import ensemble
from sklearn import svm
from sklearn import metrics
import pickle
import xgboost
import matplotlib.pyplot as plt
from sklearn.preprocessing import scale
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
```

✓ 4.2s

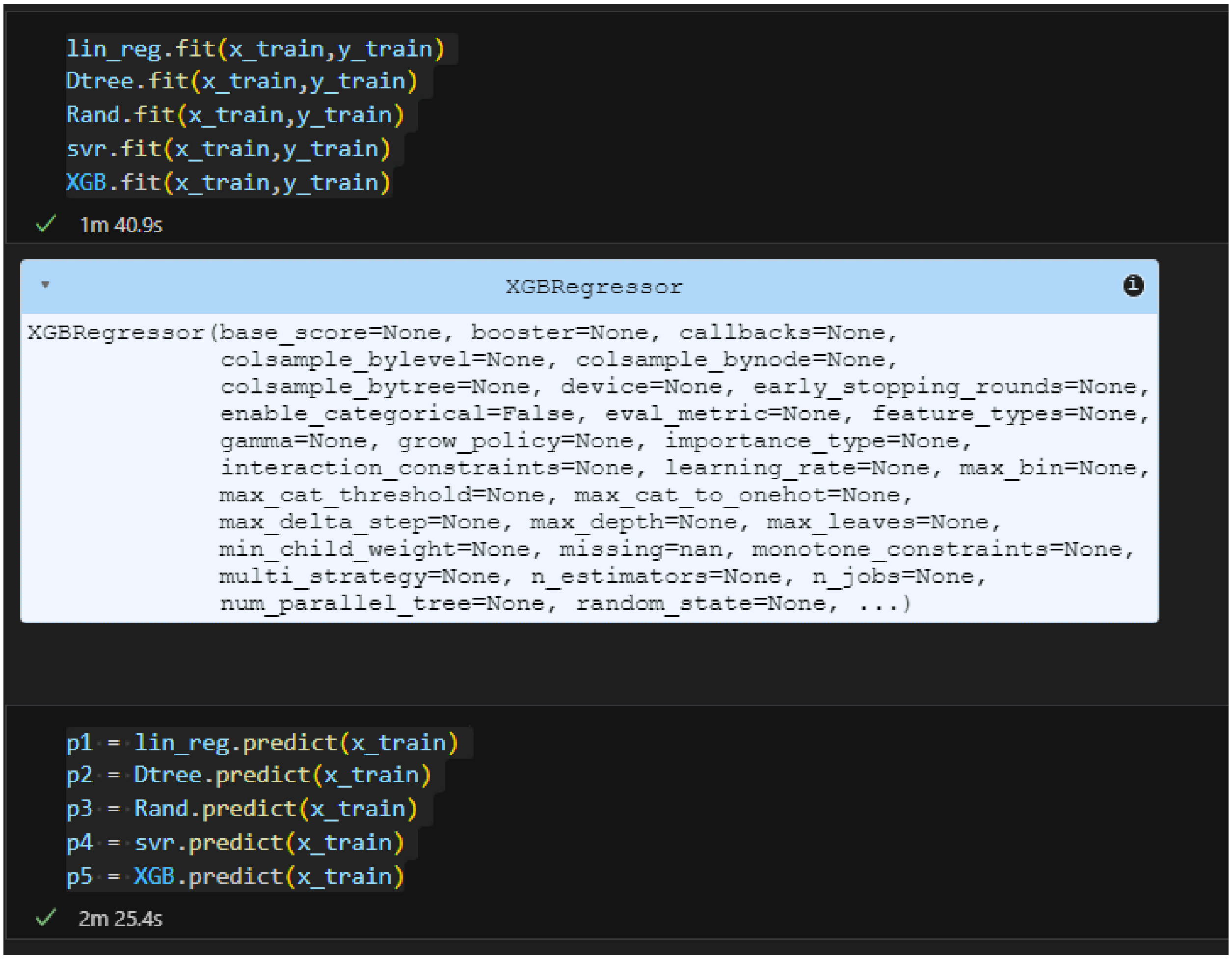


```
x_train, x_test, y_train, y_test = train_test_split(x_scaled, y, test_size=0.2, random_state=42)
```

✓ 0.0s

```
lin_reg = linear_model. LinearRegression()
Dtree = tree. DecisionTreeRegressor()
Rand = ensemble. RandomForestRegressor()
svr = svm. SVR( )
XGB = xgboost . XGBRegressor ()
```

✓ 0.0s



```
lin_reg.fit(x_train,y_train)
Dtree.fit(x_train,y_train)
Rand.fit(x_train,y_train)
svr.fit(x_train,y_train)
XGB.fit(x_train,y_train)
```

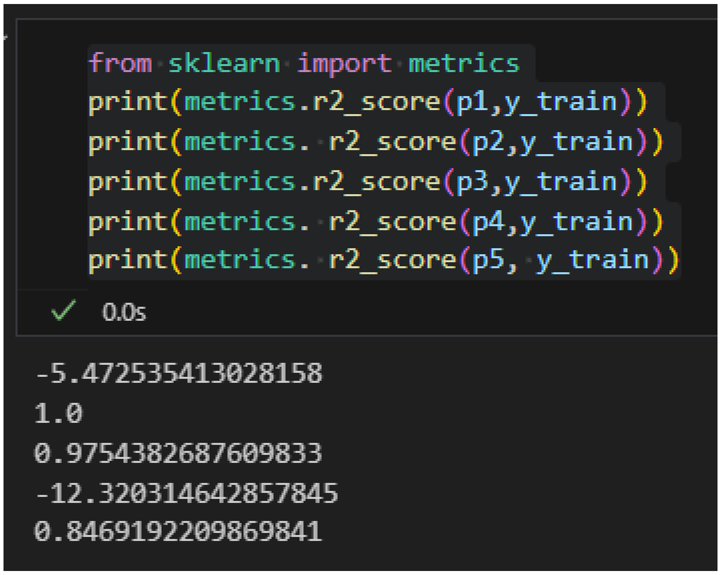
✓ 1m 40.9s

```
XGBRegressor

XGBRegressor(base_score=None, booster=None, callbacks=None,
             colsample_bylevel=None, colsample_bynode=None,
             colsample_bytree=None, device=None, early_stopping_rounds=None,
             enable_categorical=False, eval_metric=None, feature_types=None,
             gamma=None, grow_policy=None, importance_type=None,
             interaction_constraints=None, learning_rate=None, max_bin=None,
             max_cat_threshold=None, max_cat_to_onehot=None,
             max_delta_step=None, max_depth=None, max_leaves=None,
             min_child_weight=None, missing=nan, monotone_constraints=None,
             multi_strategy=None, n_estimators=None, n_jobs=None,
             num_parallel_tree=None, random_state=None, ...)
```

```
p1 = lin_reg.predict(x_train)
p2 = Dtree.predict(x_train)
p3 = Rand.predict(x_train)
p4 = svr.predict(x_train)
p5 = XGB.predict(x_train)
```

✓ 2m 25.4s

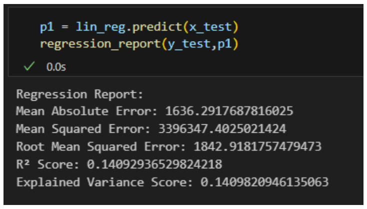
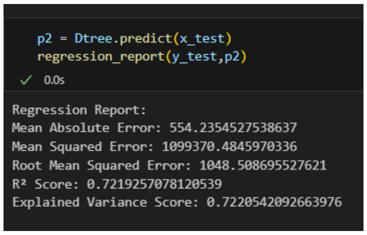


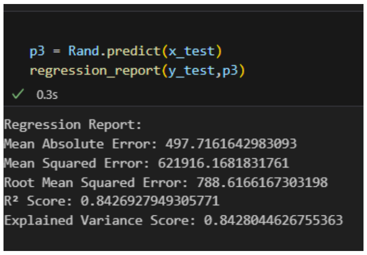
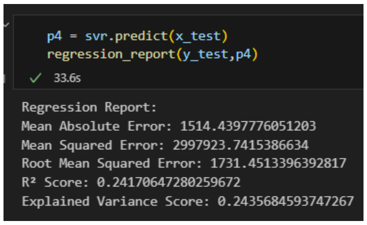
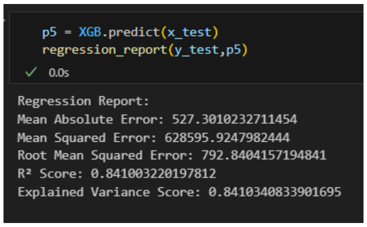
```
from sklearn import metrics
print(metrics.r2_score(p1,y_train))
print(metrics. r2_score(p2,y_train))
print(metrics.r2_score(p3,y_train))
print(metrics. r2_score(p4,y_train))
print(metrics. r2_score(p5, y_train))
```

✓ 0.0s

```
-5.472535413028158
1.0
0.9754382687609833
-12.320314642857845
0.8469192209869841
```

Model Validation and Evaluation Report:

Model	**Regression Report**	**R2_Score**
Linear Regression	 p1 = lin_reg.predict(x_test) regression_report(y_test,p1) ✓ 0.0s Regression Report: Mean Absolute Error: 1636.2917687816025 Mean Squared Error: 3396347.4025021424 Root Mean Squared Error: 1842.9181757479473 R² Score: 0.14092936529824218 Explained Variance Score: 0.1409820946135063	14%
Decision Tree Regressor	 p2 = Dtree.predict(x_test) regression_report(y_test,p2) ✓ 0.0s Regression Report: Mean Absolute Error: 554.2354527538637 Mean Squared Error: 1099370.4845970336 Root Mean Squared Error: 1048.508695527621 R² Score: 0.7219257078120539 Explained Variance Score: 0.7220542092663976	72%

Random Forest Regressor	 p3 = Rand.predict(x_test) regression_report(y_test,p3) ✓ 0.3s Regression Report: Mean Absolute Error: 497.7161642983093 Mean Squared Error: 621916.1681831761 Root Mean Squared Error: 788.6166167303198 R² Score: 0.8426927949305771 Explained Variance Score: 0.8428044626755363	84%
SVR	 p4 = svr.predict(x_test) regression_report(y_test,p4) ✓ 33.6s Regression Report: Mean Absolute Error: 1514.4397776051203 Mean Squared Error: 2997923.7415386634 Root Mean Squared Error: 1731.4513396392817 R² Score: 0.24170647280259672 Explained Variance Score: 0.2435684593747267	24%
XGB	 p5 = XGB.predict(x_test) regression_report(y_test,p5) ✓ 0.0s Regression Report: Mean Absolute Error: 527.3010232711454 Mean Squared Error: 628595.9247982444 Root Mean Squared Error: 792.8404157194841 R² Score: 0.841003220197812 Explained Variance Score: 0.8410340833901695	84%