Model Development Phase

Date	15 March 2024
Team ID	SWTID1720078183
Project Title	Predictive Modeling for Fleet Fuel Management using Machine Learning
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:

```python
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor
from sklearn.tree import DecisionTreeRegressor
from sklearn.svm import SVR
from sklearn.metrics import r2_score

# Split the data
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.3, random_state=42)

# Initialize models
models = {
    "Linear Regression": LinearRegression(),
    "Random Forest": RandomForestRegressor(),
    "Decision Tree": DecisionTreeRegressor(),
    "Support Vector Regressor": SVR()
}
```

```python
# Initialize and train a RandomForestRegressor
rf = RandomForestRegressor(random_state=42)
rf.fit(X_train, y_train)
y_pred_rf = rf.predict(X_test)
```

```python
# Initialize HistGradientBoostingRegressor
hgb_reg = HistGradientBoostingRegressor(random_state=42)
hgb_reg.fit(X_train, y_train)
y_pred_hgb = hgb_reg.predict(X_test)
```

```
: # Print results
for name, score in results.items():
    print(f"{name}: accuracy = {score:.4f}")

Linear Regression: accuracy = 0.1072
Random Forest: accuracy = 0.6021
Decision Tree: accuracy = 0.0619
Support Vector Regressor: accuracy = 0.2546
```

Model Validation and Evaluation Report:

Model	Classification Report	Accuracy	Confusion Matrix
Random Forest Regressor	```[67]: # Calculate the R2 score\nprint(f"Random Forest Regressor R² Score: {r2_rf}")\nprint(f"Random Forest Regressor Mean Squared Error: {mse}")\nprint(f"Random Forest Regressor Mean Absolute Error: {mae}")\n\nRandom Forest Regressor R² Score: 0.5889485334633298\nRandom Forest Regressor Mean Squared Error: 0.3442496727207985\nRandom Forest Regressor Mean Absolute Error: 0.42989886039886077```	60	```print("Classification Report:")\nprint(class_report)\n\nprint("Confusion Matrix:")\nprint(conf_matrix)\n\nClassification Report:\n precision recall f1-score support\n 0 0.77 0.72 0.74 32\n 1 0.41 0.44 0.43 27\n 2 0.34 0.33 0.34 30\n 3 0.59 0.61 0.60 28\n\n accuracy 0.53 117\n macro avg 0.53 0.53 0.53 117\nweighted avg 0.53 0.53 0.53 117\n\nConfusion Matrix:\n[[23 7 0 2]\n [4 12 10 1]\n [3 8 10 9]\n [0 2 9 17]]```
HistGradient Boosting Regressor	```# Calculate the R2 score for the best model\nprint(r2_best_gbr)\nprint(mse)\nprint(mae)\n\n-9.383042721509119\n13.229675735887355\n3.5036213550286974```	53	```print("Classification Report:")\nprint(class_report)\n\nprint("Confusion Matrix:")\nprint(conf_matrix)\n\nClassification Report:\n precision recall f1-score support\n 0 0.77 0.72 0.74 32\n 1 0.41 0.44 0.43 27\n 2 0.34 0.33 0.34 30\n 3 0.59 0.61 0.60 28\n\n accuracy 0.53 117\n macro avg 0.53 0.53 0.53 117\nweighted avg 0.53 0.53 0.53 117\n\nConfusion Matrix:\n[[23 7 0 2]\n [4 12 10 1]\n [3 8 10 9]\n [0 2 9 17]]```