Machine Learning Internship Program at Prodigy InfoTech:

Steve W steve94429@gmail.com 9442933527

Task 1:

Implement a linear regression model to predict the prices of houses based on their square footage and the number of bedrooms and bathrooms.

In [1]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

train_data = pd.read_csv('train.csv')


features = ['GrLivArea', 'BedroomAbvGr', 'FullBath']
target = 'SalePrice'


X = train_data[features]
y = train_data[target]


X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.2, random_state=42)


model = LinearRegression()
model.fit(X_train, y_train)


y_val_pred = model.predict(X_val)


mse = mean_squared_error(y_val, y_val_pred)
r2 = r2_score(y_val, y_val_pred)

print(f'Mean Squared Error: {mse}')
print(f'R-squared: {r2}')


test_data = pd.read_csv('test.csv')


X_test = test_data[features]


test_predictions = model.predict(X_test)


test_predictions_df = pd.DataFrame({'Id': test_data['Id'], 'SalePrice': test_predictions})
print(test_predictions_df.head())
```

```
Mean Squared Error: 2806426667.247852
R-squared: 0.6341189942328374
     Id      SalePrice
0  1461  122173.313104
1  1462  140561.538683
2  1463  201783.754896
3  1464  199183.097221
4  1465  192133.739106
```

In [ ]:
```
```