Task-1

May 21, 2025

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.metrics import mean_squared_error, r2_score, accuracy_score,␣
 ↪classification_report
from sklearn.linear_model import LinearRegression, LogisticRegression
from sklearn.ensemble import RandomForestRegressor, RandomForestClassifier
from sklearn.tree import DecisionTreeRegressor
```

[2]:
```
df=pd.read_csv("Dataset .csv")
```

[3]:
```
df.head()
```

[3]:
```
   Restaurant ID         Restaurant Name  Country Code              City  \
0        6317637        Le Petit Souffle           162       Makati City
1        6304287        Izakaya Kikufuji           162       Makati City
2        6300002  Heat - Edsa Shangri-La           162  Mandaluyong City
3        6318506                   Ooma           162  Mandaluyong City
4        6314302             Sambo Kojin           162  Mandaluyong City

                                        Address  \
0  Third Floor, Century City Mall, Kalayaan Avenu…
1  Little Tokyo, 2277 Chino Roces Avenue, Legaspi…
2  Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal…
3  Third Floor, Mega Fashion Hall, SM Megamall, O…
4  Third Floor, Mega Atrium, SM Megamall, Ortigas…

                                Locality  \
0   Century City Mall, Poblacion, Makati City
1  Little Tokyo, Legaspi Village, Makati City
2  Edsa Shangri-La, Ortigas, Mandaluyong City
3      SM Megamall, Ortigas, Mandaluyong City
4      SM Megamall, Ortigas, Mandaluyong City

                                Locality Verbose   Longitude   Latitude  \
```

```
0  Century City Mall, Poblacion, Makati City, Mak…  121.027535  14.565443
1  Little Tokyo, Legaspi Village, Makati City, Ma…  121.014101  14.553708
2  Edsa Shangri-La, Ortigas, Mandaluyong City, Ma…  121.056831  14.581404
3  SM Megamall, Ortigas, Mandaluyong City, Mandal…  121.056475  14.585318
4  SM Megamall, Ortigas, Mandaluyong City, Mandal…  121.057508  14.584450

                          Cuisines  …          Currency Has Table booking  \
0       French, Japanese, Desserts  …  Botswana Pula(P)               Yes
1                         Japanese  …  Botswana Pula(P)               Yes
2  Seafood, Asian, Filipino, Indian  …  Botswana Pula(P)               Yes
3                  Japanese, Sushi  …  Botswana Pula(P)                No
4                 Japanese, Korean  …  Botswana Pula(P)               Yes

  Has Online delivery Is delivering now Switch to order menu Price range  \
0                  No                No                    No           3
1                  No                No                    No           3
2                  No                No                    No           4
3                  No                No                    No           4
4                  No                No                    No           4

   Aggregate rating  Rating color Rating text Votes
0               4.8    Dark Green   Excellent   314
1               4.5    Dark Green   Excellent   591
2               4.4         Green   Very Good   270
3               4.9    Dark Green   Excellent   365
4               4.8    Dark Green   Excellent   229

[5 rows x 21 columns]
```

[4]:
```
df.tail(5)
```

[4]:
```
      Restaurant ID           Restaurant Name  Country Code       City  \
9546        5915730              Naml  Gurme           208   stanbul
9547        5908749             Ceviz A ac           208   stanbul
9548        5915807                   Huqqa           208   stanbul
9549        5916112              A  k Kahve           208   stanbul
9550        5927402  Walter's Coffee Roastery           208   stanbul

                                              Address     Locality  \
9546  Kemanke  Karamustafa Pa a Mahallesi, R ht m …     Karak _y
9547  Ko uyolu Mahallesi, Muhittin  st _nda  Cadd…    Ko uyolu
9548  Kuru _e me Mahallesi, Muallim Naci Caddesi, N…  Kuru _e me
9549  Kuru _e me Mahallesi, Muallim Naci Caddesi, N…  Kuru _e me
9550  Cafea a Mahallesi, Bademalt  Sokak, No 21/B, …         Moda

           Locality Verbose  Longitude   Latitude  \
9546     Karak _y,  stanbul  28.977392  41.022793
```

```
9547    Ko uyolu,  stanbul  29.041297  41.009847
9548  Kuru _e me,  stanbul  29.034640  41.055817
9549  Kuru _e me,  stanbul  29.036019  41.057979
9550        Moda,  stanbul  29.026016  40.984776

                          Cuisines  …          Currency  \
9546                       Turkish  …  Turkish Lira(TL)
9547  World Cuisine, Patisserie, Cafe  …  Turkish Lira(TL)
9548          Italian, World Cuisine  …  Turkish Lira(TL)
9549                Restaurant Cafe  …  Turkish Lira(TL)
9550                           Cafe  …  Turkish Lira(TL)

     Has Table booking Has Online delivery Is delivering now  \
9546                No                  No                No
9547                No                  No                No
9548                No                  No                No
9549                No                  No                No
9550                No                  No                No

     Switch to order menu Price range  Aggregate rating  Rating color  \
9546                   No           3               4.1         Green
9547                   No           3               4.2         Green
9548                   No           4               3.7        Yellow
9549                   No           4               4.0         Green
9550                   No           2               4.0         Green

     Rating text Votes
9546   Very Good   788
9547   Very Good  1034
9548        Good   661
9549   Very Good   901
9550   Very Good   591

[5 rows x 21 columns]
```

[5]:
```
df.shape # checking number of columns and rows
```

[5]:
```
(9551, 21)
```

[6]:
```
df.isnull()
```

[6]:
```
      Restaurant ID  Restaurant Name  Country Code   City  Address  Locality  \
0             False            False         False  False    False     False
1             False            False         False  False    False     False
2             False            False         False  False    False     False
3             False            False         False  False    False     False
4             False            False         False  False    False     False
```

```
…             …               …            …    …      …       …
9546          False             False          False  False    False     False
9547          False             False          False  False    False     False
9548          False             False          False  False    False     False
9549          False             False          False  False    False     False
9550          False             False          False  False    False     False

      Locality Verbose  Longitude  Latitude  Cuisines  …  Currency  \
0                False      False     False     False  …     False
1                False      False     False     False  …     False
2                False      False     False     False  …     False
3                False      False     False     False  …     False
4                False      False     False     False  …     False
…                …          …         …         …      …  …
9546             False      False     False     False  …     False
9547             False      False     False     False  …     False
9548             False      False     False     False  …     False
9549             False      False     False     False  …     False
9550             False      False     False     False  …     False

      Has Table booking  Has Online delivery  Is delivering now  \
0                 False                False              False
1                 False                False              False
2                 False                False              False
3                 False                False              False
4                 False                False              False
…                 …                    …                  …
9546              False                False              False
9547              False                False              False
9548              False                False              False
9549              False                False              False
9550              False                False              False

      Switch to order menu  Price range  Aggregate rating  Rating color  \
0                    False        False             False         False
1                    False        False             False         False
2                    False        False             False         False
3                    False        False             False         False
4                    False        False             False         False
…                    …            …                 …             …
9546                 False        False             False         False
9547                 False        False             False         False
9548                 False        False             False         False
9549                 False        False             False         False
9550                 False        False             False         False

      Rating text  Votes
```

```
0           False  False
1           False  False
2           False  False
3           False  False
4           False  False
…             …      …
9546        False  False
9547        False  False
9548        False  False
9549        False  False
9550        False  False

[9551 rows x 21 columns]
```

[7]:
```
drop_cols = ['Restaurant ID', 'Restaurant Name', 'Address',␣
 ↪'Locality','Locality Verbose', 'Rating color', 'Rating text']
df.drop(columns=drop_cols, inplace=True, errors='ignore')
```

[8]:
```
df.head(5)
```

[8]:
```
   Country Code              City   Longitude   Latitude  \
0           162       Makati City  121.027535  14.565443
1           162       Makati City  121.014101  14.553708
2           162  Mandaluyong City  121.056831  14.581404
3           162  Mandaluyong City  121.056475  14.585318
4           162  Mandaluyong City  121.057508  14.584450

                          Cuisines  Average Cost for two          Currency  \
0        French, Japanese, Desserts                  1100  Botswana Pula(P)
1                          Japanese                  1200  Botswana Pula(P)
2  Seafood, Asian, Filipino, Indian                  4000  Botswana Pula(P)
3                   Japanese, Sushi                  1500  Botswana Pula(P)
4                  Japanese, Korean                  1500  Botswana Pula(P)

  Has Table booking Has Online delivery Is delivering now  \
0               Yes                 No                No
1               Yes                 No                No
2               Yes                 No                No
3                No                 No                No
4               Yes                 No                No

  Switch to order menu  Price range  Aggregate rating  Votes
0                   No            3               4.8    314
1                   No            3               4.5    591
2                   No            4               4.4    270
3                   No            4               4.9    365
4                   No            4               4.8    229
```

[9]:
```
df.shape # After dropping irrelavent features
```

[9]:
```
(9551, 14)
```

[10]:
```
df.isnull().sum()
```

[10]:
```
Country Code            0
City                    0
Longitude               0
Latitude                0
Cuisines                9
Average Cost for two    0
Currency                0
Has Table booking       0
Has Online delivery     0
Is delivering now       0
Switch to order menu    0
Price range             0
Aggregate rating        0
Votes                   0
dtype: int64
```

[11]:
```
df.fillna({'Cuisines': df['Cuisines'].mode()[0]}, inplace=True) #replacing null␣
↪values with mode/median/mean
```

[12]:
```
df.head()
```

[12]:

	Country Code	City	Longitude	Latitude	Cuisines	Average Cost for two	Currency	Has Table booking	Has Online delivery	Is delivering now
0	162	Makati City	121.027535	14.565443	French, Japanese, Desserts	1100	Botswana Pula(P)	Yes	No	No
1	162	Makati City	121.014101	14.553708	Japanese	1200	Botswana Pula(P)	Yes	No	No
2	162	Mandaluyong City	121.056831	14.581404	Seafood, Asian, Filipino, Indian	4000	Botswana Pula(P)	Yes	No	No
3	162	Mandaluyong City	121.056475	14.585318	Japanese, Sushi	1500	Botswana Pula(P)	No	No	No
4	162	Mandaluyong City	121.057508	14.584450	Japanese, Korean	1500	Botswana Pula(P)	Yes	No	No

```
  Switch to order menu  Price range  Aggregate rating  Votes
0                   No            3               4.8    314
1                   No            3               4.5    591
2                   No            4               4.4    270
3                   No            4               4.9    365
4                   No            4               4.8    229
```

[13]:
```
df.isnull().sum()
```

[13]:
```
Country Code            0
City                    0
Longitude               0
Latitude                0
Cuisines                0
Average Cost for two    0
Currency                0
Has Table booking       0
Has Online delivery     0
Is delivering now       0
Switch to order menu    0
Price range             0
Aggregate rating        0
Votes                   0
dtype: int64
```

[14]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9551 entries, 0 to 9550
Data columns (total 14 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   Country Code          9551 non-null   int64  
 1   City                  9551 non-null   object 
 2   Longitude             9551 non-null   float64
 3   Latitude              9551 non-null   float64
 4   Cuisines              9551 non-null   object 
 5   Average Cost for two  9551 non-null   int64  
 6   Currency              9551 non-null   object 
 7   Has Table booking     9551 non-null   object 
 8   Has Online delivery   9551 non-null   object 
 9   Is delivering now     9551 non-null   object 
 10  Switch to order menu  9551 non-null   object 
 11  Price range           9551 non-null   int64  
 12  Aggregate rating      9551 non-null   float64
 13  Votes                 9551 non-null   int64  
dtypes: float64(3), int64(4), object(7)
```

```
memory usage: 1.0+ MB
```

[15]:
```
le=LabelEncoder()   #converting Categorical values into numericals
for col in df.select_dtypes(include='object').columns:
    df[col] = le.fit_transform(df[col])
print("Preprocessing done. No missing values.")
```

```
Preprocessing done. No missing values.
```

[16]:
```
pip install xgboost
```

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: xgboost in
c:\users\varsh\appdata\roaming\python\python312\site-packages (2.1.4)
Requirement already satisfied: numpy in c:\programdata\anaconda3\lib\site-
packages (from xgboost) (1.26.4)
Requirement already satisfied: scipy in c:\programdata\anaconda3\lib\site-
packages (from xgboost) (1.13.1)
Note: you may need to restart the kernel to use updated packages.
```

[17]:
```
from xgboost import XGBRegressor
```

[18]:
```
X = df.drop('Aggregate rating', axis=1)
y = df['Aggregate rating']

# Train/test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=42)
```

[19]:
```
xgb_model = XGBRegressor(objective='reg:squarederror', n_estimators=100,␣
 ↪max_depth=4, learning_rate=0.1)
xgb_model.fit(X_train, y_train)
```

[19]:
```
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=0.1, max_bin=None,
             max_cat_threshold=None, max_cat_to_onehot=None,
             max_delta_step=None, max_depth=4, max_leaves=None,
             min_child_weight=None, missing=nan, monotone_constraints=None,
             multi_strategy=None, n_estimators=100, n_jobs=None,
             num_parallel_tree=None, random_state=None, …)
```

[20]:
```
y_pred_xgb = xgb_model.predict(X_test)
mse_xgb = mean_squared_error(y_test, y_pred_xgb)
r2_xgb = r2_score(y_test, y_pred_xgb)
print("Task 1 with XGBoost")
```

```
print("XGBoost MSE:", round(mse_xgb, 4))
print("XGBoost R2 Score:", round(r2_xgb, 4))
```

```
Task 1 with XGBoost
XGBoost MSE: 0.0872
XGBoost R2 Score: 0.9617
```

[21]:
```
import matplotlib.pyplot as plt
from xgboost import plot_importance
# Plot feature importance by gain
plot_importance(xgb_model, importance_type='gain', max_num_features=10)
plt.title("Top 10 Features Affecting Restaurant Ratings (Gain)")
plt.show()
```