mlr-project

April 4, 2024

0.1 To predict the sales using marketing and demographic factors

[1]:
```
#Importing Necessary Libraries
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
```

[12]:
```
# Loading the dataset
data = pd.read_csv(r'C:\Users\ntpc\Desktop\Carseats.csv')
```

[13]:
```
## Checking Datatypes for each columns
data.dtypes
```

[13]:
```
Unnamed: 0       int64
Sales          float64
CompPrice        int64
Income           int64
Advertising      int64
Population       int64
Price            int64
ShelveLoc       object
Age              int64
Education        int64
Urban           object
US              object
dtype: object
```

[14]:
```
# Encoding categorical variables
data = pd.get_dummies(data, columns=['ShelveLoc', 'Urban', 'US'],␣
 ↪drop_first=True)
```

[15]:
```
# Splitting the data into training and testing sets
X = data.drop('Sales', axis=1)
y = data['Sales']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=42)
```

[16]:
```
mlr_model = LinearRegression()
mlr_model.fit(X_train, y_train)
```

[16]:
```
LinearRegression()
```

[17]:
```
# Making predictions on the testing data
y_pred = mlr_model.predict(X_test)
```

[18]:
```
# Evaluating the model
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print(f'Mean Squared Error (MSE): {mse}')
print(f'Coefficient of Determination (R^2): {r2}')
```

```
Mean Squared Error (MSE): 1.1020554847858588
Coefficient of Determination (R^2): 0.8886385473088407
```