

Sales Prediction from Advertising Ads using Linear Regression

Published by WisdomML.in

Introduction

In this project, we are going to predict the sales of a product based on the money spent on advertising through different media such as TV, radio, and newspaper. We will use Linear Regression which is a supervised machine learning algorithm for regression problems.

Objective

The main goal of this project is to build a regression model using machine learning techniques to predict future sales based on advertising budget across various platforms.

Dataset Information

The dataset contains the following columns:

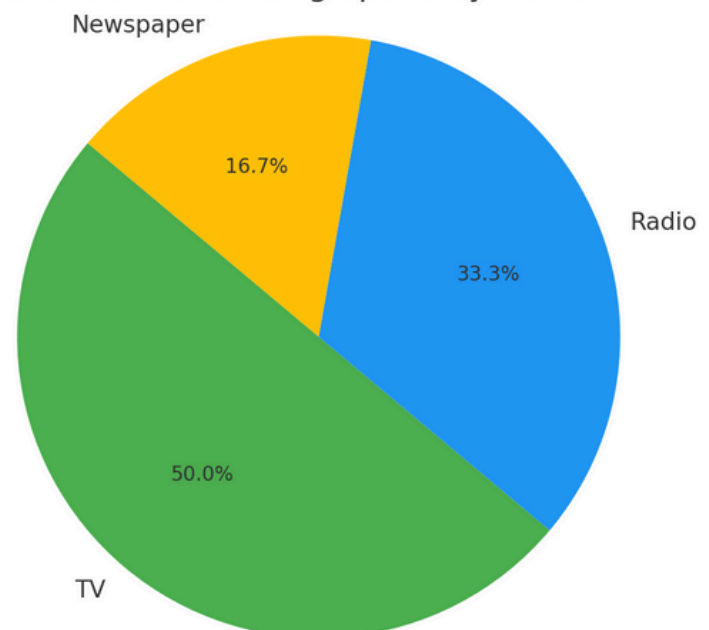
TV: Advertising dollars spent on TV for a single product in a given market (in thousands of dollars)

Radio: Advertising dollars spent on Radio

Newspaper: Advertising dollars spent on Newspaper

Sales: Sales of the product (in thousands of units)

Share of Advertising Spend by Platform



Steps Involved:

1. Import Required Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
```

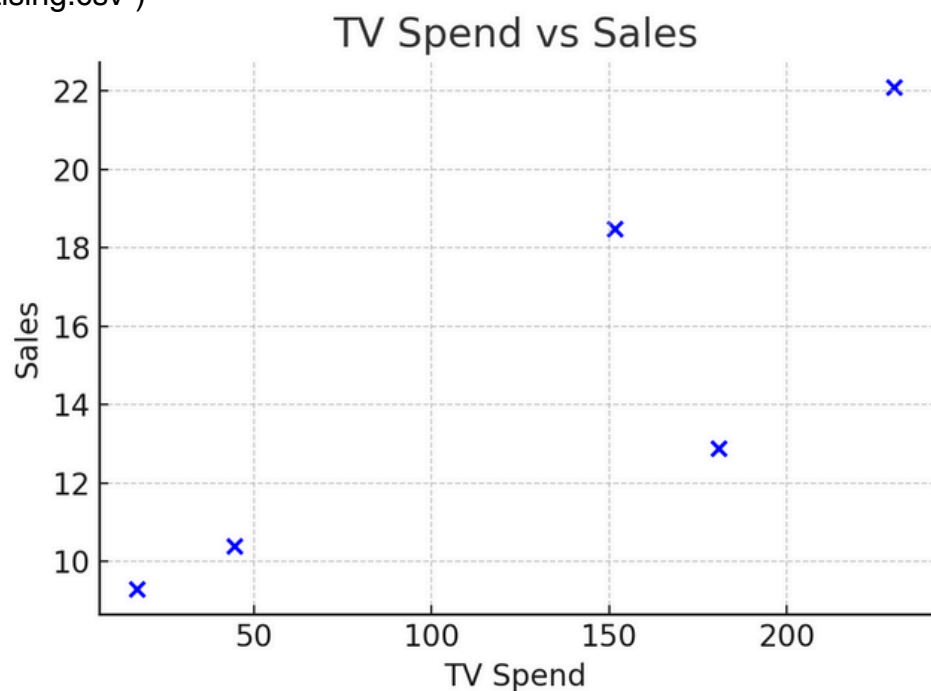
2. Load the Dataset

```
data = pd.read_csv("advertising.csv")
```

3. Data Exploration

```
print(data.head())
print(data.describe())
print(data.isnull().sum())
```

4. Visualize the Data



```
sns.pairplot(data, x_vars=['TV', 'Radio', 'Newspaper'], y_vars='Sales', kind='scatter')
plt.show()
```

5. Split the Dataset

```
X = data[['TV', 'Radio', 'Newspaper']]
```

```
y = data['Sales']
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

6. Train the Model

```
model = LinearRegression()
```

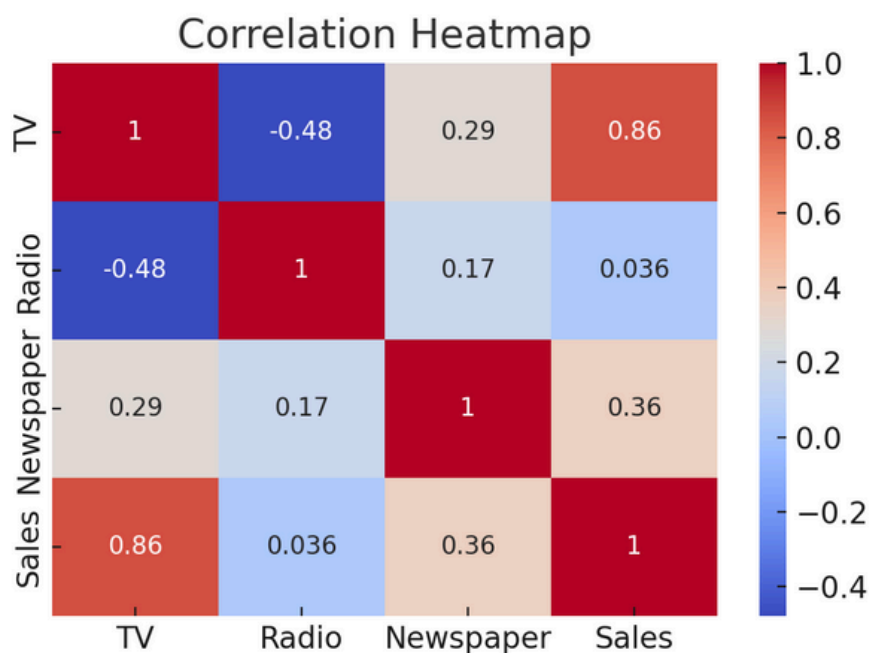
```
model.fit(X_train, y_train)
```

7. Evaluate the Model

```
y_pred = model.predict(X_test)
```

```
print("R-squared:", r2_score(y_test, y_pred))
```

```
print("MSE:", mean_squared_error(y_test, y_pred))
```



8. Plot Actual vs Predicted

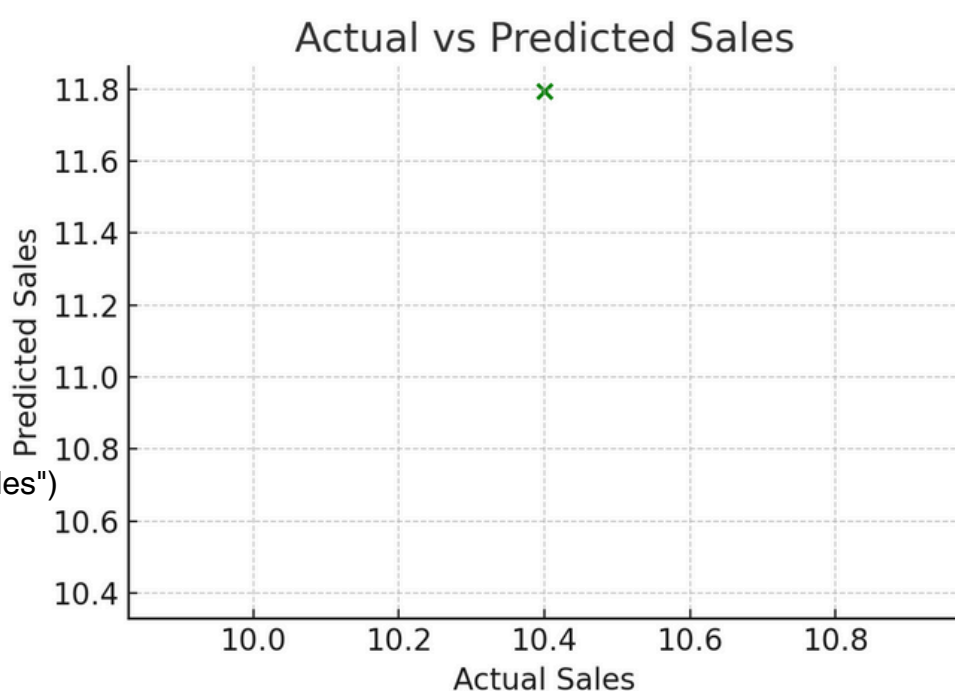
```
plt.scatter(y_test, y_pred)
```

```
plt.xlabel("Actual Sales")
```

```
plt.ylabel("Predicted Sales")
```

```
plt.title("Actual vs Predicted Sales")
```

```
plt.show()
```



Conclusion

We used a linear regression model to predict the sales based on advertising spends. The model performs well and shows that TV and Radio advertising have a greater impact on sales compared to Newspaper.