Sales Prediction with Machine Learning

Forecasting sales is a difficult problem for every type of business, but it helps determine where a business should spend more on advertising and where it should cut spending. In this article, I will walk you through the task of Sales Prediction with Machine Learning using Python.

What is Sales Prediction?

Sales prediction means predicting how much of a product people will buy based on factors such as the amount you spend to advertise your product, the segment of people you advertise for, or the platform you are advertising on about your product.

Typically, a product's sales are primarily dependent on how much you spend on advertising it, as the more people your product reaches, the more sales will increase as long as the quality of your product is good. Now in the section below, I will take you through a machine learning project on sales prediction using Python.

Sales Prediction using Python

Hope you now understand what sales forecasting is. Typically, a product and service-based business always need their Data Scientist to predict their future sales with every step they take to manipulate the cost of advertising their product. So let's start the task of sales prediction with machine learning using Python. I'll start this task by importing the necessary Python libraries and the dataset:

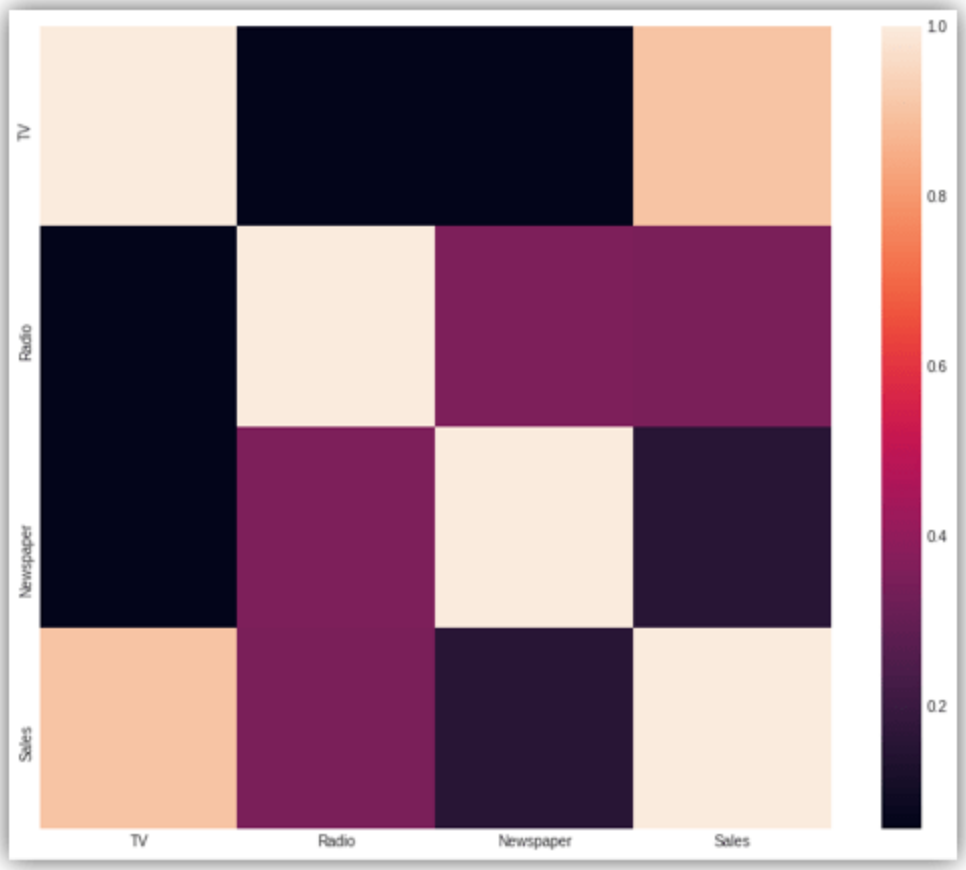
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression



data = pd.read_csv("https://raw.githubusercontent.com/amankharwal/Website-data/master/advertising.csv")
print(data.head())
```

```
   TV  Radio  Newspaper  Sales
0  230.1   37.8     69.2   22.1
1   44.5   39.3     45.1   10.4
2   17.2   45.9     69.3   12.0
3  151.5   41.3     58.5   16.5
4  180.8   10.8     58.4   17.9
```

```
print(data.isnull().sum())
```
```
TV          0
Radio       0
Newspaper   0
Sales       0
dtype: int64
```
So this dataset does not contain any null values. Now let's take a look at the correlation between features before we start training a machine learning model to predict future sales:

```
plt.style.use('seaborn-
whitegrid')
                        plt.figure(figsize=(12, 10))
                        sns.heatmap(data.corr())
                        plt.show()
```

```
plt.style.use('seaborn-
whitegrid')
                        plt.figure(figsize=(12, 10))
                        sns.heatmap(data.corr())
                        plt.show()
```



Now let's prepare the data to fit into a machine learning model and then I will use a ***linear regression*** algorithm to train a sales prediction model using Python:

```
x = np.array(data.drop(["Sales"], 1))

y = np.array(data["Sales"])

xtrain, xtest, ytrain, ytest = train_test_split(x, y, test_size=0.2, random_state=42)

model = LinearRegression()

model.fit(xtrain, ytrain)

ypred = model.predict(xtest)



data = pd.DataFrame(data={"Predicted Sales": ypred.flatten()})

print(data)
```

```
   Predicted Sales
0        17.034772
```

1	20.409740
2	23.723989
3	9.272785
4	21.682719
5	12.569402
6	21.081195
7	8.690350
8	17.237013
9	16.666575
10	8.923965
11	8.481734
12	18.207512
13	8.067507
14	12.645510
15	14.931628
16	8.128146
17	17.898766
18	11.008806
19	20.478328
20	20.806318
21	12.598833
22	10.905183
23	22.388548
24	9.417961
25	7.925067
26	20.839085
27	13.815209
28	10.770809
29	7.926825
30	15.959474
31	10.634909
32	20.802920
33	10.434342
34	21.578475
35	21.183645
36	12.128218
37	22.809533
38	12.609928
39	6.464413

Summary

So this is how we can predict future sales of a product with machine learning. Forecasting sales is a difficult problem for every type of business but using Machine Learning makes it look easy.