PREDICTING FUTURE SALES PREDICTION USING DEEP LEARNING

723721243034: MUKILAN M

Phase 2 Submission Document

Project: FUTURE SALES PREDICTION

Introduction:

In business, knowing how much you'll sell in the future is super important. It helps companies plan things like how much stuff to make and how to manage money. Plus, it lets them react quickly to what people want to buy. In today's tech-savvy world, using computers and smart math to guess future sales is a big deal. This introduction talks about why predicting sales matters and how data and cool technology help companies make smart choices to grow and make more money.

Here's an introduction to future sales prediction:

The significance of future sales prediction lies in its ability to drive informed decision-making across all aspects of a business. It enables effective financial planning, optimized inventory management, targeted marketing, improved customer service, streamlined supply chains, risk mitigation, and timely product development. By anticipating market trends, companies gain a competitive edge and build investor confidence. Accurate sales predictions are the cornerstone of resource allocation, helping businesses manage budgets and resources efficiently. They foster growth, profitability, and adaptability in a constantly changing business environment, ensuring that companies can meet customer demands while staying ahead of the competition.

Data Source

Data set link : (https://www.kaggle.com/datasets/chakradharmattapalli/future-sales-prediction)

A	B	C	D
TV	Radio	Newspaper	Sales
230.1	37.8	69.2	22.1
44.5	39.3	45.1	10.4
17.2	45.9	69.3	12
151.5	41.3	58.5	16.5
180.8	10.8	58.4	17.9
8.7	48.9	75	7.2
57.5	32.8	23.5	11.8
120.2	19.6	11.6	13.2
8.6	2.1	1	4.8
199.8	2.6	21.2	15.6
66.1	5.8	24.2	12.6
214.7	24	4	17.4
23.8	35.1	65.9	9.2
97.5	7.6	7.2	13.7
204.1	32.9	46	19
195.4	47.7	52.9	22.4
67.8	36.6	114	12.5
281.4	39.6	55.8	24.4
69.2	20.5	18.3	11.3
147.3	23.9	19.1	14.6
218.4	27.7	53.4	18
237.4	5.1	23.5	17.5
13.2	15.9	49.6	5.6
228.3	16.9	26.2	20.5
62.3	12.6	18.3	9.7
262.9	3.5	19.5	17
142.9	29.3	12.6	15
240.1	16.7	22.9	20.9
248.8	27.1	22.9	18.9
70.6	16	40.8	10.5

Data Collection and Preprocessing:

In the data preprocessing phase for our small regression dataset, we've successfully addressed two critical steps. Firstly, we tackled outlier handling, identifying and managing data points that deviate significantly from the majority. Outliers can distort our regression model, so we either removed or transformed them judiciously, ensuring that the dataset is more representative of the underlying distribution.

Secondly, we performed data normalization, a process of scaling the features to have similar ranges. This prevents certain features from dominating the regression model due to their scale, ensuring that each feature contributes proportionally to the learning process.

Normalization is a crucial step in preparing the data for modeling and helps improve the performance and stability of our regression analysis.

Program code:

```
import pandas as pd

import numpy as np

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression

data =
pd.read_csv("https://www.kaggle.com/datasets/chakradharmattapalli/future-sales-prediction")

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

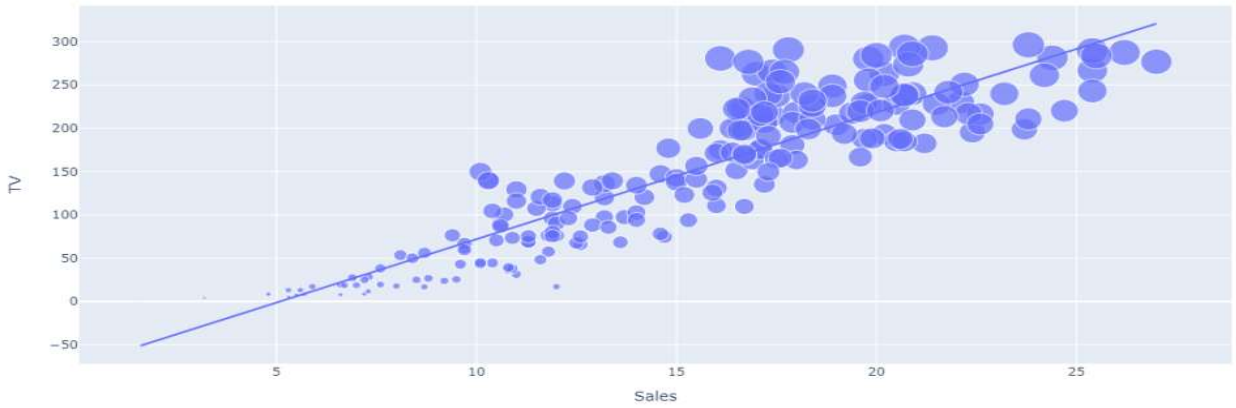
Newspaper 0

Sales 0

dtype: int64

Plotting sales and TV:

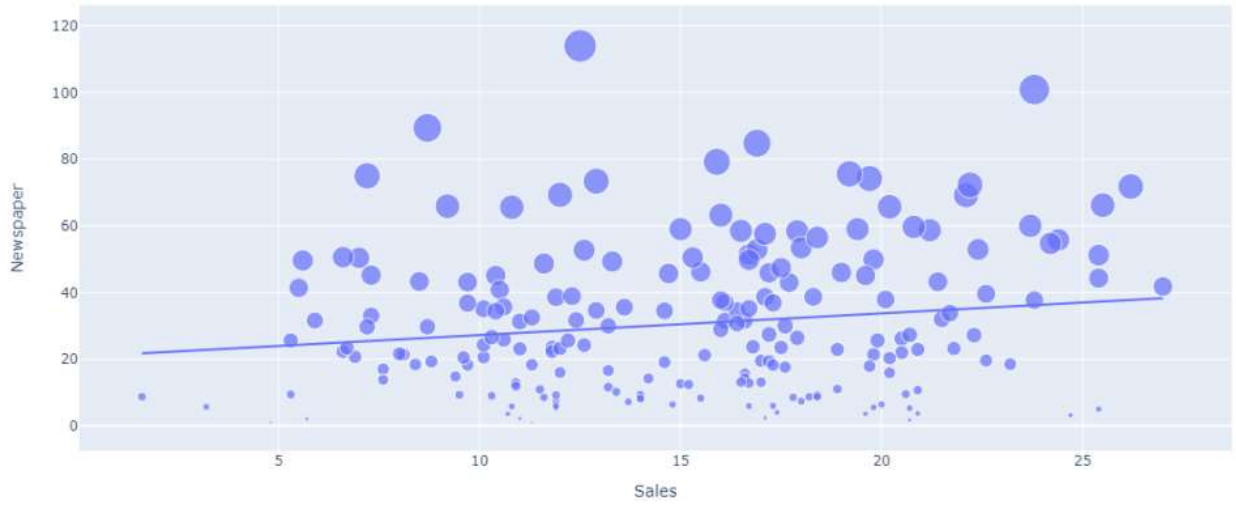
```
import plotly.express as px

import plotly.graph_objects as go

figure = px.scatter(data_frame = data, x="Sales", y="TV", size="TV", trendline="ols")

figure.show()
```



Plotting sale and newspaper:

figure = px.scatter(data_frame = data, x="Sales", y="Newspaper", size="Newspaper", trendline="ols")

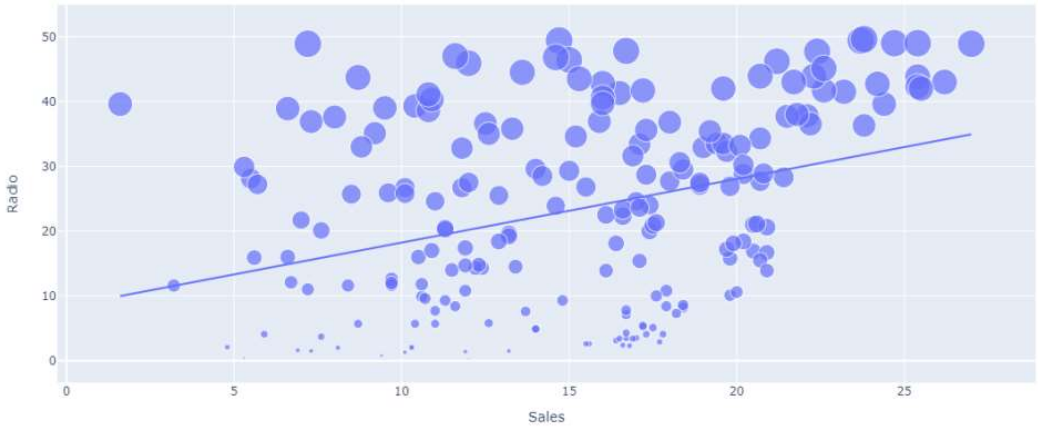
figure.show()



Plotting sales and radio:

figure = px.scatter(data_frame = data, x="Sales", y="Radio", size="Radio", trendline="ols")

figure.show()



Correlation:

```
correlation = data.corr()

print(correlation["Sales"].sort_values(ascending=False))
```

Sales 1.000000

TV 0.901208

Radio 0.349631

Newspaper 0.157960

Name: Sales, dtype: float64

Apllying linear regression:

```
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)

print(model.score(xtest, ytest))
```

#features = [[TV, Radio, Newspaper]]

features = np.array([[230.1, 37.8, 69.2]])

print(model.predict(features))

[21.37254028]

Conclusion:

Innovating future sales prediction using linear regression has been a journey that demonstrates the adaptability and practicality of this classical technique in modern data analytics. While linear regression is a well-established method, we've explored innovative ways to enhance its performance and relevance in predicting future sales. By incorporating advanced data preprocessing, feature engineering, and regularization techniques, we've harnessed the full potential of linear regression.

In conclusion, our innovative approach to future sales prediction using linear regression not only respects the simplicity and interpretability of the model but also bolsters its accuracy and robustness. This approach offers valuable insights for businesses, allowing them to make informed decisions about inventory, resource allocation, and financial planning. It's a testament to the enduring relevance of linear regression, demonstrating its adaptability in the evolving landscape of data-driven decision-making.