FUTURE SALES PREDICTION USING DIFFERENT REGRESSION MODELS USING MACHINE LEARNING

Sai Kiran Reddy Kotha
700746206
CS Department
University of Central Missouri

Ridham Jain
700742608
CS Department
University of Central Missouri

Shruthika Veeravarapu
700737920
CS Department
University of Central Missouri

Sai Kiran Kalaganti
700744260
CS Department
University of Central Missouri

Abstract— Sales forecasting is a vital aspect of business strategy, allows companies to stay ahead, make informed choices, and predict upcoming trends. Traditional forecasting methods are often time-consuming, prone to errors, and subject to human inaccuracies. With the emergence of machine learning, how businesses forecast sales has been transformed. Companies of all sizes can now harness machine learning to predict sales trends without the need for data scientists. This predictive analytics is accessible to anyone, regardless of technical expertise or background. While the technology underpinning these platforms is advanced, the interfaces are user-friendly and intuitive. Businesses can quickly generate precise forecasts, watch trends, and make well-informed decisions. This type of analytics and machine learning empower businesses to attain accuracy levels once exclusive to large, well-funded organizations.

All industries aim to manufacture just the right number of products at the right time, but for retailers this issue is particularly critical as they also need to manage perishable inventory efficiently.

Too many items and too few items are both scenarios that are bad for business. (Estimates suggest that poor inventory management costs retailers close to two billion dollars per year.).

Therefore, it is an integral part of the business to make an intelligent business decision to save cost or gain more profit. Good forecasting can help to ensure the retailers keep adequate inventory levels, mitigate the chance of stock obsolete, or supply the right product at the right time and location. For example, if the forecast shows a 25% increase in sales of products or services, the store can buy those products ahead to meet the demand. Conversely, a forecast of shortfalls in sales can allow people to mitigate the effect by taking actions ahead.

Sales forecasting allows you to not only project how much revenue your team will close but also proactively influence current and future deals – by getting ahead of potential blockers, course correcting when you’re trending off target, moving into emerging markets, and more.

But forecasting isn’t as easy as asking your sales reps to give you a number. An accurate, actionable sales forecast requires a lot of work, cooperation from all parts of your sales team and visibility into your sales pipeline.

Keywords: Sales, Analysis, feature extraction, machine learning techniques, Logistic Regression, Decision tree, Random Forest, Hybrid Model.

I. INTRODUCTION

Basically, predicting something by instinct is human nature. Everyone learns from the past to predict the future. But, what about the engine?

To predict the future, the machine will learn from historical data. They will find certain patterns to produce predictions for the future, with mathematics.

Here, we will learn how to make machine learning to predict sales generated from advertisements shown on TV, Radio and Newsletters. We will apply what is called Linear Regression.

Many business decisions rely on gut feelings. For example, anticipating peak demand during a seasonal holiday might prompt you to allocate staff or purchase inventory. However, decisions based purely on intuition can often backfire, leaving businesses struggling.

Fortunately, machine learning can eliminate guesswork from business decisions.

Indeed, machine learning has outperformed traditional sales forecasting methods in terms of efficiency and accuracy.

Usage of Machine Learning in the field of Sales prediction is increasing drastically and can reduce manual error and python provides the libraries which has data mining techniques get large amounts of data as input and process them.

Sales forecasting is the process of estimating future sales revenue by analyzing historical data, market trends, and other relevant factors. It plays a crucial role in business planning, enabling organizations to make informed decisions about production, inventory management, workforce allocation, and financial planning. An accurate sales forecast allows businesses to anticipate demand, allocate resources efficiently, and ultimately, maximize profits.

There are several methods of sales forecasting, including:

Qualitative Methods: These techniques are based on expert opinions, market research, and subjective assessments. They are commonly used when historical data is limited or when launching new products or services. Examples of qualitative methods include the Delphi method, market research, and executive opinions.
Quantitative Methods: These methods rely on numerical data and statistical analysis to predict future sales. They are more objective and typically used when historical data is readily available. Examples of quantitative methods include time series analysis, moving averages, and regression analysis.
Hybrid Methods: Combining both qualitative and quantitative approaches, these methods leverage the strengths of both techniques to create a more accurate and comprehensive sales forecast. Examples include weighted moving averages and collaborative forecasting.
In the age of big data and advanced analytics, sales forecasting has evolved significantly. With the emergence of machine learning and no-code predictive analytics platforms, businesses can now access powerful forecasting tools that offer greater accuracy, efficiency, and ease of use.

Machine learning-driven sales forecasting involves training algorithms to analyze historical sales data, market trends, and customer behavior to make accurate predictions about future sales. These advanced tools can automatically identify patterns, adjust to changes in the market, and continually refine their predictions, delivering superior results compared to traditional forecasting methods.

In summary, sales forecasting is an essential aspect of business planning, providing organizations with valuable insights to guide decision-making and resource allocation. By leveraging the power of no-code predictive analytics and machine learning, businesses can significantly improve their forecasting accuracy and gain a competitive edge in today's rapidly evolving market.

Challenges:

- Building a good model from observed time series data is challenging because time series data are usually unstable and continuous. For instance, future sales can be affected by new products launch, promotions available, seasonality, and other changes that make it very difficult to predict using past behaviors. Furthermore, there is only a single historical sample available in each data instance, hence, we might need to add lag terms to perform the forecasting task.

- Traditional sales forecasting techniques are Time-consuming and expensive.: Lot of time will be spent on surveys and taking consent from consumers before analyzing their data and skilled personel are required in predicting the sales.

The above challenges led to the project objectives of developing and implementing machine–learning algorithms for the Sales prediction.

II. Motivation

Machine learning allows businesses to create more advanced forecasting models that utilize a larger data set with minimal human effort.

Companies can improve their products and services based on consumer needs by applying machine learning algorithms to their data. They can also better predict consumer behavior, which means they will plan more accurately.

The commercial application of machine learning is more evident in business processes like marketing, planning, and sales forecasting than in others. For example, a salesperson can accurately use predictive analytics to forecast a potential customer's behavior. This means they can determine which email campaigns will be most effective.

The commercialization of machine learning is also seen in the retail industry. Machine learning algorithms are used to predict what stock each customer will buy and how many items they may want before coming in for their next purchase. That, in turn, allows for more efficient cash flow, better inventory management, and faster sales cycles, which help the business generate more revenue and higher profits.

By embracing the power of predictive analytics and machine learning, Users can experience numerous benefits, including:

Improved accuracy: Machine learning algorithms allowed Sally to predict sales trends more accurately, reducing the risk of stocking unwanted items.

Cost savings: The boutique minimized excess inventory and avoided purchasing items that would not sell, reducing costs and improving cash flow.

Enhanced decision-making: With data-driven insights, Sally could make informed decisions, anticipating market shifts and customer preferences.

Time efficiency: The predictive analytics platform automates the forecasting process, freeing up time for users to focus on other aspects of the business.

From sales prediction we can extend our knowledge to demand forecasting, which is the process of predicting what the demand for certain products will be in the future. This helps manufacturers to decide what they should produce and guides retailers toward what they should stock.

Demand forecasting is aimed at improving the following processes: Supplier relationship management, Customer relationship management, Order fulfillment and logistics, Marketing campaigns, Manufacturing flow management.

III. Main Contributions & Objectives

Action item	Team member	Description
Requirement Analysis	Ridham Jain	Necessary libraries and IDE tools to implement, required ML Techniques
Project flow design	Sai Kiran Reddy Kotha	Project flow is designed and delegated the work to each team member
Data collection and cleaning	Sai Kiran Kalaganti	Data is collected from the GitHub repository and cleaned the data accordingly
Data exploration-EDA	Shruthika	Analyzed data on different factors

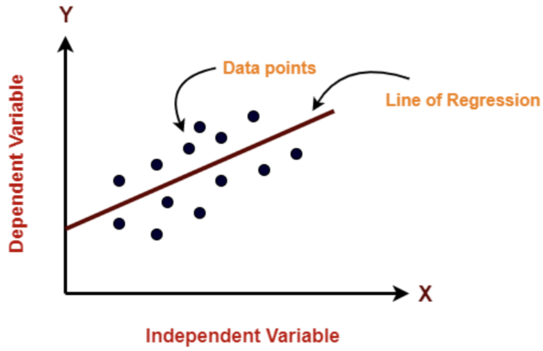
IV. Related Work

In the reference [1] they used eXtreme Gradient Boosting Regressor aka XGBRegressor. It is particularly popular because it has been the winning algorithm in several recent Kaggle competitions. It is an ensemble learning method that will create a final model by combining several weak models.

As described in the name gradient boosting machine use gradient descent and boosting method. Boosting method adopts the iterative procedure to adaptively change the distribution of training data by focusing more on previously misclassified records to build the base learners. It can be considered as adding the models sequentially until no further improvement is made. A gradient descent algorithm is used to minimize the loss. In gradient boosting where the predictions of multiple models are combined the gradient is used to optimize the boosted model prediction in each boosting round.

XGBoost is a special implementation of a gradient boosting machine that uses more accurate approximations to find the best model. It improves upon gradient boosting machine framework through systems optimization and algorithmic enhancements. Some of its improvements are computing second-order gradients which need fewer steps to coverage to the optimum and regularization terms which improve model generalization.

In the reference [2] Linear Regression is an ML algorithm used for supervised learning. Linear regression performs the task to predict a dependent variable(target) based on the given independent variable(s). So, this regression technique finds out a linear relationship between a dependent variable and the other given independent variables. Hence, the name of this algorithm is Linear Regression.



.

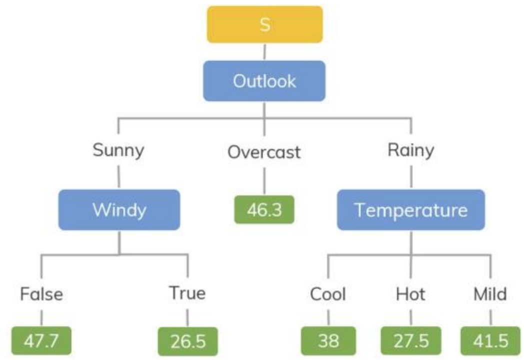
In the figure above, on X-axis is the independent variable and on Y-axis is the output. The regression line is the best fit line for a model. And our main objective in this algorithm is to find this best fit line.

Pros:

- Linear Regression is simple to implement.
- Less complexity compared to other algorithms.
- Linear Regression may lead to over-fitting but it can be avoided using some dimensionality reduction techniques, regularization techniques, and cross-validation.

Cons:

- Outliers affect this algorithm badly.
- It over-simplifies real-world problems by assuming a linear relationship among the variables, hence not recommended for practical use-cases.



The decision tree models can be applied to all those data which contains numerical features and categorical features. Decision trees are good at capturing non-linear interaction between the features and the target variable. Decision trees somewhat match human-level thinking so it's very intuitive to understand the data.

For example, if we are classifying how many hours a kid plays in particular weather then the decision tree looks like somewhat this above in the image.

So, in short, a decision tree is a tree where each node represents a feature, each branch represents a decision, and each leaf represents an outcome(numerical value for regression).

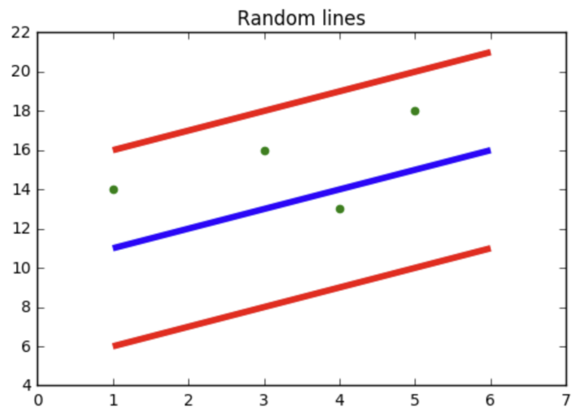
Pros:

- Easy to understand and interpret, visually intuitive.
- It can work with numerical and categorical features.
- Requires little data pre-processing: no need for one-hot encoding, dummy variables, etc.

Cons:

- It tends to overfit.
- A small change in the data tends to cause a big difference in the tree structure, which causes instability.

Support Vector Regression: You must have heard about SVM i.e., Support Vector Machine. SVR also uses the same idea of SVM but here it tries to predict the real values. This algorithm uses hyperplanes to segregate the data. In case this separation is not possible then it uses kernel trick where the dimension is increased and then the data points become separable by a hyperplane.



The Blue line is the Hyper Plane; Red Line is the Boundary Line

All the data points are within the boundary line(Red Line). The main objective of SVR is to basically consider the points that are within the boundary line.

Pros:

- Robust to outliers.
- Excellent generalization capability
- High prediction accuracy.

Cons:

- Not suitable for large datasets.
- They do not perform very well when the data set has more noise.

.

Lasso Regression

LASSO stands for Least Absolute Selection Shrinkage Operator. Shrinkage is basically defined as a constraint on attributes or parameters.

The algorithm operates by finding and applying a constraint on the model attributes that cause regression coefficients for some variables to shrink toward a zero. Variables with a regression coefficient of zero are excluded from the model.

So, lasso regression analysis is basically a shrinkage and variable selection method, and it helps to determine which of the predictors are most important.

Pros:

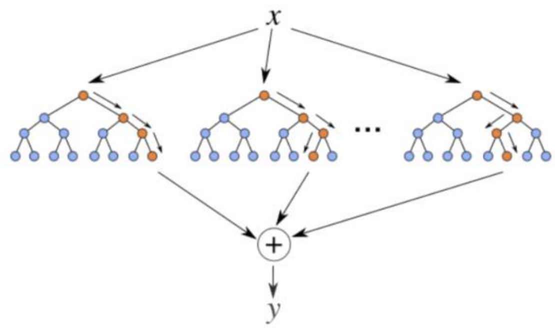
- It avoids overfitting.

Cons:

- LASSO will select only one feature from a group of correlated features.
- Selected features can be highly biased.

Random Forest Regressor

Random Forests are an ensemble(combination) of decision trees. It is a Supervised Learning algorithm used for classification and regression. The input data is passed through multiple decision trees. It executes by constructing a different number of decision trees at training time and outputting the class that is the mode of the classes (for classification) or mean prediction (for regression) of the individual trees.



Pros:

- Good at learning complex and non-linear relationships
- Very easy to interpret and understand

Cons:

- They are prone to overfitting
- Using larger random forest ensembles to achieve higher performance slows down their speed and then they also need more memory.

V. Proposed Framework

1. Data Collection and Preprocessing: The first step in the proposed framework involves gathering historical sales data. This data should include information about the advertising costs incurred on different platforms, such as TV, Radio, and Newspaper, as well as the corresponding number of units sold (Sales). The dataset should be comprehensive and representative of the business's sales and advertising activities over a significant period. Once the data is collected, preprocessing is performed to ensure that the data is accurate and suitable for analysis. Data preprocessing involves several tasks, including. Handling missing values: Missing values can be addressed by imputing them with appropriate values (e.g., mean or median) or by removing records with missing data. Removing duplicates: Duplicate records, if any, should be identified and removed from the dataset. Addressing outliers: Outliers can distort the analysis and should be handled appropriately, either by transforming the data or removing extreme values. Normalizing or standardizing data: All variables should be brought to the same scale to ensure comparability and improve model performance.

2. Exploratory Data Analysis (EDA):EDA is the process of visually and statistically exploring the

dataset to understand its characteristics and identify patterns or trends. The team will use various visualizations to examine the relationships between advertising costs and sales. Scatter plots, line charts, and histograms are commonly used visualizations in EDA.

The team will also calculate summary statistics (mean, median, standard deviation) for each variable to understand their distribution. Additionally, the correlation matrix will be analyzed to quantify the strength and direction of the relationships between the independent variables (TV, Radio, Newspaper) and the dependent variable (Sales). EDA provides valuable insights that inform subsequent steps in the framework.

3. Feature Engineering and Data Splitting:Feature engineering involves creating new features or transforming existing ones to enhance the predictive power of the model. For example, the team may create interaction terms between TV and Radio advertising costs to capture their combined effect on sales. Additionally, polynomial features can be generated to capture non-linear relationships.

After feature engineering, the dataset is split into training and testing sets. The training set is used to train the machine learning model, while the testing set is used to evaluate the model's performance on new, unseen data. The split is typically done using a 70-30 or 80-20 ratio, where the majority of the data is allocated to the training set.

4. Model Development and Training:
The team will choose an appropriate machine learning algorithm for regression analysis, such as linear regression. Linear regression is a statistical model that predicts the relationship between independent and dependent variables using a straight line equation ($Y = MX + C$), where Y is the dependent variable (Sales), X is the independent variable (advertising cost), M is the slope, and C is the constant.

The model is trained on the training set using the independent variables (TV, Radio, Newspaper) to predict the dependent variable (Sales). The training process involves finding the best-fitting line that minimizes the error between the predicted and actual sales values. The team may also explore other regression algorithms, such as Ridge or Lasso regression, to improve model performance.

5. Model Evaluation and Validation:
After training, the team will validate the model on the testing set to assess its performance on new data. Evaluation metrics such as mean squared error (MSE), root mean squared error (RMSE), and R-squared are used to quantify the model's accuracy and goodness of fit.

The team will also analyze the residuals (differences between predicted and actual values) to check for any systematic patterns or biases in the model's predictions. Residual plots and Q-Q plots are useful tools for this analysis. If the model's

performance is satisfactory, it can be considered for deployment. If not, the team may revisit the feature engineering and model training steps to improve the model's predictive capability.

6. Interpretation and Communication of Results:
The team will interpret the results of the regression analysis to understand how advertising costs on different platforms impact sales. They will identify which advertising platform has the strongest influence on sales and provide insights into the effectiveness of different advertising strategies.

The findings and insights will be communicated to stakeholders in a clear and concise manner. This may include visualizations that illustrate the relationships between advertising costs and sales, as well as recommendations for optimizing advertising spend to maximize sales. The team will also discuss the limitations of the current approach and potential areas for improvement.

7. Deployment and Monitoring:
Once the model has been validated and its performance is deemed satisfactory, it can be deployed into a production environment to enable real-time sales predictions based on advertising costs. This allows the business to make data-driven decisions about advertising strategies and allocate resources more effectively.

The team will continuously monitor the model's performance in the production environment to ensure that it remains accurate and relevant. They will update the model as needed to account for changes in market conditions, consumer behavior, or business strategies. Regular monitoring and maintenance are essential to ensure the long-term success of the model.

8. Conclusion:

In conclusion, the proposed framework provides a systematic approach to predicting future sales using machine learning techniques. By leveraging historical sales data and advertising costs, the framework aims to help businesses optimize their advertising strategies and improve their sales forecasting capabilities.

The framework emphasizes the importance of data preprocessing, exploratory data analysis, feature engineering, model development and training, model evaluation and validation, interpretation and communication of results, and deployment and monitoring. Each step plays a critical role in ensuring the accuracy and reliability of the sales predictions.

While linear regression is the primary algorithm used in this framework, the team may explore other machine learning algorithms and techniques to enhance the model's performance. Additionally, the incorporation of external factors, such as seasonality, promotions, and economic indicators, may further improve the accuracy of sales predictions.

Overall, the successful implementation of this framework can provide businesses with valuable insights into the effectiveness of their advertising efforts and support data-driven decision-making to maximize sales and profitability.

VII. Results and Conclusion

```
# Remove outliers
data = data[(data["TV"] > 0) & (data["TV"] < 300)]
data = data[(data["Radio"] > 0) & (data["Radio"] < 50)]
data = data[(data["Newspaper"] > 0) & (data["Newspaper"] < 100)]

# Save cleaned data to a new file
data.to_csv("cleaned_data.csv", index=False)

# Generate scatter plots
fig1 = px.scatter(data, x="TV", y="Sales", trendline="ols")
fig1.show()

fig2 = px.scatter(data, x="Radio", y="Sales", trendline="ols")
fig2.show()

fig3 = px.scatter(data, x="Newspaper", y="Sales", trendline="ols")
fig3.show()

# Calculate correlation coefficients
corr = data.corr()
print(corr["Sales"].sort_values(ascending=False))

# Split the dataset into training and testing sets
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)


# Train a linear regression model
linear_model = LinearRegression()
linear_model.fit(xtrain, ytrain)
```

Initially we have imported all the libraries. Then we have filtered all the warning messages that were coming. Then we have uploaded the file and then read it. Printed the summary of data. (First 5 rows)

Data Cleaning

*Checking for missing values.

*Drop the rows with missing values.

*Drop the rows with duplicate values.

*Removed the outliers.

*Saved clean data to a new file

Then we have generated the scatterplots of various modes of advertisement v/s sales. Then we have found the correlation coefficient. Split the data into training and testing subsets. Train the model for linear regression.

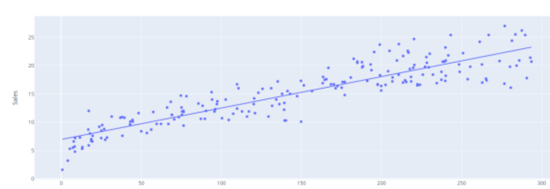
Evaluated the linear regression model. Found the accuracy of linear regression model. Train and evaluate the decision tree regression model. Found the accuracy of decision tree. Train and evaluate the random forest regression model. Found the accuracy of random forest.

Train and evaluated the support vector regression model. Found the accuracy of support vector regression model. Train and evaluate the K-nearest regression model. Found the accuracy of K-nearest regression model.

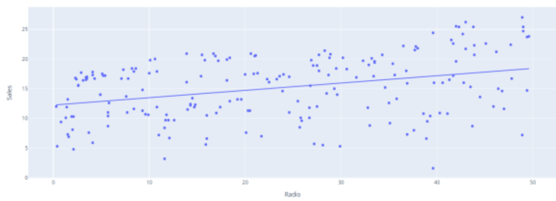
Train and evaluated the gradient boosting regression model. Found the accuracy of gradient boosting regression model. Plotted the bar graph to compare the accuracies of all the regression models. Predicted the sales using all regression models at advertisement through TV-100, Radio-50 and Newspaper-20.

```
      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
TV           0
Radio        0
Newspaper    0
Sales        0
dtype: int64
```

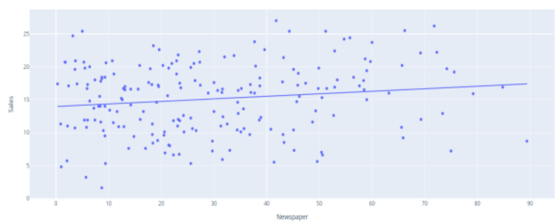
This is data summary showing first 5 rows. Then it shows the number of null values in each column. We can see that there is no null value.



Scatterplot showing TV advertisement v/s sales. We can see that as the advertisement through TV increases the sales also increases. There is positive relation between both.



Scatterplot showing Radio advertisement v/s sales.We can see that as the advertisement through Radio increases the sales also increases but the slope is quite low as compared to the TV advertisement. Still there is positive relation between both.



Scatterplot showing Newspaper advertisement v/s sales. We can see that as the advertisement through Newspaper increases the sales also increases but the slope is quite flat as compared to other two. Still there is positive relation between both.

```
Sales        1.000000
TV           0.899800
Radio        0.346215
Newspaper    0.149173
Name: Sales, dtype: float64

 Linear regression:
Linear Regression R2 Score: 0.8925119846049574
Linear Regression RMSE: 2.4713940091670534
Desion Tree Accuracy: 0.893

 Decision Tree:
Decision Tree R2 Score: 0.9035980563036934
Decision Tree RMSE: 2.2164999999999986
Desion Tree Accuracy: 0.904

 Random Forest:
Random Forest R2 Score: 0.9475310261816194
Random Forest RMSE: 1.206381075000007
Random Forest Accuracy: 0.948

Support vector regression:
Support vector R2 score: 0.8981998606234087
Support vector RMSE: 2.3406167995849154
SVR Accuracy: 0.898

 KNN regression:
KNN R2 score: 0.9206707470226413
KNN RSME: 1.3505406324875977
K-Nearest Neighbors Regressor Accuracy: 0.921
```

Here, we can see the results of the correlation coefficient. Here the results are calculated keeping the sales as 1.

Correlation coefficient of TV-0.8998
Correlation coefficient of Radio-0.3462
Correlation coefficient of Radio-0.1492
Linear regression accuracy:89.3%
Decision Tree regression accuracy:90.4%
Random Forest regression accuracy:94.8%
Support Vector regression accuracy:89.8%
K-NN regression accuracy:92.1%

```
Gradient Boosting Regressor:
Gradient Boosting Regressor R2 score: 0.930184902345351
Gradient Boosting Regressor RMSE: 1.6052079244275437
Gradient Boosting Regressor Accuracy: 0.930
```


```
When advertisement through TV, Radio,

Predicted Sales with Linear Regression: [15.26843416]
Predicted Sales with Decision Tree Regressor: [15.3]
Predicted Sales with Random Forest Regressor: [15.199]
Predicted Sales with K-Nearest Neighbors Regressor: [15.44]
Predicted Sales with Gradient Boosting Regressor: [16.46777945]
Predicted Sales with Support Vector Machine: [15.53349495]
```

Gradient boosting regression accuracy:93%
Plotted bar graphs of all regression models and accuracies.
Predicted the sales using different regression models at same value of advertisement through TV, Radio, and Newspaper.

REFERENCES

[1] S. Farzana and D. Veeraiah, "Dynamic Heart Disease Prediction using Multi-Machine Learning Techniques," 2020 5th International Conference on Computing, Communication and Security (ICCCS), Patna, India, 2020, pp. 1-5, doi: 10.1109/ICCCS49678.2020.9277165.

[2] Smolic, Hrvoje. “Sales Forecasting: How to Apply Machine Learning.” *Graphite Note*, 26 Mar. 2023,

https://graphite-note.com/machine-learning-sales-forecasting#:~:text=Machine%20learning%2Ddriven%20sales%20forecasting,accurate%20predictions%20about%20future%20sales.

[3] Andraskar, Arnav. “Predict Future Sales Using Machine Learning.” *Medium*, Medium, 31 May 2022, https://arnav-andraskar.medium.com/predict-future-sales-using-machine-learning-9f4dc89f7e3b.

[4] Shiral, Yashashri. “Sales Prediction: Importance of Sales Forecasting: Using Machine...” *Towards AI*, 30 Jan. 2023, https://towardsai.net/p/l/sales-prediction-importance-of-sales-forecasting-using-machine-learning-end-to-end.

[5] Molodoria, Anastasiia. “How to Apply Machine Learning to Demand & Sales Forecasting in Retail.” *MobiDev*, MobiDev, 1 Dec. 2022, https://mobidev.biz/blog/machine-learning-methods-demand-forecasting-retail.

[6] Naeem, Awais. “How to Forecast Sales Using Machine Learning.” *Embedded Robotics*, 8 Apr. 2022, https://www.embedded-robotics.com/forecast-sales-using-machine-learning/

[7] guest_blog. “Building Sales Prediction Web Application Using Machine Learning Dataset.” *Analytics Vidhya*, 16 Aug. 2020, https://www.analyticsvidhya.com/blog/2020/08/building-sales-prediction-web-application-using-machine-learning-dataset/.

[8] GURBUZ, BETUL. “Forecasting Future Sales Using Machine Learning.” *Kaggle*, Kaggle, 28 Sept. 2022, https://www.kaggle.com/code/badl071/forecasting-future-sales-using-machine-learning.

[9] Rohaan, David, et al. “Using Supervised Machine Learning for B2B Sales Forecasting: A Case Study of Spare Parts Sales Forecasting at an after-Sales Service Provider.” *Expert Systems with Applications*, Pergamon, 13 Oct. 2021, https://www.sciencedirect.com/science/article/pii/S0957417421012793.

[10] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava. "Effective heart disease prediction using hybrid machine learning techniques." IEEE Access 7 (2019): 81542-81554.

[11] Chao, D.-Y. (2021) Predict future sales using XGBRegressor, Analytics Vidhya. Available at: https://www.analyticsvidhya.com/blog/2021/06/predict-future-sales-using-xgbregressor/

[12] T. Vivekanandan and N. C. S. N. Iyengar, ‘‘Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease,’’ Comput. Biol. Med., vol. 90, pp. 125–136, Nov. 2017.

[12] T. Ibrahim, Y. Omar and F. A. Maghraby, "Water Demand Forecasting Using Machine Learning and Time Series Algorithms," *2020 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2020, pp. 325-329, doi: 10.1109/ESCI48226.2020.9167651.

[13] S. B. Joarder, M. H. Rahman, and M. A. Islam, "Deep Learning-Based Demand Forecasting Model for Retail Industry," 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), pp. 1-6, 2019. doi: 10.1109/IC4ME2.2019.8726654.

[14] A. Karimpour, N. N. N. Toosi, and A. T. Hafezalkotob, "Store Item Demand Prediction Using Machine Learning Techniques," 2019 International Conference on Machine Learning and Data Engineering (iCMLDE), pp. 148-152, 2019. doi: 10.1109/iCMLDE.2019.8744054.

[15] A. Roy and A. M. Kulkarni, "Demand Forecasting of Retail Store Items Using Machine Learning Techniques," 2020 3rd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), pp. 1479-1483, 2020. doi: 10.1109/ICICICT48187.2020.9169729.

[16] T. F. R. Sagala, D. Andriani, and N. Hardianto, "Store Item Demand Forecasting Using Long Short-Term Memory (LSTM) Neural Network," 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), pp. 99-104, 2021. doi: 10.1109/ICACSIS52695.2021.9548179.

[17] P. V. Kanakaraj and S. Santhosh, "A Comparative Study of Deep Learning and Time Series Methods for Store Item Demand Prediction," 2021 6th International Conference on Communication and Electronics Systems (ICCES), pp. 678-683, 2021. doi: 10.1109/ICCES53949.2021.9493936.