

Elaboration and Implementation of an Online Marketing Sales Prediction Model Based on Machine Learning Techniques

V RajeshkumarPitani¹, Dr. Harsh Lohiya²

¹ Research Scholar, Dept. of Computer Science, Sri Satya Sai University of Technology and Medical Sciences, Sehore Bhopal-Indore Road, Madhya Pradesh, India.

² Research Guide, Dept. of Computer Science, Sri Satya Sai University of Technology and Medical Sciences, Sehore Bhopal-Indore Road, Madhya Pradesh, India

ABSTRACT

Most business management systems rely significantly on an information system and request sales prediction. What happens with the sales estimate has a big impact on the firm. A data mining method or algorithm is a very effective way to extract unknown information from a training dataset in order to increase the accuracy and efficiency of sales forecasting. This review paper provides a comprehensive research and analysis of understandable predictive models that may be used to provide improved future sales predictions. Because of the large amount of data and the accuracy of sales forecasting, the most recent forecasting techniques are quite simple to utilize. These problems can be tackled utilizing a variety of machine learning techniques. The linear regression model provides the most accurate forecasts and predictions of future sales. Sales predictions are acceptable sales targets established by various analysis techniques after taking into account all of the aspects that may effect future sales. These elements include external influences such as economic shifts and demand trends, as well as internal factors such as marketing techniques and the product itself.

INTRODUCTION

Machine Learning (ML) is an effective method for forecasting sales. The sales forecast is mostly used to determine how much money is required to complete new plans and projects within budget. By improving sales rates, machine learning may be used to improve system advantages and ensure profitability. It is one of the most effective methods for assembling a complete set of data in order to eliminate onerous duties in the business. According to the research, sales forecasting is a modern, effective, and modern form of business intelligence. A detective's role is to locate lost data, gaps in data, and difficult challenges so that a business can increase its

profitability. ML approaches employ several models, such as GARCH, SARIMA, and others, to introduce various algorithms that aid in the accuracy of business plans.

ML is important since it helps to eliminate the prediction that sales forecasts will be inaccurate. This sales prediction rate is important for keeping the supply chain management system in excellent working order. In order to retain management talent in an organization, there has been an increase in ML during the last few decades. Using AI and computer algorithms aids in the design of various programs that can run independently within the company. Python and Spyder were discovered to be two creative programs that have been utilized to anticipate sales in the prediction. To complete the system, the ML system assists with feature engineering, judgment, data exploration, and model creation. Algorithms of several types exist, including Random Forest, Design Tree, linear regression, and ridge regression. All of these good algorithms will be used to improve sales prediction.

The ML techniques are vital for determining whether or not a company is a solid idea based on future sales. Furthermore, this system is required to develop future effective sales management techniques. Sales forecasting requires the utilization of both financial and human resources in the management field. An firm can use ML to build plans to keep up with sales and demand. It will also be effective to develop a market strategy to assist the firm in making sound judgments and staying on track. ML can increase an organization's profitability by 15%, which is a solid approach to improve business outcomes. It will also be effective to establish professional relationships with customers and suppliers in order to boost management performance. Machine Learning is also vital for creating marketing campaigns that will assist the organization in increasing sales. ML may also forecast how many services and goods a business will require to earn money. To be lucrative, a business must make accurate predictions, evaluate data, and establish robust processes.

A. Challenges in sales of products

E-commerce is the practice of conducting business through the internet and purchasing goods from anywhere in the globe. People are increasingly purchasing what they require on a daily basis online, particularly around COVID-19. Many firms attempt to sell their items through e-commerce websites with the assistance of business management specialists and IT solutions.

There are numerous IT firms that work on e-commerce technologies to sell their clients in selling their products and developing e-commerce websites. Online buyers can now tailor the goods they want thanks to virtual technologies. They can, for example, virtually try on clothes, design the interior of a place, and arrange tickets utilizing virtual tours. These qualities make customers happier and more excited when they buy products online. When creating an e-commerce website, the designer must consider what features should be included, how to make the website website-friendly, how to display the products, and so on. The main issue with e-commerce is that all of the things are stylish, up-to-date, and priced lower than they would be in person. There are large groups of people whose sole purpose is to sell and buy products. They devise new strategies to implement, and all of them are created, tested, and released to the public.

B. Machine learning algorithms

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that can learn from data and test its findings. ML attempts to educate the computer to learn so that it may answer complex problems without the assistance of a knowledge. There are four major categories in Machine Learning.

- “Supervised learning
- Unsupervised learning
- Semi-Supervised learning
- Reinforcement Learning”

Each of these machine learning methods can be used to create a wide range of algorithms that solve difficult issues. ML Algorithms will assist the client in predicting what will sell in the coming month. This allows the client to stock up on products and plan ahead of time how to sell them. There are numerous types of machine learning algorithms, each of which falls into one of the following categories:-

- “Regression algorithms
- Decision tree algorithms
- Bayesian algorithms

- Clustering algorithms
- Artificial neural network algorithms
- Deep learning algorithms
- Predictive algorithms”

C. AI for sales

AI (artificial intelligence) techniques will assist the firm in running smoothly and communicating with customers and clients in an intelligent manner. Businesses were able to produce more money and reduce costs where it made sense by replacing them with AI as a result of AI. By obtaining data, analysts can examine it, develop marketing strategies, and make conclusions.

AI techniques are used to forecast future sales. Each month, business owners can view and analyze their product sales. They could also predict how sales would progress in the future. Because of this capability, marketing teams can determine what went wrong when sales decline and devise new marketing strategies to increase sales. By analyzing data, AI will be able to generate forecasts and product recommendations. There are numerous advanced AI techniques that can be used to increase sales and assist a firm develop. Artificial intelligence can be applied in both business-to-business and business-to-consumer engagements. AI research has been ongoing for a long time. There is no need to employ software that requires people to manually enter data. Instead, AI can greatly assist and improve the process.

D. Need for prediction

We may make decisions using projections, which is a vital part of managing a business. The prediction's accuracy must be high in order to make sound decisions. If the accuracy is low, it may harm the business when implemented. There are numerous algorithms for creating predictions in Machine Learning-

- “Random forest
- Generalized linear model(GLM) for two values
- Gradient linear model(GBM)
- K-Means

- Prophet”

E-commerce websites that sell items can use prediction to guess which ones will sell well. Many factors must be considered while making a prediction, including the weather, temperature, holidays, festivals, discounts, and so on. As a prediction, while forecasting sales, all of these factors must be considered. At the moment, all e-commerce sites make projections regarding future sales and what will happen.

LITERATURE REVIEW

Bhuvanewari, A., & Venetia, T. A. (2021). Online purchasing has surged in popularity, yet vendors have few options for increasing sales. A machine learning algorithm can be used to identify what should be sold in a given month in order to increase sales. When the forecast is finished, a dashboard will be produced to display which products should have been sold to achieve high sales. With the assistance of an expert, the sales are billed and examined. However, not everyone has the financial capacity to hire professionals in this scenario. Vendors rely on previous experience. People who have been running businesses for a few years need help because they lack experience. Each month, a sales prediction is made to display vendors in growing their businesses, and a dashboard displays the items that will be sold in that month for an offer. To make predictions, machine learning algorithms employ the Random Forest Algorithm. This decision tree-based technique is the most effective way to forecast. The purpose of this project is to develop a random forest model to forecast how many products will be sold each month between January 2013 and October 2015.

Patangia, M. S. et al A et al (2020)A retailer must understand what their customers desire and provide those needs. Market basket analysis is one method for determining which items can be combined. Market basket analyses provide businesses with a wealth of information on sales related to a specific category of goods. When customers buy bread, they frequently buy other products to go with it, such as milk, butter, or jam. It makes sense to place these groups near to other in a shopping mall so that customers may easily access them. Groups of similar goods must

also be placed adjacent to each other to remind customers of related things and direct them logically around the center. Market basket analysis is used to determine which products are purchased together, how the supermarket should be set up, and how to create advertising campaigns. As a result, it is critical to examine how Market customers behave, which can be accomplished via the use of various data mining approaches. Informed decisions may be made regarding where to place a product, how much to charge for it, who to promote it, how profitable it is, and whether or not there are any successful products that have little in common. You can find comparable products and place them next to each other or try to sell them together.

Cheriyam, S. et al (2018)Data mining techniques are extremely effective at converting massive volumes of data into meaningful information that can be used to forecast costs and sales. This is the foundation of a good budget. Sales predictions are key components of many corporate decision-making processes, including operations, marketing, sales, production, and finance. The study continues with a new angle that focuses on how to choose the optimal strategy for accurately anticipating sales. The first dataset analyzed for this research had a high number of entries, however the final dataset utilized for analysis had much less entries than the first due to the removal of useless data, duplicate entries, and sales data. The data-tuning process and forecasts are illustrated with images of the outcomes.

Bajaj, P., et al (2020)Forecasting sales has always been a significant project to focus on. To sustain marketing organizations running well, all vendors must have a strong way to predict forecasting. Manually performing this operation may result in major errors that would harm the organization's management and, more critically, would take a long time, which is not desirable in today's fast-paced environment. A significant portion of the global economy is dependent on business sectors, which are expected to provide enough items to suit the demands of the entire world.

Chavan, S., et al (2020)A large element of purchasing management is predicting how well a product will sell. One of the most difficult challenges that organizations confront today is the need to operate in a fast-paced, globalized, and unpredictable commercial environment.

Customers are becoming more fussy about price and quality, so producers can no longer simply compete on pricing. Forecasting sales is critical for determining how much inventory to keep on hand, and accurately predicting future demand for items has long been a challenge for enterprises. If things are difficult to obtain or if there are more goods than people demand, the overall profit may suffer. As a result, forecasting how well items will sell can help to keep losses to a minimal. Based on this research, our project anticipates the use of machine learning algorithms to develop a prediction model capable of accurately predicting online product sales. Our project's purpose is to anticipate how many things will sell using current data such as online reviews, ratings, promotional techniques and feelings, and other aspects.

PROPOSED SYSTEM

We will examine how a company's sales have evolved over time and generate data on the company's sales for the upcoming quarter and for a specific product. We will utilize linear regression and logistic regression for this type of sales prediction project and judge the results based on training, testing, and validation data sets.

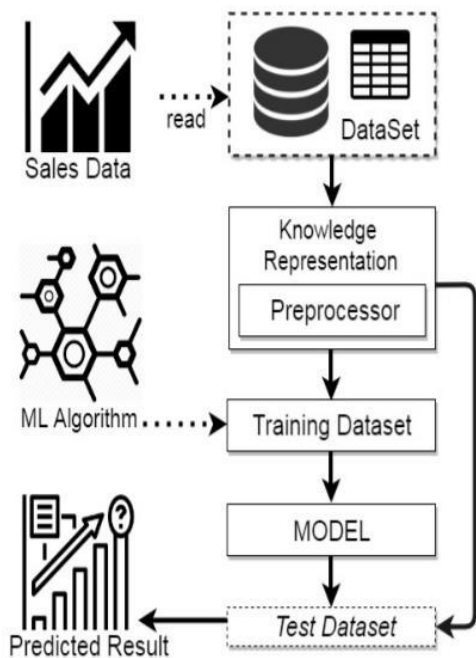


Fig 1 system architecture

The study of how to utilize computers and statistics to automate the process of learning from knowledge is known as machine learning. Our project involved creating sales estimates, and some of them were examined. We've discussed many machine learning algorithms that can be utilized in various aspects of the business, such as retailing, logistics, marketing, and so on, depending on the situation. The study provides important knowledge regarding machine learning techniques.

Linear Regression

The most popular and widely used Machine Learning algorithm is linear regression. It is used to demonstrate that the target or dependent variable is directly related to the response or independent variables. The linear regression model is based on the following equation:

$$\hat{y} = \theta_0 + \theta_1x_1 + \theta_2x_2 + \theta_3x_3 + \dots + \theta_nx_n$$

“where, y is the target variable, θ_0 is the intercept, $x_1, x_2, x_3, \dots, x_n$ are independent variables and $\theta_1, \theta_2, \theta_3, \dots, \theta_n$ are their respective coefficients”.

“The main aim of this algorithm is to find the best fit line to the target variable and the independent variables of the data. It is achieved by finding the most optimal values for all θ . With best fit it is meant that the predicted value should be very close to the actual values and have minimum error.

Error is the distance between the data points to the fitted regression line and generally can be calculated by using the following equation”:

$$\text{Error} = y - \hat{y}$$

where, y is the actual value and \hat{y} is the predicted value”.

K-Neighbors Regressor

The KNN algorithm for regression employs a supervised learning mechanism. It predicts the target based on how similar it is to previously seen examples. Distance is used to determine how similar two items are, and the most frequent method is Euclidian distance.

The K most similar instances, or neighbors, of the testing point in the entire dataset are used to make predictions. The KNN algorithm uses the Euclidean distance formula to calculate the distance between the mathematical values of these points:

$$\sqrt{\sum_{i=1}^n (y_i - x_i)^2}$$

The value of K chosen should not be too tiny, since this may produce noise in the data and, as a turn, overfitting. The typical answer is to set aside some data to test the model's accuracy. Then, set K to 1, perform the training part of modeling, and use all of the test set samples to determine how accurate the prediction is. Rep this procedure while increasing K, and select K so that the model functions best with it.

Comparative analysis of Performance Metrics

The average was derived in this section based on the ten iterations that were considered for each metric.

Average Accuracy Score

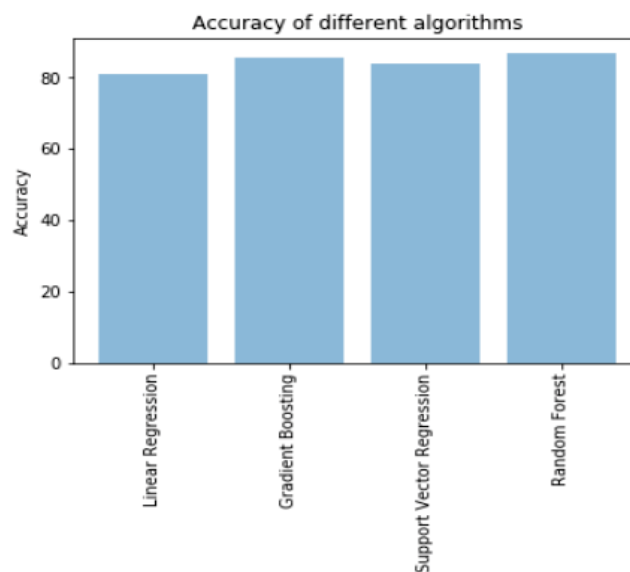


Figure 2: Average Accuracy score plot

“Figure 2 shows the average accuracy score of the 10-fold stratified cross-validation obtained by the Simple Linear Regressor is 81.2 percent, followed by the Gradient Boosting Regressor with 86.27 percent accuracy score, then SVR with the 84.82 percent accuracy score and finally Random Forest Regressor with 87.72 percent accuracy score. From figure 2, it can be seen that Random Forest Regressor is the best performer with approximately 88 percent accuracy score compared to other methods, and Simple Linear Regressor is the poor performer with an accuracy score of 81.2 percent”.

Average Mean Absolute Error

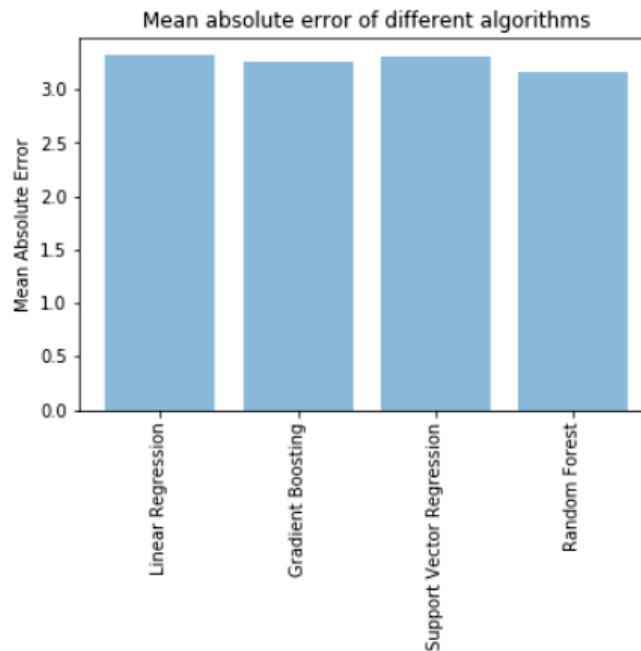


Figure 3: Average Mean Absolute Error plot

“Figure 3 shows the average MAE of the 10-fold stratified cross-validation obtained by Simple Linear Regressor is 3.21 error, followed by Gradient Boosting Regressor with 3.19 error, then SVR with 3.21 error, and finally Random Forest Regressor with 3.15 error respectively. From figure 3, it can be shown that Random Forest Regressor is the best performer with less error relative to other approaches and with the highest error, Simple Linear Regressor is the poor performer”.

CONCLUSION

To avoid product shortages, each client and each company must know what they want ahead of demand. Businesses will demand increasingly accurate predictions as time goes forward. This demand is expected to expand at an exponential rate. As a result, much research is being undertaken in this area in order to make accurate sales predictions. Better predictions are made in terms of how much money a firm makes. In this article, a two-level statistical model that decreases the mean total error value was used to attempt to accurately predict sales of a product made on a certain passage. For the Big Mart dataset, the two-level statistical model outperformed the single-model predictive approach and made more accurate predictions.

REFERENCES

Bhuvanewari, A., & Venetia, T. A. (2021). Predicting periodical sales of products using a machine learning algorithm. *International Journal of Nonlinear Analysis and Applications*, 12(Special Issue), 1611-1630.

Patangia, M. S., Mohite, M. R., Shah, M. K., Kolhe, M. G., Mokashi, M. M., & Rokade, M. P. Sales Prediction of Market using Machine Learning. *International Journal of Engineering Research & Technology (IJERT)* <http://www.ijert.org> ISSN: 2278-0181 IJERTV9IS090345
Published by :

www.ijert.org Vol. 9 Issue 09, September-2020

Cheriyana, S., Ibrahim, S., Mohanan, S., & Treesa, S. (2018, August). Intelligent sales prediction using machine learning techniques. In 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE) (pp. 53-58). IEEE.

Cheriyana, S., Ibrahim, S., Mohanan, S., & Treesa, S. (2018, August). Intelligent sales prediction using machine learning techniques. In 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE) (pp. 53-58). IEEE.

Research paper

© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, S Iss 1, 2022

Bajaj, P., Ray, R., Shedge, S., Vidhate, S., &Shardoor, N. (2020). Sales Prediction Using Machine Learning Algorithms. International Research Journal of Engineering and Technology (IRJET), 7(06).

Chavan, S., Panchal, S., Sawant, T., & Shinde, J. (2020). Predicting Online Product Sales using Machine Learning.

Professor DevenKetkar(2018). "A Forecast for Big Data Sales based on Random Forests and Multiple Linear Regression", IJEDR 2018, VOL. 6, ISSUE 4. ISSN: 2321-9939.

Singh Manpreet, BhawickGhutla, Reuben Lilo Jnr, AesaanFS Mohammed, and Mahmood A. Rashid. "Walmart's Sales Data Analysis-A Big Data Analytics Perspective." In 2017 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), pp. 114-119. IEEE, 2017.

Panjwani, Mansi, Rahul Ramrakhiani, Hitesh Jumnani, Krishna Zanwar, and RupaliHande. Sales Prediction System Using Machine Learning. No. 3243. EasyChair, 2020.

Cheriyana, Sunitha, Shaniba Ibrahim, SajuMohanana, and Susan Treasa. "Intelligent Sales Prediction Using Machine Learning Techniques." In 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE), pp. 53-58. IEEE, 2018.