International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 07 | July - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

INTALLIGENT SALES PRIDICTION

[1] AFTHAB MAHABOOB A, [2] Dr. SHANKARA GOWDA B B

[1] 4th Sem Student MCA, Department of MCA, BIET, Davangere.

[2] Associate Professor and HOD Department of MCA, BIET, Davangere.

ABSTRACT:

Predictions and decision analysis must be integrated in an intelligent decision analysis system. The majority of business organizations rely heavily on a knowledge base and demand forecasts of trends in sales. The exactness in deals gauge gives a major effect in business. To improve forecasting accuracy and efficiency, data mining methods are very effective tools for uncovering buried information in a large dataset. This research focuses on the in-depth investigation and evaluation of understandable predictive models with the goal of improvingpredictions regarding future sales. With the volume of data and precision required for sales forecasting, traditional forecasting systems struggle. Various data mining methods could be used to solve these issues. The concepts of sales data and sales forecast are the subject of a brief discussion in this paper. The research work's concluding section provides a description of the various sales prediction methods and metrics. On the premise of an exhibition assessment, a most ideal prescient model is proposed for the deals pattern gauge. The reliability and accuracy of effective methods used for prediction and forecasting are summed up in the findings. The Gradient Boost Algorithm was found to be the best-fitting model, providing the highest level of forecasting and future sales prediction accuracy.

1. INTRODUCTION

The reliable sales trend prediction mechanism, which is implemented using data mining techniques to achieve the best revenue possible, is one of this study's primary goals. The present business handles tremendous archive of information. It is anticipated that the amount of data will continue to increase at an exponential rate. The measures are required to accommodate the transaction's processing speed as well as the anticipated rise in data volume and customer behaviour. New data mining methods and an intelligent sales trend prediction model with the highest possible level of accuracy and dependability are desperately needed in the e-commerce sector. Deals anticipating gives knowledge into how an organization ought to deal with its labour force, income also, assets. It is a crucial requirement for business planning and decision-making. It enables businesses to effectively plan their business strategies.

The company can increase revenue generation and market growth through accurate predictions. The foundation of sound budgeting is data mining, which converts huge amounts of data into usefulinformation for cost prediction and sales forecast. Forecasts of sales are important inputs to many organizational decision-making activities in operations, marketing, sales, production, finance, and other functional areas.

To serve an association's inner assets successfully, prescient deals information is significant for organizations while searching for getting venture capital. The studies move forward from a fresh perspective that focuses on selecting an effective method for accurately forecasting sales. Starting dataset thought about in this research had countless sections, yet the last dataset utilized for examination having a lot more modest size contrasted with the unique because of the riddance of non-usable information, excess passages and unimportant deals information.

In Section I, the methods for predictions and data mining are discussed, and in Section II, a literature review on sales forecasts is presented. With a visual representation of the generated results, the data tuning process and predictions are highlighted in Section III. The prescient examination and strategy on deals cost likewise talked about.

The results of various machine learning-based performance evaluations of prediction algorithms are presented. the analysis and conclusion of theresult include a summary of the research findings and future scope.

2. RELATED WORK

Businesses are always looking for a better model or method for data mining and critical data maintenance in order to be competent enough and make more money. Due to the exponential growth of the enormous volume of data used in e-commerce transactions, the business industry faces significant challenges in identifying an accurate data mining technique and an efficient prediction strategy. Deals information examination faces part of issues and significant parts of deals capabilities are ID of item quality, cost obsession, net deals acknowledgment and send off of new item.

A near report on information tuning and different bunching calculations on deals information is plainly made sense of in. As examined in grouping of information is vital in navigation. Clustering algorithms use distance metric-based similarity measures to identify distribution patterns. Clustering techniques are extremely helpful. Using supervised and unsupervised learning, it is possible to convert information from a large data set into a manageable format with the right data mining techniques. It is possible to make good business decisions with the right sales prediction technique. deals with the concepts and algorithms. Implementing XGBoost and adding support for the SigOpt Bayesian Optimization method can reduce the prediction error, as suggested by Korolev and Ruegg. Different datamining techniques can be used to forecast sales, allowing for the prediction of sales on any given day at any store.

We used a variety of data mining techniques to forecast store sales for this project. In order to familiarize ourselves with the task we had previously studied, the task required us to predictthe sales that would take place on any given day at any store.

3. METHODOLOGY

The fundamental motivation behind this exploration is to assess and examine the utilization of information digging strategies for deals gauging, to create models which are complete and solid.

3.1 Information Assortment and Planning.

The dataset utilized for this examinationdepends on an e-fashion store, for the three successive long periods of deals information. The e-fashion store's past sales records from 2015 to 2017 were compiled in order to make sales predictions. The information base incorporates Class, City, Sort of things and its depiction, number of things, Amount, Quarter, Deals Income, Year, SKU depiction, Week, Year. The initial dataset had a lot of entries, but theone that was ultimately chosen had a much smaller size than the original dataset because non-usable, redundant, and irrelevant data had been removed.

3.2 Exploratory Analysis

An exploratory analysis was carried out following the preprocessing of the data in order to clearly comprehend the nature ofour data.

Fig 1: Stage of Data Mining

Understanding, preparation, modelling, evaluation, and deployment are the stages of the data mining model.

the visualization of the quarter-by-quarter increase in sales revenue from 2015 to 2017 is shown as 1,2,3,4, respectively. This showsthat in quarter 4 of 2016 the business incomeis high and in quarter 3 of 2015 shows an extraordinary reduction in the income produced.

3.3 Outlier detection

sales	Year of Data		
Quarter	2015	2016	2017
1	2,425,084	3,338,276	3,032,690
2	1,623,062	2,852,142	2,447,278
3	1,254,177	2,885,557	3,306,000
4	1,795,108	4,199,956	2,107,128

Table1: Data set

This procedure does all of the necessary model optimization and data preprocessing. Exception identification cycle can be used to

send the model or as a beginning stage for additional improvements and supportive in showing nonexclusive data which is autonomous of the models. The quality of the data, particularly the quality of each data attribute, is the primary focus. In addition, these think about getting rid of data attributes that don't add much value.

3.4 Prediction

Prediction is concerned with things that will happen in the future. the utilization of AI calculations works on the insight of the framework without manual mediation.

According to Ethem alp Aydin's definition, "Machine Learning (ML) is used to optimize the performance criterion using sample data or the past experience."

On the training dataset, we tested the performance of three different machine learning algorithms inthis study. The most accurate algorithm for the prediction is chosen based on performance accuracy.

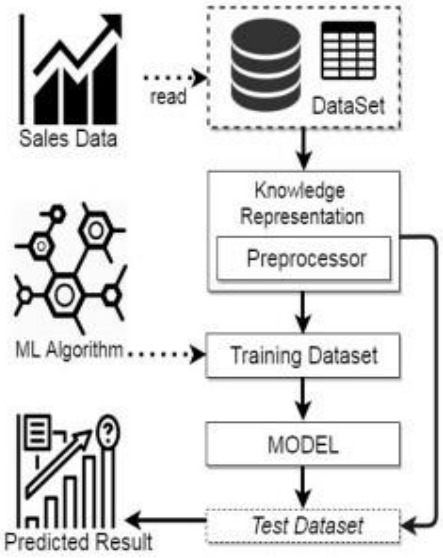

Fig 2: System Architecture4.RESULTS

AND ANALYSIS

The performance of the classification algorithms is primarily focused on classification accuracy, accuracy in each class, and the confusion matrix, which displays the number of predictions for each class in relation to the instances of each class. Root Mean Square Blunder, Mean Square Blunder,

Outright mistake are determined and normal of the blunder is displayed in the result in the Table III as the Mistake Rate. This action serves to personality whether the given expectation is off-base by andlarge.

The comparison of the three algorithms based on their prediction performance reveals that Gradient Boost Algorithm has an overall accuracy of 98%, followed by Decision Tree Algorithms, which have an overall accuracy of nearly 71%, and Generalized Linear Model, which has an overall accuracy of 64%. Last but not least, the empirical evaluation of the three chosen algorithms reveals that Gradient Boosted Tree provides the model with the best fit. The classification accuracy rate can be as high asone hundred percent, but the GBT model was analyzed and found to be about ninety-eight percent accurate. With the help of a robust data set and models like Grabit, if the GBT implementation can be improved further, the accuracy rate will improve.

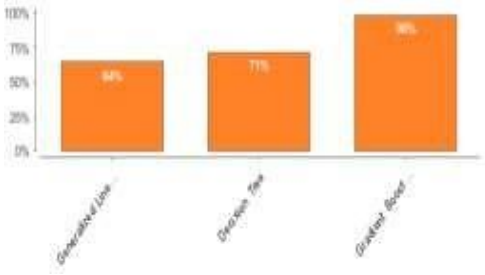

Fig3: Comparing Two model

5.CONCLUSION

The researchers came to the conclusion that businesses cannot handle huge amounts of data without an intelligent sales prediction system. Business choices are in light of speed and exactness of information handling procedures. An efficient mechanism for data tuning and decision making will be provided by the machine learning approaches discussed in this research paper. Companies must be equipped with cutting-edge methods for anticipating attractive sales turnover in order to be successful in business and accommodate a variety of customerbehaviours. Nearly 85,000 records were used in our research to compare algorithms. Since the hour of execution was gigantic and to oversee such a huge arrangement of records are intricate, a portion of the

records were disposed of, during the examination stage. At the same time, fields and qualities, utilized in this examination were deficient for the further investigation. It was the most difficult obstacle we encountered during the research. However, we had applied effective ML methods for prediction and forecasting to weigh our efforts thoroughly. Using Big Data as a tool for predictive analytics in sales forecasting can speed up the current studies. The enormous information examination and determining are estimated as the imperative fields in the advanced business situation

REFERENCES

[1] Huang, Q., and F. Zhou (March 2017). Research on retailer information

bunching calculation in light of flash. The Proceedings of the AIP Conference (Vol. 1820, No. 1, p. 080022). AIP Distributing.

[2] A. Sayl, I. Ozturk, and M. Ustunel are the authors. Brand reliability examination

framework utilizing K-Means calculation. 1(3), Engineering Technology and Applied Sciences Journal.

[3] Maingi, M. N., A Survey of Sales Data Mining Clustering Algorithms.

[4] Sastry, S. H., P. Babu, and M. S. Prasada Investigation and

Forecast of Deals Information in SAP-ERP Framework utilizing Bunching

Calculations. arXiv preprint arXiv:1312.2678.

[5] Arya, N., and V. Shrivastava an investigation of various clustering algorithms using data on retail sales. Int. Comput. J. Commun. Netw, 1(2).

[6] Digital Rajagopal, Using a data mining technique, clustering customer data. arXiv preprint arXiv:1112.2663.

[7] Wu, H. C., Tsai, C. W., and C. F. Tsai a brand- new method for data clustering in large database data mining. Algorithms, Networks, and Parallel Architectures, 2002. I-SPAN'02. Proceedings.

International Conference on 315-320). IEEE.

[8] Kaur, N., and A. K. Mann. Audit paper on grouping strategies. Computer Science and Technology: AGlobal Journal

[9] Sales Prediction Using Effective Mining Techniques by N. Shah, M. Solanki, A. Tambe, and Dhangar

[10] Korolev, M., and Ruegg, K. (2015). Trees were able to predict store sales thanks to gradient.

[11] Sales Forecasting for Retail Chains by A. Jain, M. N. Menon, and S. Chandra.