

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/365344800>

BMSP-ML: big mart sales prediction using different machine learning techniques

Article in *IAES International Journal of Artificial Intelligence (IJ-AI)* · June 2022

DOI: 10.11591/ijai.v12.i2.pp874-883

CITATIONS

2

READS

1,631

7 authors, including:



Rao Faizan Ali

Universiti Teknologi PETRONAS

33 PUBLICATIONS 401 CITATIONS

[SEE PROFILE](#)



Amgad Muneer

Universiti Teknologi PETRONAS

68 PUBLICATIONS 655 CITATIONS

[SEE PROFILE](#)



Ahmed Almaghthwi

King Khalid University

9 PUBLICATIONS 74 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Information Security Policy Compliance in Oil and Gas organizations [View project](#)



Training Centre Management System [View project](#)

BMSP-ML: big mart sales prediction using different machine learning techniques

Rao Faizan Ali¹, Amgad Muneer², Ahmed Almaghthawi³, Amal Alghamdi⁴, Suliman Mohamed Fati⁵,
Ebrahim Abdulwasea Abdullah Ghaleb²

¹Department of Computer Science, Faculty of Software Engineering, University of Management and Technology, Lahore, Pakistan

²Department of Computer and Information Sciences, Faculty of Science and Information Technology, Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia

³Department of Computer Science, Collage of Science and Art at Mahayil, King Khalid University, Muhayel Aseer, Saudi Arabia

⁴Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia

⁵Department of Information System, College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia

Article Info

Article history:

Received Aug 24, 2021

Revised Nov 4, 2022

Accepted Dec 4, 2022

Keywords:

Linear regression

Machine learning

Random forest

Ridge regression

Sales prediction

ABSTRACT

Variations in sales over time is the main issue faced by many retailers. To overcome this problem, we attempt to predict the sales by comparing the previous sales data of different stores. Firstly, the primary task is to recognize the pattern of the factors that help to predict sales. This study helps us understand the data and predict sales using many machines learning models. This process gets the data and beautifies the data by imputing the missing values and feature engineering. While solving this problem, predicting the monthly sales value is significant in the study. In addition, an essential element is to clear the missing data and perform proper feature engineering to better understand them before applying them. The experimental results show that the random forest predictor has outperformed ridge regression, linear regression, and decision tree models among the four machine learning techniques implemented in this study. The performance of the proposed models has been evaluated using root mean square error (RMSE).

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Amgad Muneer

Department of Computer and Information Sciences, Universiti Teknologi PETRONAS

32610 Seri Iskandar, Perak, Malaysia

Email: muneeramgad@gmail.com

1. INTRODUCTION

The sale is a lifeline of every business sales prediction that significantly impacts companies. Accurate predictions benefit the organization to maintain the standard and increase the company's lifestyle by using different strategies [1]. Typically, a prediction is based on the knowledge of previous studies with a deep focus on the conditions and then applies various factors, including customer's taste, culture, marketplace, and many more. In short, we can say that our prediction depends upon the previous study results [2]. Every business needs to be good in profit, and profit does not mean that stock sales are at maximum but also avoid the extra stock. Every retailer must maintain the stock according to the requirements and check the flaws and drawbacks that lead the sales down. Therefore, the proposed study deals with the same problem by predicting the store's sales [3].

Therefore, this paper divides the prediction into different phases, including preprocessing the data. In the first section, we explore and impute the missing values of the data by using statistics, then we use

feature engineering to explore more data and are ready to pass in models. Splitting the data into train and test variables with some ratio is better. Scikit-learn gives us the reliability to split the data into the desired ratio. This process gives a benefit to avoiding over and to underfit. After, we applied machine learning models to predict the results. We call the regressor of each method and then predict and compare it with the test values. We also see the factors like weather, culture, and events, which depend on the sales [1].

We need a large amount of data for any prediction. To forecast sales, we need a statistical model such as regression techniques and random forest models for prediction [4]. To access more accuracy, we need models that are combined different approaches. Many hybrid models give better results in predictions [2], [5]. Xia and Wong [3] highlight the difference between classical and modern methods. Most models are not asymmetric in real-world sales data because they are linear models. However, in current cases, these issues could be handled using root mean square error (RMSE) [6] and mean absolute error (MAE), which have been employed as a performance measure of this study. We performed many machine learning models one by one and compared the values to check the models' performance. We use the general flow of the data science in which we hypothesize the data, explore the data, impute the missing values, do some feature engineering and then select the model that best fits the data sets. In (1) and (2) shows the performance measure of MAE and RMSE.

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_{\text{predict}} - x_{\text{actual}}| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (|x_{\text{predict}} - x_{\text{actual}}|)^2} \quad (2)$$

This paper is organized: section 2 presents a comprehensive literature review and discusses the related works. Section 3 describes the characteristics and nature of the data set utilized and discusses a comparison performed over the data set based on the results of three algorithms. Section 4 provides the model performance evaluation, and finally, section 5 concludes the paper.

2. LITERATURE REVIEW

The critical factor of every big mart or store chain is sales. The sale factor has two faces, profit and loss. Therefore, to maintain the mart's standard, the sales graph has to be good every year. This can only be possible by forecasting the sales correctly or approximately well. In this section of the literature view, we see many previous searches and methods used to predict sales as perfectly as possible by many researchers. Which are restate: Ait-alla *et al.* [7] suggested a scientific framework for the manufacture and design of a robust clothing distributor. The author's focal point is decision-making power at various distribution centers. Another proposed model and claims that his model is best to perform robustly and perfectly according to the customer's demand. A different mechanism is proposed for machine learning in different fields [8]. Langley and Simon [9] presented a technique for the business that used data mining called rule induction technique. They give a solution based on the mining of data.

On the other hand, the author describes the distribution related to pharmaceutical firms. This paper focuses on two major issues: the stock never goes out of stock, and the second is related to the customer's satisfaction on behalf of sales [10]. Punam *et al.* [11] proposed a model based on two-level statistical analyses that help forecast store sales. In another study proposed by [12], the author uses the neural network to predict clog's weekly sales store. This forecasting helps the owner to balance the stock regularly. Das and Chaudary [12] proposed a model based on a linear and non-linear statistical framework for forecasting trading. They use both regressions in the model and separate the results based on the given data sets.

Additionally, many researchers research the sales market prediction, same as Beheshti-Kashi *et al.* [13] work on the fashion market and predicting sales using general prediction algorithms. Asur and Huberman [14] contribute to the famous blogging site 'Twitter' They also focus on the revenue of movies at the box-office. Another study proposed by O'Connor [15] contributes to the music field. They train their model to check that the chatter matter in the selection of music and which type of music gives the best sales. Bermingham and Smeaton [16] work on Twitter sites' political affairs and predict the election results on behalf of posts. Prediction of the stock exchange is discussed in this paper [17]. Many papers are not related to sales prediction; however, to get the idea of forecasting, we read that paper [18], forecasting the long-term generation of electric power based on hourly monitoring and the previous year's record.

Similarly, another research on the rental demand for bicycles depends upon the environmental weather, and researchers use the Regression model for this purpose [19]. The right location gives an extra advantage to sales of retail stores. Karna *et al.* serve their services by selecting the right place or location for the store to impact monthly sales [20]. They are many factors that participate in forecasting. Data is directly

proportional to the prediction; the prediction will be wrong if the data set is incorrect. Under the variability of items, it is difficult to predict sales. There are many factors that give a negative result in the case of uncertainty. Yuan *et al.* [21] deal with the effective sales of E-commerce and discuss the design and management that help predict the sales on behalf of user behaviors. Yuan *et al.* [21] proposed a model that focuses on predicting current sales and tries to sort out the best relation b/w the customer's choice and stock selection. Fakharudin *et al.* [22] a researcher that helps publisher to overcome their problem of publishing the right quantity of newspapers, use the neural network in their techniques and train their model on behalf of previous data sets of newspapers and the confusion matrix of their model gives the result value between the forecasting and the original value is strongly co-related. Significant components of almost every data science research are data mining, machine learning, and data visualization, so that is why we used to read many research papers related to major components. One of the research is data mining on the urban traffic prediction we used in our research [23]. This paper focuses on the classifications model and studies traffic behaviors to predict traffic affairs.

Even for small business, sales analysis is not an easy task. This research is based on the inspection of the customer's bargain and updates the critical adjustments based on the analysis of the customer's bargain data [24]. The authors set the plans by adding a forum in the mobile application so that customers will buy directly from sellers. Nagamma *et al.* [25] conduct their research to find the relationship between ticket booth revenue and the film's online viewers. They use a support vector machine (SVM) classification model to train and predict the online viewers' movie revenue. South Korean researchers named Kim and Youn use the huff model to predict agricultural product sales [26]. Their search is based on the value of λ , which is reversed in the case of 1 and store mass is directly affected by travel time if the value of λ is below 0.5.

3. METHODOLOGY

Forecasting mart sales is an exciting problem; data science also helps improve sales and business. We divide our proposed research method into the following steps: i) hypothesis making, ii) exploration of data, iii) cleaning the data, and iv) feature engineering. In order to build the model, at the very first, we have looked at the dataset hypothetically and extracted all the possible information to understand the data and research. Next, we explore our data's attributes, and the values for the conclusion. Missing values are not a big problem with data science, so we clean our data before feature engineering. It is quite possible that we need more variables that are not present in our data for analysis, so we create new variables for better understanding, and finally, we create a model to predict the sales.

3.1. The problem assertion

We have collected the data of 1560 of ten stores because we want to predict the sales of a certain product in different stores and find the reason for the difference in sales. The critical factor is determining the attributes that increase sales. Let us consider some hypotheses that can depend upon the sales factor.

3.2. Mart hypothesis

- According to the general hypothesis, store sales depend upon whether the stores are in urban or rural areas.
- Another factor is population; it is clear that more population affects the sales.
- The size of a store may also affect the sales; the more significant the store would print the store's quality on the customer's mind.
- Opponent factor how many challenges are in the market.
- Advertisement also a pillar of sales.
- Location matters. The sale would be satisfied if the store is near a dense population.
- Customers satisfaction.

3.3. Product hypothesis

- People always need quality products related to some brands.
- Designing the product with good packaging also prints the excellent quality of the customer's mind.
- Stores have a related product or not., *i.e.*, stores should have the product related to daily usage.
- Discount Promotions also attract customers.
- Product price.

These are some hypotheses which we considered important; this is not enough. There are also many factors which depend on sales, but we write according to our search. These hypotheses are for a better understanding of the data. Hence, the dataset detail is provided in the next section.

3.4. Data exploration

In this step, we start exploring our data by comparing the data set's attributes and our hypothesis and checking the difference as given in Table 1 and Figure 1. We exclude that it differs slightly from our hypothesis and the original data. Hence, that is why we unvaried explore quality in original data, hypothesis, and the other way round. We got our data by web scrapping. Our data is split b/w two files named train/test, so we generally concatenate and split it before the model step. For this, we use the Concat function and shape to check the file size before and after concatenation. Therefore, before concatenation, the form of the train and test files are different than after concatenation. Now our data set is ready for the cleaning data. First, check the missing values in the dataset and the number of missing data in each attribute.

Table 1. Attributes vs hypothesis

Attributes	Explanation	Hypothesis link
Product ID	Product identification	Unique identification
Product W	Product weight	Absent in hypothesis
Product Ingredient	Good quality ingredients	Relation with quality
Product Display	How much importance of a product in the display	Present in hypothesis, related to the display.
Product Type	Product category	Absent in our hypothesis
MPR	What is the price	Present in hypothesis
Year of Establishment	How old the store	Absent in our hypothesis
Size	How large the store	Has relation with the capacity of the store
Location	Whether the store in rural or urban	Present in hypothesis
Product sale	The monthly average sale of a product	In hypothesis

Outline of data for now:

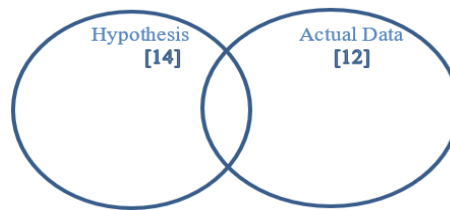


Figure 1. Effects of selecting different switching under dynamic conditions

We have two types of variables to predict any value: target and process variables. In our case, we have to predict the outlet sales, so our target variable is outlet sales, but from Table 2, we can see that there are 5680 missing values, so we have to fill these missing values before building a model to predict sales as perfect as well. Describe () function is used for the statistical understanding; let us suppose we want to check the dataset's attribute's min, max, and avg. Therefore, the data exploratory analysis used in this study is presented in Table 3.

Additionally, the exploratory data analysis was performed to provide us with statistical information, some vital information that we have to understand. This makes it unclear that if a product is being sold, how can it be possible that its appearance is zero. If we use since to define the year of establishment, then it looks cool to mention the range of years. Besides, we know that there are 1559 items in ten stores and sixteen unique product types. We ignore the Product identification and source because these attributes cannot predict sales. The selected frequencies of different nominal variables from the employed dataset are shown in Table 4. The occurrence rate of kinds for variable Product_FAT_Content.

The information in section 3.4 gives us some information about the data. Many values use similar observations *i.e.*, LF is for Lowfat, regular is for regular. Many are meaningful numbers, so it is better to concatenate them for valuable outcomes. Now our data is ready for the cleaning process.

3.5. Data cleaning

Outliers and missing values significantly impact data, so treating outliers and handling the missing values is essential. Data cleaning is how we improve our dataset by cleaning the data. We have two essential attributes in our datasets and need to fill in the missing values. The employed codes give a comparison of missing values before and after. For example, product weight has 2439 missing values before and zeroes after

implementing the above code. Output speaks that there are no more missing values. Further, another attribute has a missing value, Outlet_size, by using mode in aggfunc.

Therefore, we have two essential attributes and need to fill in the missing values. The implemented codes give a comparison of missing values before and after. For example, product weight has 2439 missing values before the implemented method is used, and the zeroes method is performed. Output speaks that there are no more missing values. Further, another attribute with a missing value, Outlet_size, uses mode in aggfunc. Now attribute "Outlet_Type" also cleaning with zero missing value as shown in Table 5. Now, our data is ready for feature engineering. The feature engineering step is discussed in the next section.

Table 2. Checking for missing values in the dataset

Attribute	Missing value
Fat_Content	0
Product Identifier	0
MPR	0
Outlet sale	5680
Product Type	0
Product Display	0
Weight of Product	2440
Year of Establishment	0
Outlet Identifier	0
Location	0
Size	4015
Outlet Type	0
Source	0

Table 3. Data exploratory analysis

	Product MRP	Product outlet sales	Product appearance	Product weight	Establishment year
Count	14204.0000	8523.0000	14204.0000	11765.000	14204.0000
Mean	141.004977	2181.288914	0.065953	12.792854	1997.830681
Std	62.086938	1706.499616	0.051459	4.652502	8.371664
Min	31.29000	33.290000	0.00000	4.55500	1985.0000
25%	94.012000	834.247400	0.027036	8.710000	1987.0000
50%	142.247000	1749/331000	0.054021	12.600000	1999.0000
75%	185.855600	3101.296400	0.094037	16.750000	2004.0000
Max	266.888400	13086.964800	0.328391	21.350000	2009.0000

Table 4. Frequencies of different nominal variables

LowFat	8484
Regular	4825
LF	521
Reg	196
Low fat	177

Table 5. Mode for each Outlet_Type

GroceryStore	Small
SuperMarket 1	Small
SuperMarket 2	Medium
SuperMarket 3	Medium

3.6. Feature engineering

Feature engineering helps us understand the data for better analysis. Here, we will create some new variables from the original data. When we explore our data for analysis, we observe that many shades have to be resolved. For example, remember talking about the concatenation of two Supermarkets because we almost think that both have the same sales. So, check whether it is true or not. The output shows the difference between supermarkets sales, and there is a significant difference b/w them, that why this idea will not be valuable to combine the data of markets.

After data cleaning and data rangling, we were going to build our predictive model. This model plays a vital role in predicting sales. Our research uses six decision tree models, random forest, linear regression, and many more. So, let us start by using the baseline model initially. The baseline is independent of the forecasting model requirement; we use the mean of all the sales to forecast the sales. This simplifies it because we can predict by taking an average of all mart sales using the baseline model. On the other hand, linear regression is a beneficial and essential prediction model, and the library scikit-learn gives many machines learning models, one of which is the regressions model. Figure 2 shows the linear regression baseline model results. Now our public leaderboard is 1202, which shows the result better than the baseline model, but despite this, we can see many larger coefficients in magnitude, so we attempt to use ridge models. The public leaderboard is 1203, which looks more helpful than linear regression. Figure 3 shows the ridge model results with their mean and the attributes of the selected feature.

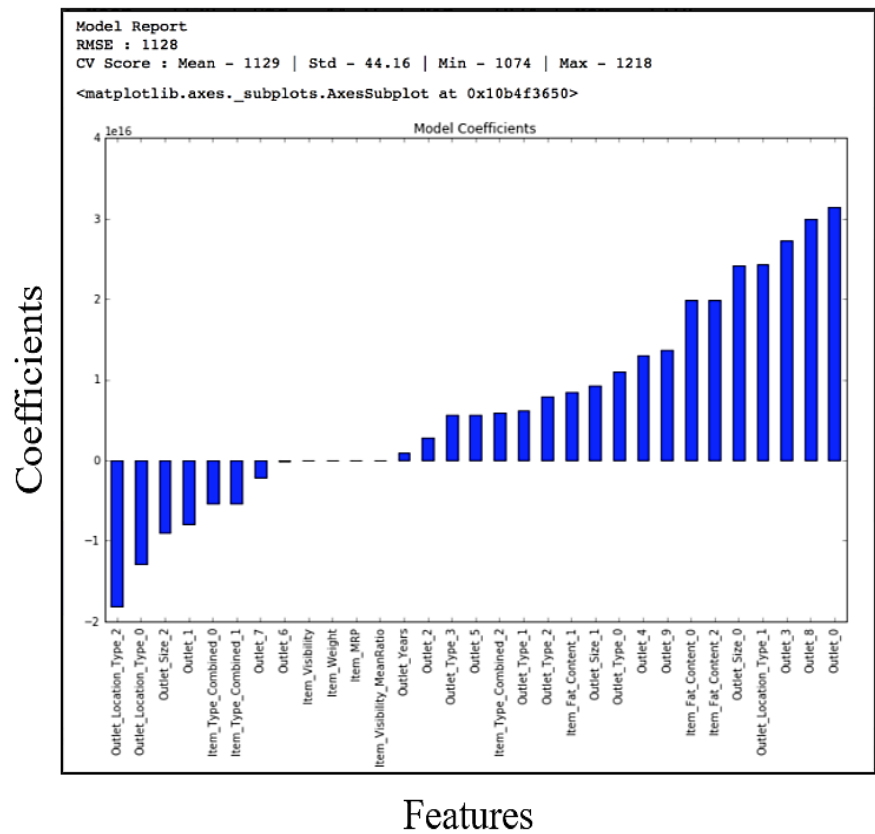


Figure 2. Linear regression baseline model results

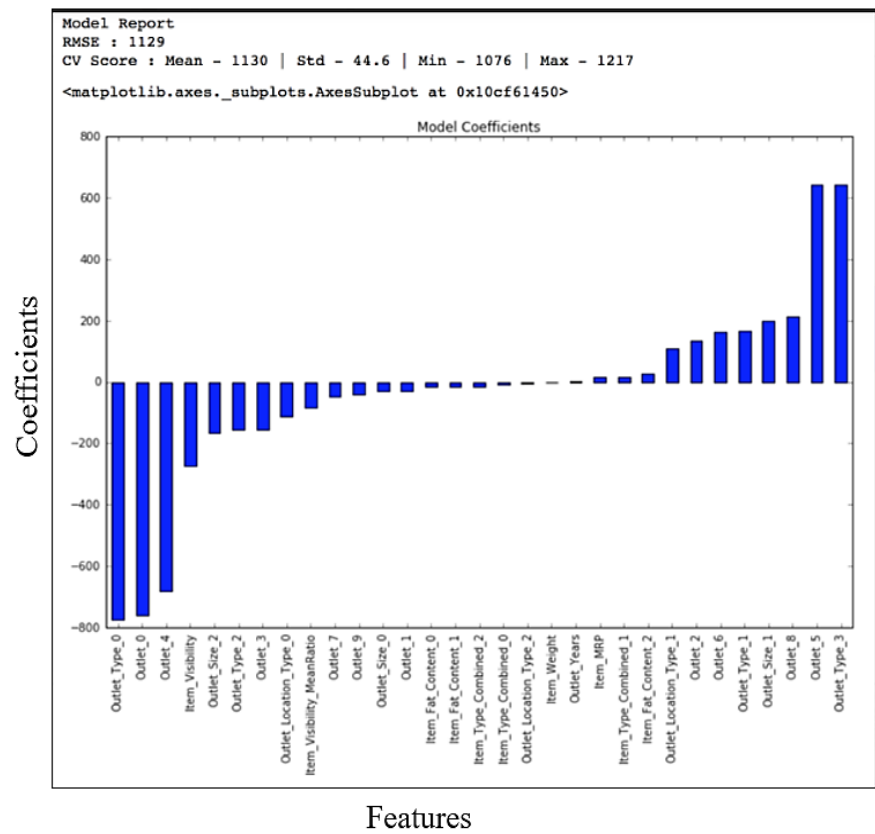


Figure 3. Ridge model results

The decision tree is also a beneficial model for predictions. That is why we include this model in our study and compare the results with the previous model by visualization. Now sees the comparisons between different methods of visualization. In this experiment, we got a score on the public leaderboard is 1157. Hence, our result is not satisfying; that is why we try another model named the random forest model. This gives more satisfying results in forecasting because the public leaderboard shows 1157. Therefore, the achieved result is satisfying because of 1157 scores after using the decision tree model.

4 EXPERIMENT RESULTS AND DISCUSSION

Big marts' sales prediction is conducted by using many algorithms. We use machine learning algorithms to solve our dataset. Initially, we want to predict the sales of the mart by studying the sales of different marts with specific attributes, so that is why we set the "Price" attribute with the dependent variable, and we see above there are more than 10 attributes that we use as some independent variables. Our dataset contains two different train and test files; we concatenate our files to understand the data better. A hypothesis is necessary to check the possible attributes of the data. Also, it gives the understanding between the data scientist and the prediction. Therefore, we created some hypotheses and then compared them with the existing data. We saw a little bit of difference b/w the hypothesis and the data, and then we adjusted the data with our hypothesis attributes. Next, we moved to explore the data. In this part, we check the basic statistics of the dataset and missing values in the data.

We found that three attributes have many missing values, and we will resolve these missing values in the coming section. Moved to the nominal variables, we checked the unique values in the data and found that there are 4-5 categorical variables. After, we need to impute the missing values further. Attributes named 'Product-weight' and 'Outlet-Size' filled the missing values, 2439 and 4016, respectively. We use mean and mode for the missing values product weight and outlet size, respectively. Feature engineering is the process of creating some new attributes for a better understanding of the data.

Once we were ready with the data, we had to make a model. We used six machine learning models to predict sales and compare the results. When we applied the baseline model algorithm, we got a score of 1772 on the public leaderboards, which was not satisfying. When working with machine learning (ML) models, it is good to split the files into train and test, but using the built-in function of sci-kit learn lib is a good idea. The advantages of using a split function are avoiding over-fitting and under-fitting, so use a split function before using any machine learning algorithm. The linear regression model gives the 1201 score in the public leaderboard. Ridge regression gives the 1202 public leaderboard score, which looks better than linear regression. Decision tree gives the RMSE of 1057 and the mean common vulnerabilities and exposures (CVE) of 1090. This shows that the model is overfitting. Take a quick example of four top variables with eight deep lengths 151 minimum sample leaves. Now the value of RMSE has risen from 1057 to 1070. Random forest gives the 1153 value of the public leaderboard. Improvement can be possible when we change the value of max-depth and the number of trees, which may increase the computational burden of using the number of trees. Table 6 provides a comparative analysis of the leaderboard values.

Table 6. A comparative analysis between different proposed methods

Model name	Public leaderboard score	RMSE
Baseline model	1772	Not consider
Linear regression	1201	1128
Ridge regression model	1202	1129
Decision tree model	1161	1507
Random forest model	1153	1070

Additionally, our findings predict the sales on behalf of the other marts sales data. Our results are more accurate and near the original test data. Thus, this study method performs much better in using and predicting sales. We perceive different scenarios in which different models are best among all the models. In our case, all the model's public result is above 1150, so all are good enough to use. Figure 4 provides a comparative analysis between different proposed methods.

Among all, the tree classifier shows high values of RMSE. Ridge Regression scores higher in the public leader board than the remaining models. Therefore, the outcomes show the importance of results in our research. While the proposed model has many benefits, we think there are some limitations. Our research only predicts sales based on specific attributes, but it is not good enough to use globally; for example, we do not include disasters in our research, so our prediction is invalid in case of disaster. For prediction, we use many machine learning models and then evaluate and compare the result of the final use. Data is always not

in a manageable manner, so we need to beautify our data before modelling. This paper tells us the sequence flow of research and further attain outcomes with machine learning models. For new retailers, this research gives the benefit in an investment sense. Initially, how much income is needed to start a new retail store or big mart and the monthly outcomes.

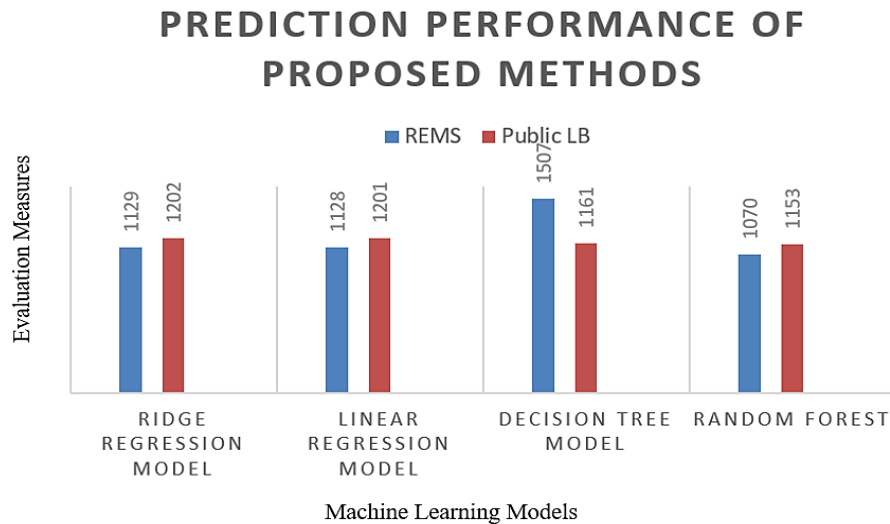


Figure 4. A comparative analysis between different proposed methods

5 CONCLUSION

This research proposed a framework that predicts mart's sales using a machine learning model and different techniques. This research uses the data of various marts and then combines and analyses the data so that any mart can check the product's demand and sales overall. This forecasting helps the retailers to set the stock quantity more accurately. Next, we use many models to study the scores outcomes. Considering the outcomes, it is suitable for present data, but extensive data might be unsuitable or change the model selection. For more accuracy, we need a massive amount of data with minimum outliers. Therefore, the authors seek to check the individual product demand in a particular area in future work. Further, in the future, a retailer checks the score of a specific product by entering product attributes and its store's information, like location, and culture. Also, we consider an online App for the customer's review regarding the stores and specific products for future work. This App works as a ranking App. Customers rank the stores by giving feedback; this helps the other customers to move on towards the stores. Sort of forum allows checking the demand for a specific store's specific product. Also, this kind of portal helps the retailer compare the stocks and scores of the public leader board.




REFERENCES

- [1] S. Thomassey, "Sales forecasts in clothing industry: the key success factor of the supply chain management," *International Journal of Production Economics*, vol. 128, no. 2, pp. 470–483, Dec. 2010, doi: 10.1016/j.ijpe.2010.07.018.
- [2] N. Liu, S. Ren, T.-M. Choi, C.-L. Hui, and S.-F. Ng, "Sales forecasting for fashion retailing service industry: a review," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–9, 2013, doi: 10.1155/2013/738675.
- [3] M. Xia and W. K. Wong, "A seasonal discrete grey forecasting model for fashion retailing," *Knowledge-Based Systems*, vol. 57, pp. 119–126, Feb. 2014, doi: 10.1016/j.knsys.2013.12.014.
- [4] A. Muneer and S. M. Fati, "A Comparative analysis of machine learning techniques for cyberbullying detection on twitter," *Future Internet*, vol. 12, no. 11, p. 187, Oct. 2020, doi: 10.3390/fi12110187.
- [5] A. Muneer, S. Mohd Taib, S. Mohamed Fati, A. O. Balogun, and I. Abdul Aziz, "A hybrid deep learning-based unsupervised anomaly detection in high dimensional data," *Computers, Materials & Continua*, vol. 70, no. 3, pp. 5363–5381, 2022, doi: 10.32604/cmc.2022.021113.
- [6] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, and M. Sartin, "Combining content-based and collaborative filters in an online newspaper," 1999.
- [7] A. Ait-Alla, M. Teucke, M. Lütjen, S. Beheshti-Kashi, and H. R. Karimi, "Robust production planning in fashion apparel industry under demand uncertainty via conditional value at risk," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–10, 2014, doi: 10.1155/2014/901861.
- [8] I. Bose and R. K. Mahapatra, "Business data mining — a machine learning perspective," *Information & Management*, vol. 39, no. 3, pp. 211–225, Dec. 2001, doi: 10.1016/S0378-7206(01)00091-X.
- [9] P. Langley and H. A. Simon, "Applications of machine learning and rule induction," *Communications of the ACM*, vol. 38, no. 11, pp. 54–64, Nov. 1995, doi: 10.1145/219717.219768.




- [10] A. Ribeiro, I. Seruca, and N. Durão, "Improving organizational decision support: detection of outliers and sales prediction for a pharmaceutical distribution company," *Procedia Computer Science*, vol. 121, pp. 282–290, 2017, doi: 10.1016/j.procs.2017.11.039.
- [11] K. Punam, R. Pamula, and P. K. Jain, "A two-level statistical model for big mart sales prediction," in *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*, Sep. 2018, pp. 617–620, doi: 10.1109/GUCON.2018.8675060.
- [12] P. Das and S. Chaudhury, "Prediction of retail sales of footwear using feedforward and recurrent neural networks," *Neural Computing and Applications*, vol. 16, no. 4–5, pp. 491–502, May 2007, doi: 10.1007/s00521-006-0077-3.
- [13] S. Beheshti-Kashi, H. R. Karimi, K.-D. Thoben, M. Lütjen, and M. Teucke, "A survey on retail sales forecasting and prediction in fashion markets," *Systems Science & Control Engineering*, vol. 3, no. 1, pp. 154–161, Jan. 2015, doi: 10.1080/21642583.2014.999389.
- [14] S. Asur and B. A. Huberman, "Predicting the future with social media," in *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Aug. 2010, pp. 492–499, doi: 10.1109/WI-IAT.2010.63.
- [15] A. Bermingham and A. F. Smeaton, "On using Twitter to monitor political sentiment and predict election results," in *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011)*, 2011, pp. 2–10.
- [16] B. O'Connor, R. Balasubramanyam, B. R. Routledge, and N. A. Smith, "From tweets to polls: linking text sentiment to public opinion time series," in *Proceedings of the Fourth International Conference on Weblogs and Social Media*, 2010, pp. 23–26.
- [17] E. Gilbert and K. Karahalios, "Widespread worry and the stock market," in *Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010*, 2010, pp. 23–26.
- [18] R. K. Agrawal, F. Muchahary, and M. M. Tripathi, "Long term load forecasting with hourly predictions based on long-short-term-memory networks," in *2018 IEEE Texas Power and Energy Conference (TPEC)*, Feb. 2018, pp. 1–6, doi: 10.1109/TPEC.2018.8312088.
- [19] Y. Feng and S. Wang, "A forecast for bicycle rental demand based on random forests and multiple linear regression," in *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, May 2017, pp. 101–105, doi: 10.1109/ICIS.2017.7959977.
- [20] M. Karna, A. Rai, B. Karki, L. Karna, M. Rai, and S. Adhikari, "A study on selection of location by retail chain: big mart," *International Journal of Research -GRANTHAALAYAH*, vol. 7, no. 1, pp. 383–395, Jan. 2019, doi: 10.29121/granthaalayah.v7.i1.2019.1067.
- [21] H. Yuan, W. Xu, and M. Wang, "Can online user behavior improve the performance of sales prediction in E-commerce?," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2014, pp. 2347–2352, doi: 10.1109/SMC.2014.6974277.
- [22] A. S. Fakharudin, M. A. Mohamad, and M. U. Johan, "Newspaper vendor sales prediction using artificial neural networks," in *2009 International Conference on Education Technology and Computer*, 2009, pp. 339–343, doi: 10.1109/ICETC.2009.67.
- [23] S. K. Nejad, F. Seifi, H. Ahmadi, and N. Seifi, "Applying data mining in prediction and classification of urban traffic," in *2009 WRI World Congress on Computer Science and Information Engineering*, 2009, pp. 674–678, doi: 10.1109/CSIE.2009.906.
- [24] A. Zacharia, "Sales prediction for online shopping," *International Journal of Innovative Science and Research Technology*, vol. 4, no. 5, pp. 671–674, 2019.
- [25] P. Nagamma, H. R. Pruthvi, K. K. Nisha, and N. H. Shwetha, "An improved sentiment analysis of online movie reviews based on clustering for box-office prediction," in *International Conference on Computing, Communication & Automation*, May 2015, pp. 933–937, doi: 10.1109/CCAA.2015.7148530.
- [26] P. J. Kim and M. K. Youn, "Retail sales forecasting of agricultural products marts in a mid-sized city using the Huff model," *Journal of Food Agriculture and Environment*, vol. 8, no. 2, pp. 945–948, 2010.

BIOGRAPHIES OF AUTHORS






Rao Faizan Ali    received the bachelor's degree in Computer Science from COMSATS University Islamabad, Pakistan, and the M.Phil. Degree in Computer Science from the University of Management and Technology, Lahore, Pakistan. He is currently pursuing the Ph.D. degree with University Technology PETRONAS, Malaysia. He has eight years of experience in teaching and research. He has been with various computer science positions in financial, consulting, academia, and government sectors. He is currently working as a Research Officer with the Department of Computer and Information Sciences, University Technology Petronas, Perak, Malaysia. He can be contacted at email: rao_16001107@utp.edu.my.






Amgad Muneer    received the B.Eng. degree (Hons.) in mechatronic engineering from the Asia Pacific University of Technology and Innovation (APU), Malaysia, in 2018. He is currently pursuing the master's degree in information technology with Universiti Teknologi PETRONAS, Malaysia. He has authored several ISI and Scopus journal articles/conference papers. He is currently working as a Research Officer with the Department of Computer and Information Sciences, University Technology PETRONAS, Perak, Malaysia. His research interests include machine and deep learning, image processing, the Internet of Things, computer vision, and condition monitoring. He is a Reviewer in some international impact-factor journals, and he has published more than 30 scientific publications. He can be contacted at email: muneeramgad@gmail.com.






Ahmad Almaghthawi    received his bachelor's degree in Computer Science from Taibah University in 2015. He has a master's degree in the program Computer Science and Artificial Intelligence at Jeddah University. Currently, he works as adjunct lecturer at college of computer science and artificial intelligence at Jeddah University. His scientific interests are related to artificial intelligence, image and video processing, machine learning, and in IoT. He can be contacted at email: ahmed.almaghthawi.1991@gmail.com.






Amal Alghamdi    Currently, she is a master student in computer science and artificial intelligence at Jeddah University. She received her bachelor's degree in Computer Science from the Al-Baha University in 2014. Her interests in artificial intelligence, machine learning and deep learning. She can be contacted at email: dr.amal.alghamdi@gmail.com.



Suliman Mohamed Fati    obtained his BS.c. (2002), MS.c. (2009), and Ph.D. (2014) from Ain Shams University-Egypt, Cairo University-Egypt. And Universiti Sains Malaysia (USM)–Malaysia respectively. He is a Senior lecturer in Information Technology at INTI International University, Malaysia. His research interests focus on internet of things, machine learning, social media mining, cloud computing, cloud computing security, and information security. He has authored over 20 ISI and Scopus journal/conference papers, books, and book chapters. He is the lead editor for an edited book shortly be published by Wiley entitled “A Comprehensive Guide to IPTV Delivery Networks”. Besides, He is a member of different professional bodies as IEEE, IACSIT, IAENG, and Institute of Research Engineers and Doctors, USA. He is a reviewer in many international impact-factor journals, and also technical committee program member in a number of international. He can be contacted at email: amgad_20001929@utp.edu.my.



Ebrahim Abdulwasea Abdullah Ghaleb    received the B.Sc. and M.Sc. Bachelor of information technology (Hons) in Networking Technology Infrastructure University Kuala Lumpur, Malaysia, and He hold Master. Degree in Information system from The National University of Malaysia (Malay: Universiti Kebangsaan Malaysia, abbreviated as UKM). He is a Ph.D. student on information system with UTP Universiti Teknologi PETRONAS. He has authored or coauthored more than 9 refereed journal and conference papers, with Sustainability, Journal of Theoretical & Applied Information Technology, Solid State Technology and International Congress of Advanced Technology and Engineering, IEEE and Springer. His research interests include the applications of big data, healthcare evolutionary and heuristic optimization techniques to power system planning, operation, and control. He can be contacted at email: ebrahim_1800342@utp.edu.my.