

# BIGMART SALES PREDICTIVE MODEL ANALYSIS USING MACHINE LEARNING ALGORITHMS

Dr. V. NAGA GOPIRAJU<sup>1</sup>, K. MANASWI<sup>2</sup>, CH. VELANGINI JYOTHI<sup>3</sup>, B. VENKATA PAVAN KUMAR<sup>4</sup>

<sup>1</sup>Associate Professor,CSE,Chalapathi Institute of Technology,Guntur, India

<sup>2</sup>UG Student,CSE,Chalapathi Institute of Technology,Guntur, India

<sup>3</sup>UG Student,CSE,Chalapathi Institute of Technology,Guntur, India

<sup>4</sup>UG Student,CSE,Chalapathi Institute of Technology,Guntur, India

**ABSTRACT:** The aim is to build a predictive model and predict the sales of each product at a particular outlet. Using this model, Big Mart will try to understand the properties of products and outlets which play a key role in increasing sales. The sales forecast is based on Big Mart sales for various outlets to adjust the business model to expect outcomes. The resulting data can then be used to prediction potential sales volumes for retailers such as Big Mart through various machine learning methods. The estimate of the system proposed should take account of price tag, outlet and outlet location. A number of networks use the various machine-learning algorithms, such as linear regression and decision tree algorithms, and an XGBoost regressor, which offers an efficient prevision of Big Mart sales based on gradient. At last, hyper parameter tuning is used to help you to choose relevant hyper parameters that make the algorithm shine and produce the highest accuracy.

## 1. INTRODUCTION

Big malls & marts, in today's contemporary world, collect data connected to the sales of commodities or products with their many dependent or independent aspects as a crucial step to assist estimate future demand and inventory management. The dataset is a combination of item characteristics, customer data, and data linked to inventory management in a data warehouse, all constructed using numerous dependent and independent factors. In order to achieve more accurate forecasts and discover new and fascinating findings, the data is then tweaked and analyzed further. Random

forests and simple or complex linear regression models may then be used to

predict future sales. For the purpose of establishing new hubs, Big Mart patterns are closely scrutinized by data scientists. To get the best outcomes, data scientists use a machine to anticipate the transactions of Big Mart and then test different patterns by shop and product. Forecasting market trends is essential for many businesses that largely depend on their information base. Shopping centers and stores are always looking for ways to make it easier for the storeowner to pull in more customers throughout the day, so that the total volume of sales can be

calculated for purposes of inventory management, logistics and transportation management.

Different Machine Learning techniques including Linear Regression, Random Forest, Decision Tree, Ridge Regression, and XGBoost are used to gauge the number of trades in various Big Mart sectors depending on future customer demands. It is conceivable that the location of the shop is in a rural area or an urban area, thus it is likely that the conclusion of the transaction depends on these factors: the kind of business, the population in the area surrounding the store, the city in which the store is situated. The store's capacity, as well as the area's demographics and other factors, should be taken into account while analysing sales. Sales predictions are critical at a shopping mall since every company faces considerable demand. When it comes to creating and refining a company's market strategy, the ability to make better predictions is always beneficial. As a result, the company can predict client demand and better manage its inventory by tracking every item in its stores and Big Mart. Big Mart is a massive retail chain with locations in almost every country.

## 2. LITERATURE SURVEY

**“Nikita Malik, & Karan Singh, Sales Prediction Model for Big Mart”,**

Nikita Malik [1] has discussed about sale prediction using machine learning. She used Machine learning algorithm (linear regression, Random Forest etc.). She has done analysis on few products and established some correlation between

product and store. The accuracy is between 70% to 80%.

**Aditi Narkhede , Mitali Awari , Suvarna Gawali & Amrapal Mhaisgawali , Big Mart Sales Prediction Using Machine Learning Techniques**

Aditi Narkhede [2] has collected dataset of Big Mart and has used ML algorithm to find RMSE value. She has done some calculation to find RMSE value and is quite easy to use. She made calculation look easy to use.

**Kumari Punam, Rajendra Pamula, Praphula Kumar Jain (2018), A Two-Leval Statistical Model for Big Mart Sales Prediction**

Rajendra Pamula [3] has discussed about flow chart diagram to make us understand things easily. Here he used Machine learning and data mining as well. The accuracy is between 50 to 60%. He used some complex calculation to get the output which is not easy to understand.

**Sunitha Cheriyan, Shaniba Ibrahim, Saju Mohanan, Susan Treesa (2019), Intelligent Sales Prediction Using Machine Learning Techniques**

Saju Mohanan [4] has used data mining techniques and machine learning algorithm for sale prediction. He used decision tree and generalized linear model for prediction. The accuracy of the model is between 60 to 70%. He also has drawn system architecture to make things simple but the output is in very complex form.

## 3. EXISTING SYSTEM

Using Big Mart's sales data, it is possible to discover the importance of specific qualities of a product and how they effect sales. "Using

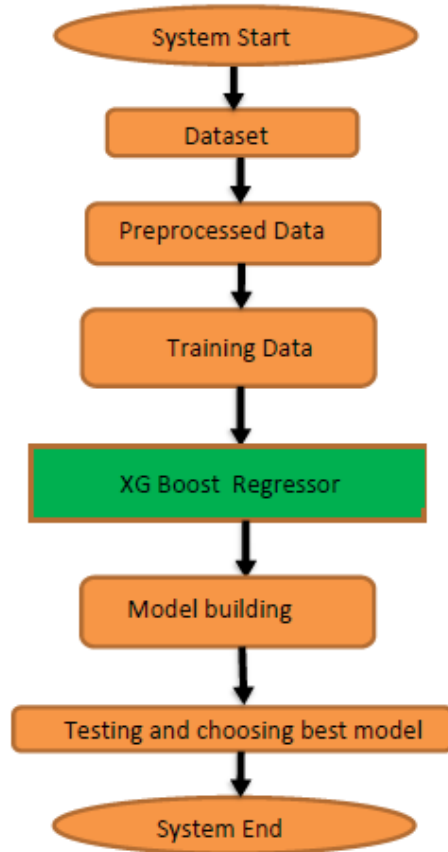
a predictive model, BigMart can assist reach this aim by discovering the main elements that may boost sales in each store and what adjustments can be made to the product or shop's attributes.

**4. PROPOSED SYSTEM**

The aim of the project is to understand the sales data, test a few hypotheses and finally build a predictive model to identify the sales of each product at a particular store, with the various products and store attributes available in the competition. Using this model, BigMart can try to understand the properties of products and stores which play a key role in increasing sales.

Raw data collected at big mart would be pre-processed for missing data, anomalies and outliers. Then an algorithm would be trained on this data to create a model. This model would be used for forecasting the final results.

**5. SYSTEM ARCHITECTURE**



**6. SCREEN SHORT**

|   | Item Identifier | Item Weight | Item Fat Content | Item Visibility | Item Type             | Item MRP | Outlet Identifier | Outlet Establishment Year | Outlet Size | Outlet Location Type | Outlet Type       | Item_Outlet_Sales |
|---|-----------------|-------------|------------------|-----------------|-----------------------|----------|-------------------|---------------------------|-------------|----------------------|-------------------|-------------------|
| 0 | FDA15           | 9.30        | Low Fat          | 0.016047        | Dairy                 | 249.8092 | OUT049            | 1999                      | Medium      | Tier 1               | Supermarket Type1 | 3735.1380         |
| 1 | DRC01           | 5.92        | Regular          | 0.019278        | Soft Drinks           | 48.2692  | OUT018            | 2009                      | Medium      | Tier 3               | Supermarket Type2 | 443.4228          |
| 2 | FDN15           | 17.50       | Low Fat          | 0.016760        | Meat                  | 141.6180 | OUT049            | 1999                      | Medium      | Tier 1               | Supermarket Type1 | 2097.2700         |
| 3 | FDX07           | 19.20       | Regular          | 0.000000        | Fruits and Vegetables | 182.0950 | OUT010            | 1998                      | NaN         | Tier 3               | Grocery Store     | 732.3800          |
| 4 | NCD19           | 8.93        | Low Fat          | 0.000000        | Household             | 53.8614  | OUT013            | 1987                      | High        | Tier 3               | Supermarket Type1 | 994.7652          |

**7. CONCLUSION**

In this work, basics of machine learning and the associated data processing and modeling algorithms have been described, followed by their application for the task of sales

prediction in Big Mart. On implementation, the prediction results show the correlation among different attributes considered, suggesting that should follow similar patterns for improved sales.

Multiple instances parameters and various factors can be used to make this sales prediction more innovative and successful. Accuracy, which plays a key role in prediction based systems, can be significantly increased as the number of parameters used are increased. Also, a look into how the sub-models work can lead to increase in productivity of system.

## 8. FUTURE SCOPE

Accuracy, which plays a key role in prediction-based systems, can be significantly increased as the number of parameters used is increased. Also, a look into how the sub models work can lead to increase in productivity of system. The project can be further collaborated in a web-based application or in any device supported with an in-built intelligence by virtue of Internet of Things (IoT), to be more feasible for use. Various stakeholders concerned with sales information can also provide more inputs to help in hypothesis generation and more instances can be taken into consideration such that more precise results that are closer to real world situations are generated. When combined with effective data mining methods and properties, the traditional means could be seen to make a higher and positive effect on the overall development of corporation's tasks on the whole. One of the main highlights is more expressive regression outputs, which are

more understandable, bounded with some of accuracy. Moreover, The flexibility of the proposed approach can be increased with variants at a very appropriate stage of regression model building.

There is a further need of experiments for proper measurements of both accuracy and resource efficiency to assess and optimize correctly.

## REFERENCES

- [1]. Nikita Malik, & Karan Singh (2020), Sales Prediction Model for Big Mart, Vol 3. Issue 1 (<https://msig-gsip.org>).
- [2]. Aditi Narkhede , Mitali Awari , Suvarna Gawali & Amrapal Mhaisgawali, Big Mart Sales Prediction Using Machine Learning Techniques (<http://www.ijrsred.com>)
- [3]. Kumari Punam, Rajendra Pamula, Praphula Kumar Jain (2018), A Two-Level Statistical Model for Big Mart Sales Prediction. (<https://ieeexplore.ieee.org/document/8675060>).
- [4]. Sunitha Cheriyan, Shaniba Ibrahim, Saju Mohanan, Susan Treesa (2019), Intelligent Sales Prediction Using Machine Learning Techniques (<https://ieeexplore.ieee.org/document/8659115>).
- [5]. Pavan Chatradi, Meghana, Avinash Chakravarthy V, Sai Mythri Kalavala
- [6]. Makridakis, S., Wheelwright, S.C., Hyndman, R.J.: Forecasting methods