International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

BIGMART SALES PREDICTION

[1]*Uma K R,MCA Maharaja Institute of Technology Mysore*

[2]*Prof. Amos R, Assistant Professor MCA Maharaja Institute of Technology Mysore*

Abstract –The intention is to build analytical model discovery outsales of individually product at a precisestock. Via BigMart will try to appreciate properties of product in exactstore

***Key Words*:**XG Boost Regressor.

1. INTRODUCTION

To determine present day to day sales, future sales can be predicted using machine learning concepts and algorithm. By using data like past seven days sales, season, economic factors etc. this model will predict.

on certain day after being provided sales within factors like past seven days, sales outlet, categories, outlet type, outlet location, item id, item MRP. This helps companies to develop, improve business strategies and gain proper market knowledge.

Estimating future sales in an important aspect of any business. Accurate prediction of future sales help companies to develop and improve business strategies as well as to gain proper market knowledge. Standard sales forecast helps companies to analyses the situation which has occurred before and then apply customer purchase inferences to identify inadequacies data, as well as to prepare good plan for the next year.

A detailed knowledge of past opportunities permits one to plan for future market needs and increase the possibility of success regardless of external factors, firms which see sales modeling as its first step towards improved performance, we use XG Booster approach to predict sales for best accuracy.

2. METHODLOGY

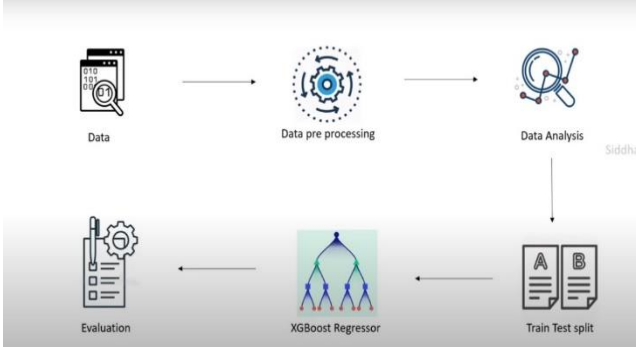

Figure: Methodology

- **Data:**Can be unprocessed facts, values, text, sound, or images that is not beaning analyzed, where in ML data is very important aspect to train the system.

- **Data Processing:** A process of preparing the raw data into to a useful information for ML model its import to process data by using any ML algorithm or mathematical

modeling or statistical knowledge, so entered process can be automated.

- **Data Analysis:** A procedure to cleaning, examining, converting then modeling data with valuable information for good decision making.

- **Train Test Split:** Used for evaluate the performance of ML algorithm that are appropriate for prediction-based application this will be in fast predicting model.

- **XG Boost Regressor:** XG Boost stands for Extreme Gradient Boosted Decision Tree machine learning library. It provides parallel tree boosting used for regression, classification model to predict good accuracy**,** an efficient implementation of gradient boosting that can be used for regression predictive modelling.

- **Evaluation:** Process of evaluating metrics to understand model performance, strength, weakness. Model evaluation is important to assess efficiency during research phase plays major role in prediction.

3. MODEL BULDING

The algorithm used in this frame work is, XG Boost approach is developed using inbuilt decision tree and an advance of gradient booting algorithm. This algorithm stands for principle of boosting algorithm placed in a gradient decent boosting framework.

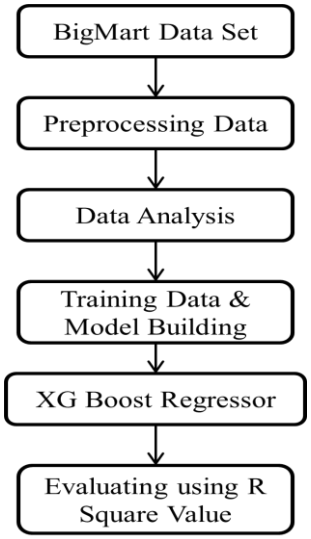

Figure: Work Flow

This approach works very accurately beating almost all other algorithm in providing accurate prediction. XG boost is portable which can run on any OS. It tents to over fitting of data uses automatic allocation; it can be defined as an extension to gradient Boosting algorithm.

Features of XG Boosting are: parallelized tree building, Effectual handling of missing factsinbuilt cross validation capability of tree pruning and cache awareness, integrated on XG Boost and with cloud, networking etc.

3.1 ADVANTAGES

- It is extremely flexible.
- Parallel process.
- Earlier advanced than Gradient boosting.
- It supports regularization.
- Handles missing data with its built-in features.

3.2 DISADVANTAGES

- XG Boost doesn't perform well on spare and unstructured data.
- The overall method is hardly scalable.
- XG Boost can't predict the future very well, but suited for task such as the following:
- Classification problems, especially those related to CNN, Deep learning this doesn't work.

4. RESULT

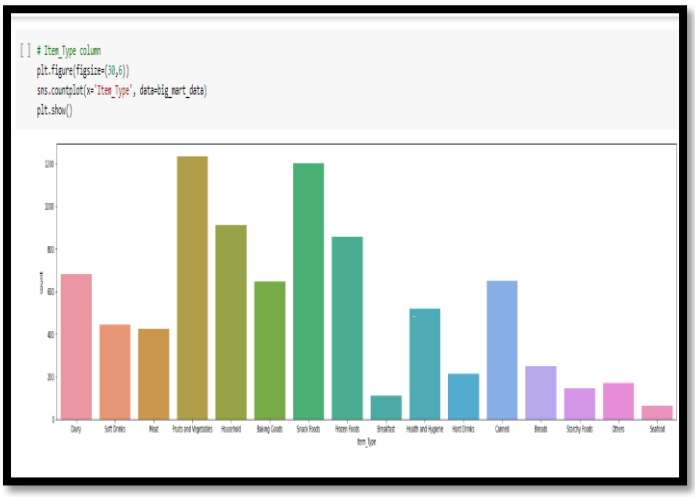

Figure: Performance Evaluation Through Bar graph

The above graph describes bar graph are plotted that represents categorical data with rectangular bar lengths proportional to values represents, a bar shows comparisons among discrete categories. From the above graph we can observe the different items or food type we have such as soft drinks, meat, fruits, vegetables etc. Hence Fruits and vegetables have high demand and Sea Food has low demand.

5. CONCLUSIONS

The objective of this framework is to predict the future sales from given data of the previous year's using learning techniques. We have discussed how different machine learning models are built using XG booster algorithms.

These algorithms have been applied to predict the final result of sales. We have addressed in detail about how the noisy data in been removed and the algorithm used to predict the result.

Based on the accuracy predicted by different models we conclude that the XG Booster approach are best models. Our predictions help big marts to refine their methodologies and strategies which in turn help to increase their profit.

6. ACKNOWLEDGEMENT

I would like to express my special regards and gratitude to Prof. Amos R, who guide and helped me in completing this survey paper. I came to know about many new things am really thankful to them.

7. REFERENCES

- "Improvising big mart sales prediction", Meghana N, Pavan Chatradi..
- " A comparative study of BigMart Sales Prediction", International Gopal Behera, Neeta.

- "Sales prediction model for BigMart", Nikita Malik, Karan Singh.
- "BigMart sales using machine learning with data analysis", Ayesha Syed, Jyothi.
- "A forecast for BigMart sales based on random forest and multiple linear regression", Heramb Kadam, Rahul Shevade, Prof. Deven Kelkar.

DataSet:

- https://www.kaggle.com/vivekhn/sales-prediction-for-big-mart-outlets

Text Book:

- Python Machine Learning – Second Edition: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow and Edition, Kindle Edition.

  Author By: Sebastian Raschka, Vahid Mirjalili. Format: Kindle Edition.

8. BIBLIOGRAPHY

Author 1:

I Uma K R born and brought up in Mysore, currently located in Kuvempunagar Mysore. I have completed MCA at Maharaja Institute of Technology Mysore. I have completed my internship at ONTRIVER Mysore. My area of interest are java, DBMS, HTML and CSS, Python, ML. My Hobbies are playing sports, listening music, occasionally reading novel.

Author 2:

My name is Amos R. I was born on 12th June 1981 at the sugar city Mandya, Karnataka. I have been brought up and been residing in the Mandya since then. I earned by Masters Degree MCA from Viveswaraya Technological University, Belgavi. I obtained a Bachelor Degree in B.Sc from University of Mysore and M.Phil from Vinayaka Mission University, Salem.

Teaching is something I have been doing my entire adult life. During these last 15 years able to blend my passions and talents to do something that has made a difference in my life and I hope, is making a difference in the lives of the next generation.

I am the 4th generation of my family involved in the noble profession of teaching. Started my carrier in the IT Industry and later got into Academic field, it's almost 15 years ever since. I am currently working as Assistant Professor in the Department of MCA at Maharaja Institute of Technology Mysore, Mysuru

More than 15 papers have been so far published in both National and International Journals.