# Forecasting of Walmart Sales Prediction Using Machine Learning Algorithms

**P.Poonguzhali[1], Dr.S.Pakkir Mohideen[2]**
[1]Dept of Computer Applications
[2]Associate Professor, Dept of Computer Applications
[1, 2] B.S Abdur Rahman Crescent Institute of Science & Technology, Vandalur,Chennai

***Abstract-*** *The paper is based on sales prediction for walmart stores by using Machine learning algorithms. Department stores like Walmart have uncountable product and money transactions every day. Because of their rapid transaction rates, keeping a balance between the inventory and customer demand is the most important decision for the managers. Therefore, making an accurate sales prediction for different products becomes an essential need for stores to optimize their profits. Most of the existing sales predictions only depend on extrapolating the statistical trend. The previous studies on sales prediction require a lot of extra information like product analysis and customer. A more simplified model is needed by the department store to predict the product sales based on only the historical sales record. The new emerging machine learning methods enable us to make more accurate predictions like that. Our data is from the Kaggle M5 Competition. A comprehensive study of sales prediction is done using machine learning algorithms such as Linear regression, Lasso Regression using sales records, price and calendar information. This project will optimize the accuracy by all means for weekly sales prediction.*

***Keywords***- sales records, linear regression, lasso regression, weekly prediction.

## I. INTRODUCTION

Sales forecasting has been a significant area to concentrate upon. An optimal and efficient way of forecasting has become indispensable for all the vendors and to sustain the efficiency of the marketing organizations. Manual influx of the task could lead to drastic errors leading to deprived management of the organization and importantly time consuming. A main part of the universal budget relies upon the business sectors, which are literally expected to produce fitting quantities of products to meet the overall needs. The major focus of business sectors is targeting the market audience. The process of prediction involves analysing data from various sources such as consumer behaviour and market trends. These prediction process including predicting how much sold in a given periods and predicting the future demand. Machine learning is domain where the machines increase the ability to performs human in specific tasks. They used to do specific task in a logical way and gain improved results for the progress of the existing society. Sales forecasting also machine learning has showed to be a boon and it helpful in predicting the future sales with accuracy. In our paper we have proposed the machine learning algorithms towards the historical data collected from various Walmart stores. The objective for this project is to estimate precisely for the product unit sales forecasting in the Walmart sales. To perform predictions on several products that are sold in Walmart and machine learning algorithms have been implemented along with the methods to increase the precision. Three different machine learning models are used to forecast weekly sales for following a 28-day period. The metric of evaluating the models is Root Mean Square Error (RMSE). The outcome from RMSE could support the stated hypothesis and assist the business analyst to improve planning on different aspects of the business level, for instance inventory distribution, distribution management, inventory storage solutions, product fulfilment, etc. The problem is misleading business forecasts on product sales could potentially cause opportunity and revenue loss for Walmart.

### *A. Methodology*

#### 1. Linear regression:

Linear regression is one the most popular and easiest machine learning algorithms. It is a statistical based method is used for predictive analysis and its makes predictions for numeric or continuous variables. This algorithm shows an linear relationship between independent and dependent variables. This regression provides a sloped straight line representing the relationship between the variables. Direct relapse may be a straight show, Eg, a show that accept a straight relationship between the input factors (x) and the single yield variable (y). More particularly, that y can be calculated from a direct combination of the input variables (x). In measurements, straight relapse may be a straight approach to modeling the relationship between a scalar reaction and one or more informative variables. Linear Relapse is the foremost commonly and broadly utilized calculation Machine Learning

calculation. It is utilized for setting up a direct connection between the target or subordinate variable and the reaction or autonomous factors. The direct relapse demonstrate is based upon the taking after condition: y θ0 + θ1x1 + θ2x2 + θ3x3 +……+ θnxn (1) where, y is the target variable, θ0 is the caught, x1,x2,x3,….,xn are autonomous factors and θ1, θ2, θ3,……, θn are their individual coefficients. The most point of this algorithm is to discover the finest fit line to the target variable and the independent factors of the information. It is accomplished by finding the foremost ideal values for all θ. With best fit it is implied that the anticipated esteem ought to be exceptionally near to the real values and have least mistake. Mistake is the separate between the information focuses to the fitted relapse line and for the most part can be calculated by utilizing the taking after condition: Blunder = y - y , where, y is the genuine esteem and y is the anticipated value

**2. Lasso regression:**

LASSO stands for Least Absolute and Selection Operator and it is regularization technique to reduce the difficulty of the model. It is similar to the ridge regression which a penalty term contains only the absolute weights instead of a square of weights, It is called L1 regularization. Lasso regression is technique to reduce the complexity of the models. It takes absolute values when it can shrink the slope to 0, Ridge regressions only shrink it near to 0. The equation for the cost purpose of lasso regression. Features in this technique are totally neglected for model evaluation. Regression which helps us to reduce the overfitting in the model as well as feature selection.

## II. EXISTING SYTEM

Sales forecasting is usually done by collecting the sales data of a shop of a time period and make predictions using various prediction techniques. There are many factors which affects the sales forecasting which includes direct and indirect competition, state and city holidays, population changes, sales promotions etc. The above factors create a great deviation in sales prediction in existing system which is not providing accurate results as expected. The confidence level has not taken for all algorithms. The holiday factors which is important in sales prediction is no considered. Thus, the sales various on using different machine learning algorithms. Keras algorithm is used for sales prediction of Wall Mart stores from the historical sales data for 45 stores located in different regions. Each store contains many departments and participants must project the sales for each department in each store.

**Drawback:**

- Very low in result accuracy for sales prediction.
- Time-consuming and unpleasant experience.

## III. PROPOSED SYSTEM

In proposed Framework Walmart deals forecast is anticipated with the machine learning calculations towards the information collected from the past deals of a basic supply store. The objective here is to imagine the design of deals and the amounts of the items to be sold based on a few key highlights accumulated from the crude information we have. In case of deals estimating totally foreseeing the deals level, stock and offers data are get known. It is supportive in anticipating long term deals more precisely. The capacity to anticipate information precisely is amazingly profitable in a endless cluster of spaces such as stocks, deals, climate or indeed sports. Displayed here is the consider and usage of a few gathering classification calculations utilized on deals information, comprising of week after week retail deals numbers from diverse divisions in Walmart retail outlets all over the Joined together States of America. The hyper parameters of each show were changed to get the finest Cruel Supreme Mistake (MAE) esteem and R2 score. The number of estimators hyper parameter, which specifies the number of choice trees utilized within the demonstrate, plays a especially critical part within the assessment of the MAE esteem and R2 score and is managed with in an attentive way. A comparative examination of the three calculations is performed to demonstrate the most excellent calculation and the hyper parameter values at which the leading comes about are gotten.

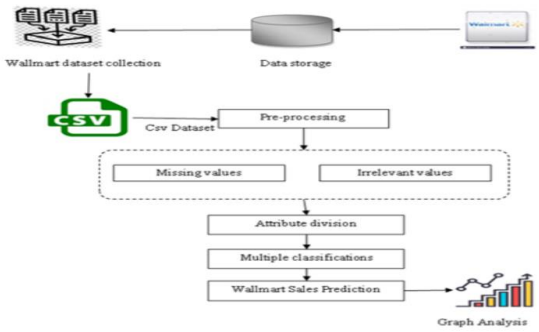


Fig. 2 Overall Architecture Diagram

## IV. LITERATURE SURVEY

Social media have changed the world in which we lives proposed by Tesco and Walmart [1]. In spite of the fact that a few considers have revealed shapes of client engagement on social media, there's a shortage of scholarly inquire about on client engagement inside the basic need segment. This consider hence points to address this hole within the writing and shed light on the different ways clients

lock in with basic supply stores on Facebook. Netnography is utilized to pick up an understanding of the conduct of clients on the Facebook page of Tesco and Walmart. The discoveries of this ponder uncover that cognitive, emotional and behavioural client engagement are showed which clients can both make and annihilate esteem for the firm.

The point of this paper [10] is to analyse the deals of the enormous superstore, and foresee their future deals for making a difference them to extend their benefits and make their brand indeed way better and competitive as per the advertise patterns by creating client fulfilment as well. The method utilized for forecast of deals is the Linear Regression Algorithm, which may be a popular calculation within the field of Machine Learning. The deals information is from the year 2011-13 and forecast of data for the year 2014 is done. At that point, real-time information of the year 2014 is additionally taken and the actual data of the year 2014 has been compared to the anticipated information to calculate the precision of expectation. This can be done so as to approve our comes about with the genuine ones. This in turn would offer assistance them take fundamental activities.

## IV. MODULES

**1. Dataset collection**:

The dataset comes from the Kaggle stage and comprises of information from an American retail organization, Walmart Inc. The dataset was utilized for a machine learning competition in 2014. It comprises information from 45 Walmart division stores primarily cantered around their deals on a week after week premise. The dataset has 282,452 passages that will be utilized for preparing the models. Each section has properties as takes after: the related store (recorded as a number), the comparing division (81 offices, each entered as a number), the date of the beginning day in that week, departmental week after week deals, the store measure, and a Boolean esteem indicating on the off chance that there's a major occasion within the week. The major occasions being one of Thanksgiving, Labour Day, Christmas or Easter. Together with the previously mentioned qualities may be a parallel set of highlights for each section counting Customer Cost List, unemployment rate, temperature, fuel cost, and special markdowns.

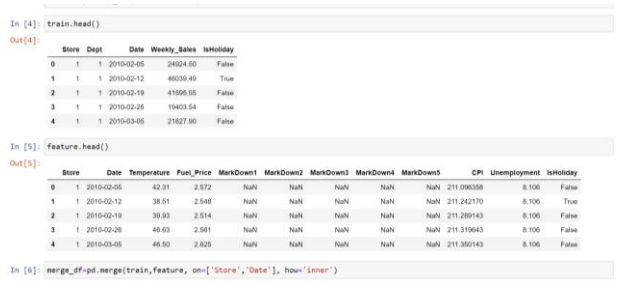


Fig. 4.1.1 Datasets

In Fig4.1.1 the image dataset for the normal is imported from the folder with the help of data generator function and classified into Store, dept, Weekly sales, Is holidays.. Here the imported testing image dataset contains about 465 data's.

**2. Data Pre-processing:**

Information pre-processing could be a portion of information mining, which includes changing crude information into a more coherent organize. Crude information is more often than not, conflicting or fragmented and ordinarily contains numerous mistakes. The information pre-processing involves checking out for lost values, seeking out for categorical values, part the dataset into preparing and test set and finally do a highlight scaling to constrain the range o factors. Data pre-processing may be an information mining method which is utilized to convert the crude information in a valuable and effective format. Steps Included in Information Pre-processing: Data cleaning

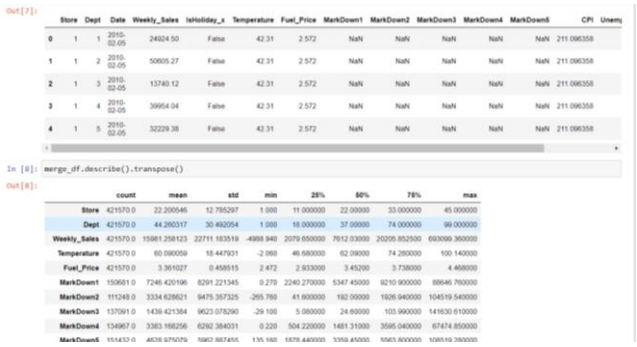


Fig. 4.2 Dataset Pre-processing

**3. Data Acquisition:**

Information Securing comprises of two words: Information alludes to the crude actualities, figures, or piece of realities, or measurements collected for reference or examination. Securing indicates to securing information for the scheme. There are four strategies of securing information: collecting modern information; converting/transforming bequest information; sharing/exchanging information; and

obtaining information. Retailer's to begin with need is ordinarily to get it their clients to be able to fulfill their needs so that these clients will return to the store for future needs, in this way expanding the item requests and including to the trade esteem. These businesses need this data to arrange where and when contribute beneficially.

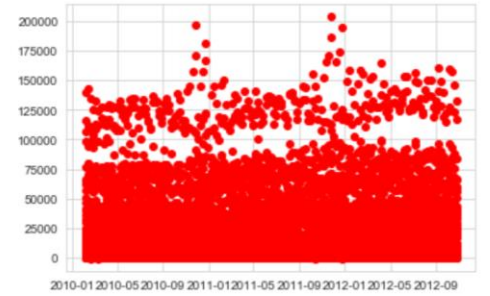


Fig. 4.3.1 Data Analysis for Weekly sales

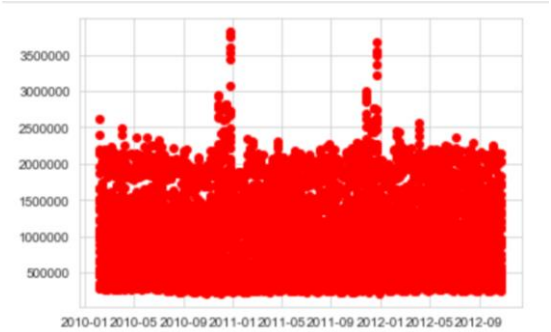


Fig. 4.3.2 Weekly sales for particular Datatime

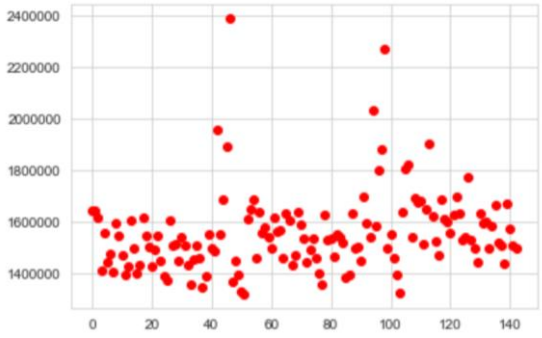


Fig. 4.3.3 Analysis for Trained dataset

**4. Model Fitting**:

Model fitting is a measure well a machine learning model which generalize to similar data to that on which it was trained. A model that is well-fitted produces more accurate outcomes.

**4. Evaluation**:

Firstly, we compare the performance of different classifiers to each other we perform cross-validation and used performance measures such as Precision, Recall, F measure, and Accuracy to determine the effectiveness of the proposed approach. Models performed best overall to achieve 90% accuracy.

## V. RESULT

The result that we are getting from our implemented algorithm provides better results than previous algorithm with reference to great accuracy and also the implemented algorithm provides stronger results when compared with others existing algorithm.

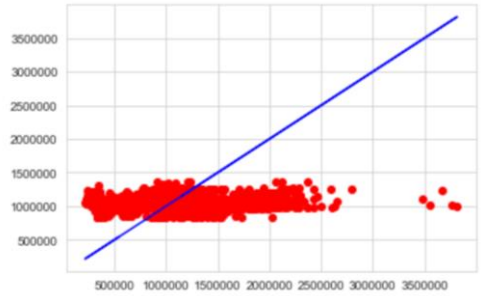


Fig. 5.1 Test Accuracy in Linear Regression

In Fig5.1 shows the test accuracy in Linear regression and it also shows the graph for Model accuracy and Model loss that is the error of the model

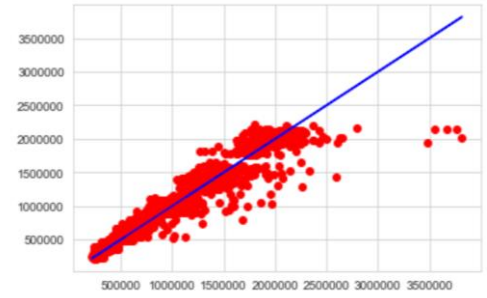


Fig. 5.2 Test Accuracy in Linear Regression with Training dataset

In Fig5.2 shows the test accuracy in Linear Regression with training dataset and it also shows the graph for Model accuracy.

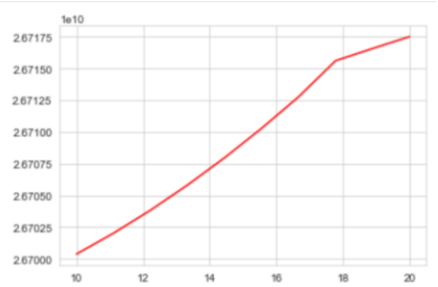


Fig. 5.3 Test Accuracy in Lasso Regression

In Fig5.3 shows the test accuracy in Lasso Regression and it also shows the graph for Model accuracy.

## VI. CONCLUSION AND FUTURE ENHANCEMENT

Machine Learning calculations such as Logistic Regression and Lasso Regression calculation have been utilized to anticipate the deals of different outlets of the Enormous Bazaar. Different parameters such as Root Cruel Squared Blunder (RMSE), Fluctuation Score, Preparing and Testing Exactness's which decided. The winning accommodation for the Kaggle competition had a Cruel Supreme Blunder (MAE) of around 2130. As a reference, a accommodation where all the anticipated values of week by week deals are 0's, the MAE is found to be roughly 21000. In this ponder, the final 20% of the preparing dataset was utilized as the neighbourhood test-set. The Angle Boosting calculation was taken as a standard and the MAE was found to be 5771.5, with a R2 score of 0.80 that suggests that 80% of the anticipated values were exact. These were the most excellent comes about gotten with the n_estimatorshyperparameter, which alludes to the number of choice trees that are utilized for relapse, set at 200. The other hyperparameters were set to their default values and the results in greater accuracy.

## REFERENCES

[1] Baba, Norio, and HidetsuguSuto. "Utilization of artificial neural networks and GAs for constructing an intelligent sales prediction system."In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks.IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, vol. 6, pp. 565-570. IEEE, 2000.

[2] Cheriyan, Sunitha, Shaniba Ibrahim, SajuMohanan, and Susan Treesa. "Intelligent Sales Prediction Using Machine Learning Techniques."In 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE), pp. 53-58.IEEE, 2018.

[3] Fawcett, Tom, and Foster J. Provost. "Combining Data Mining and Machine Learning for Effective User Profiling."In KDD, pp. 8-13. 1996.

[4] Friedman, Jerome H. "Stochastic gradient boosting." Computational Statistics & Data Analysis 38.4 (2002): 367-378.

[5] Giering, Michael. "Retail sales prediction and item recommendations using customer demographics at store level." ACM SIGKDD Explorations Newsletter 10, no. 2 (2008): 84-89.

[6] Kulkarni, Vrushali Y., and Pradeep K. Sinha. "Random forest classifiers: a survey and future research directions." Int J AdvComput 36.1 (2013): 1144-53.

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

[8] Ragg, Thomas, Wolfram Menzel, Walter Baum, and Michael Wigbers. "Bayesian learning for sales rate prediction for thousands of retailers."Neurocomputing 43, no. 1-4 (2002): 127-144.

[9] Sekban, Judi. "Applying machine learning algorithms in sales prediction." (2019).

[10] Singh Manpreet, BhawickGhutla, Reuben Lilo Jnr, Aesaan FS 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.