

Buyer's Prediction For Mall Outlets

¹Shraddha Dhuri, ²Rahul Yadav, ³Niranjan Salunke, ⁴Deepali Shrikhande
¹Student, ²Student, ³Student, ⁴Professor
Vidyalankar Institute of Technology

Abstract - The study of consumer behavior helps everyone as all are consumers. It is essential for marketers to understand consumers to survive and succeed in this competing marketing environment. When a consumer has an unfulfilled need, the buying process begins to fulfill the needs. The need may be activated by internal or external factors. The intensity of the need will indicate the speed with which a person will move to fulfill the want. On the basis of need and its importance, forms the order of priority. Marketers should give required information about selling points. Customer Relationship Management systems are used to allow organizations to gain new customers, establish a continuous relationship with them and increase customer confiscation for more profitability. Now-a-days, consumers have become more interested in the quality of service (QOS) that organizations can deliver. Services provided by various vendors are not highly distinguished which increases competition between vendors to maintain and increase their service. CRM systems use machine-learning to analyze customer's individual and behavioral data to give an organization a competing advantage by increasing customer confiscation rate. Those models can predict consumers who are awaited to churn and reasons of churn. Predictions are used in making targeted marketing goals and service offers.

keywords - Machine Learning, Prediction, Classification, Data Mining, Data Extraction.

I. INTRODUCTION

India is one of the arising markets in the global economy. Since liberalization, India has witnessed growth at an unconventional rate. With modifications in almost all the sectors, India has seen growth in infrastructure, capital markets, banking, insurance, etc. This upgrade has given rise to a new section in the country in the form of retail industry. With growth of industry, the employment levels have increased and that helps to increase disposable earnings of the common customer. Each decade calls upon company management to think about its objectives, strategies, and delicacy. Rapid changes can easily expire yesterday's used strategies for conducting business. Marketing is one of the problems that are facing by current businesses. Marketing is a social and managerial process by which personal and groups obtain what they need and want through making, offering, and exchanging deliverables of value with others. Marketing intelligence starts with the fact of human essentials and wants. There is no doubt that people's essentials and wants are staggering. They satisfy their essentials and wants with goods and services i.e. product. A product is everything that can be offered to satisfy a need or want. Manufacturers often make the mistake by paying extra attention to their main products than to the services produced by these products. They see themselves as selling a product instead of giving an answer to a requirement. The gradual increase in GDP and the buying power of citizens provides an excellent opportunity for planned retailing. The fastidious pace of growth of the Indian economy is the drive for Indian consumerism. Projections by analysts suggest that the country has the potential to be labeled the fastest-growing economy and outpace the developed economies by 2050. India presents a big market with its young citizens just starting to embrace indicative lifestyle changes. The demographic and economic data widely quoted are undoubtedly attractive but so as to assess the true nature of retail openings, we'd like to know the high transformation that is occurring in Indian consumer behavior thanks to varying lifestyles, uprising aspirations and thus the aspect of a changing subculture. Against the effects of accelerating modern retail changes, India offers to be a well-favored place for global corporations and leading retailers seeking coming up markets overseas. Retailing here is receiving global recognition and attention and this emerging market is witnessing a significant change in its growth and investment pattern.

II. LITERATURE SURVEY

Research [1] indicates that the cost of holding a customer is less than attracting new ones. This is due to marketing costs required to call to new customers. For this reason, together with the increase of competition it has become crucial that the current customers base is owned. Normally, customers churn gradually and not abruptly. This means that by analyzing customers' history buying patterns one can adopt a careful approach in predicting churn. Since all transactions are inserted through POS and stored in databases, understanding customers' needs and patterns is possible as data is accessible.

According to [2], execs are dedicating marketing budgets to focus on customer confiscation campaigns. Various models used to predict churn specialize in analytics and renowned machine learning algorithms including Random Forest and Logistic Regression. This paper focuses on 2 aspects when predicting churn within the grocery retail industry. The first is the features which will be passed on to the model. Instead of using consumer buying trends to cluster the individual subject, these values will be created as features and are handed over to the model. Therefore, for every customer different features are created to permit the model to find out and identify patterns per individual. For this reason, 2 data sets are created to test and evaluate how data should be illustrated to predict churn. The second aspect is the implementation of the algorithm. The originality of this study is the use of deep learning to predict churn within the grocery industry. To our knowledge, this is often the primal study which implements deep learning within this industry. The power of using deep learning is that it can disclose hidden patterns within the available data set. An important aspect within the business is to have a good understanding of customers' needs, whereby

comprehensive views of their patterns may be analyzed. When customers are satisfied with the service or products, customer faith increases [3].

Authors [4], further discuss that earnings and margins will increase if customers are valued when compared to the cost of attracting new customers. Applying statistical techniques and machine learning algorithms on available data sets may guide companies in identifying hidden trends (tendency) and customer buying patterns. Implementing data processing methods to predict churn may give companies a competing edge up improves the connection with customers. Using customer churn models which appropriately classify churn, companies have added worth. Churn is a term used in the marketing field to signify that a customer has moved to a competitor or has stopped transacting. Churn may be defined as customers who have a high probability to stop transacting with the organization [1] or as described by [6]: churn may be identified when a customer's purchasing value falls below a threshold across a predefined period of your time. Within the Grocery Retail Industry, the identification of the delicate moment a customer will churn is hard to define.

The output of this layer is sent to a max pooling layer [7] discusses that this layer creates a smaller and compact feature map of the input. Next, is a fully connected layer whereby all nodes from the max pooling layer are connected to the neurons within this layer. The number of fully connected layers differs on the depth of the data.

Yuta Kaneko et. Al [8] proposed A Deep Learning Approach for the Prediction of Retail Store Sales system used deep learning to construct a model that predicts the increment and decrement in the sales of a retail store and tested its practicability. System used three years of POS data from supermarkets for the analysis, treating 29 months of its data for learning, while the remaining 7 Months of Information were used for verification. As a result, the predictive accuracy of the increment or decrement in sales for the following day

varied between 75% to 86% according to the changes in the number of product attributes. The predictive accuracy was topmost when the model was designed using the Category 1 data, which consisted of 62 attributes.

Yi Zuo et. Al. [9] proposed Prediction of Consumer Purchasing during a department store Using ML Techniques. System employs 2 representative machine learning methods: Bayes classifier and support vector machine (SVM) and investigates the performance of them with the data in the real world. It is a method for extracting consumer behavior. Utilizing RFID data acquired for individuals in Japanese supermarket, we tested several important methodological issues related to the use of RFID data in support vector machines (SVMs) to predict purchasing behaviour.

Liu Bing and Shi Yuliang [10] Prediction of buyer's Purchase Intention supported ML. System present stage, a naive Bayesian algorithm has the advantage of straightforward implementation and high classification efficiency. However, this method is dependent on (Allotment) Distribution of samples in the sample space, and has the potential of unstableness. To this end, the decision tree method is

introduced to deal with problems of interest classification, and the innovative use of Local storage technology in HTML5 to obtain the required experimental data. Classification method uses the information (data) entropy of the training data set to construct the classification model, through the simple search of the classification model to complete the classification of unknown data items.

Jinggui Liao, Yuelong and Saiqin Long have proposed the MRPrePost parallel algorithm adapted for mining big data [11]. It is a parallel calculation which is actualized utilizing the Hadoop stage. The MRPrePost is an enhanced Pre-Post calculation which utilizes the map's reduced structure. The MRPrePost calculation is employed to get the affiliation runs by mining the vast datasets. The MRPrePost calculation has three stages. In the initial step the database is isolated into the information squares called the shards which are allotted to every specialist hub. In the second step the FP-tree is developed. In the last advance the FP-tree is mined to acquire the successive item sets. Trials have demonstrated that the MRPrePost calculation is the quickest.

Sheela G and Bharat T Frequent item set Mining for bulky Data in social media using the ClustBig FIM algorithm [12]. essential datasets are mined using the Map-reduce framework in the proposed opinion. Enormous FIM figuring is modified in accordance with the ClustBig FIM computation. ClustBig FIM computation gives all-round and speed which are wont to get obliging information from far-reaching datasets. The profitable data can be used to settle on better decisions in the business development. The proposed ClustBig FIM estimation has 4 essential steps. In the underlying advance the proposed computation uses K-infers count to make the bundles. In the second step the relentless item sets are mined from the gatherings. By building up the prefix tree the general TID summary is obtained. The sub trees of the prefix tree are mined to get the standard item sets. The proposed ClustBig FIM figuring is aroused being more successful appeared differently in reference to the large FIM calculation.

R. Priyanka and S. P. Siddique I. proposed A Survey on Infrequent Weighted Item-set Mining Approaches [13]. This paper handles the issues of finding the rare and weighted item set. The periodic item-set mining issue is discovered item sets whose repeat of the data isn't precisely or equal to the most outrageous edge. This paper audits a distinctive procedure for mining rare item sets. Finally, a relative strategy for each method is shown. Data Mining is described as Extraction interesting illustrations or gaining from a colossal measure of

data. Data burrowing is the procedure for exposing data from different points of view and gathering into accommodating information. Finding of typical illustrations concealed during a database expects a pivotal part during a few data processing tasks. There are two sorts of models in data mining. One is an insightful model which uses data which uses data with known results to develop a model that can use explicitly to expect esteems. Another is an obvious model, which depicts the case in existing data. Course of action is a model or classifier created to anticipate class names. It made out of two phases: coordinated learning of a planning set of data to make a model,

and after that gathering the data as demonstrated by the model. It relies upon a perceptive model. Backslide examination is a quantifiable theory that is often used for numerical desire.

III. FEASIBILITY STUDY

System Feasibility :

The very first phase in any system development life cycle is early justification. In the preliminary investigations we examine the project viability, the possibility of the system being useful to the automation process. This project has been tested in the following areas of feasibility:

- **Operational Feasibility :**

It was decided that the proposed project could be created as a web application that will meet the operating environment of various applicants from all over the place. The reasons for this conclusion are:

- Business method adopted is acceptable to all users.
- The end user have been involved in the planning and development.
- Manual errors will be reduced.
- It is a user-friendly interface and ensures interoperability.
- It is operationally feasible project considering both the hardware and software factors.

- **Technical Feasibility :**

It is decided that project is technically feasible because of the following reasons:

- Necessary technology exists to do what is suggested.
- The website is available freely and completely customizable and flexible.
- The system could be expandable and enhanced if so decided

- **Economic Feasibility :**

As a part of it, expenses and outcomes associated with the proposed system are compared and the project is economically feasible only if palpable or intangible advantages outweigh the costs. The system development costs will be eloquent. So the proposed system is economically feasible.

:

IV. METHODOLOGY

- **Problem Statement :**

In the previous system of malls we only get to know about the sales of any particular shop located in the mall. But in that system, owners didn't get an idea about the buyer's whether they would purchase anything or not. To overcome this drawback we enforced new technology which will come to know about the visitor who entered the mall will purchase anything from any shop/food court/game zone inside that mall.

- **Scope :**

Customer's behavior has been observed by numerous scientists consisting psychologists, sociologists, economists, behavior analysts and anthropologists. These all scientists are exercising to know the consumer deciding process, both personal and in groups. It studies characteristics of every consumer like demographics, and changing variables in an effort to know people's wants. It also tries to assess strike on the buyer from groups like family, friends, reference groups, and society generally. This project not only gives the sales prediction but also tells that will that customer will that customer purchase anything or not. This makes the project more reliable.

- **Aim and Objectives :**

A. **Keep track of customer :**

Once a customer registers himself in the database. We can keep track of customers every time when customer entered mall data is saved in the database.

B. **Modification of data :**

Users can modify the database according to their needs. Users can change or insert data from the registration desk. When he visits again the record will be updated.

C. **Restrict unknown visits :**

For security reasons when someone who is not in database visits he will need to register first. This data will be stored in a database for future visits.

D. **Security :**

We are storing customer information. Authentication is done every time when customers visit the mall. Authentication is necessary without authentication customer data is not entered in the database.

E. **Centralized data :**

Data is stored centralized due to which it is easily excessive for all shops for generating bills and maintaining history.

V. PROPOSED SYSTEM

One of the most common financial decisions that each of us makes on a nearly daily basis involves the purchasing of various products, goods, and services. In some cases the decision on whether or not to make a purchase is based largely on price but in many instances the purchasing decision is more complex, with many more considerations affecting the decision-making process before the ultimate commitment is formed. Retailers understand this well and attempt to make use of it in an effort to gain an edge in a highly competitive market. Specifically, in an effort to make purchasing more likely, in addition to balancing the salability and profit in setting the selling price of a product, companies frequently introduce additional elements to the offer which are aimed at increasing the perceived value of the purchase to the consumer. The system which we are going to make contains a web application, local server, database which will help you to predict the buyers prediction. How many times a customer has purchased/shopped depending on prediction history in database shopkeeper will give an offer/discount to a customer.

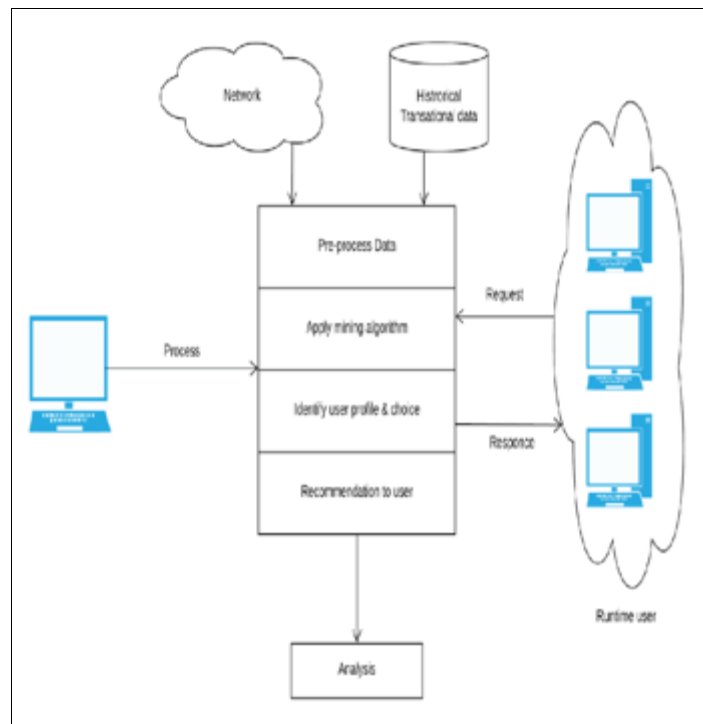


Fig 1 : Block Diagram

Following are the features provided by the system:

- Visitor first enters the mall .After entering he will go to registration desk if he is a new user he will first register herself by entering all credentials in the registration form
- If a visitor is a registered user he will then authenticate herself .
- After registration visitor information will be stored in the database. Visitors then start shopping /window shopping in the mall.
- Prediction analysis will be performed in the database. Depending upon the visitor's prediction, the shopkeeper will give an offer/discount to the customer to attract him. Due to which the shopkeeper will also get benefits to increase the sales.

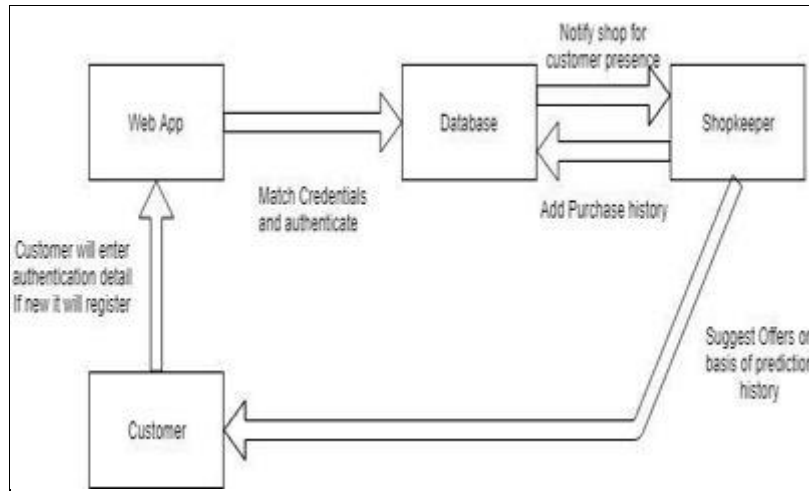


Fig 2 : Flow Diagram

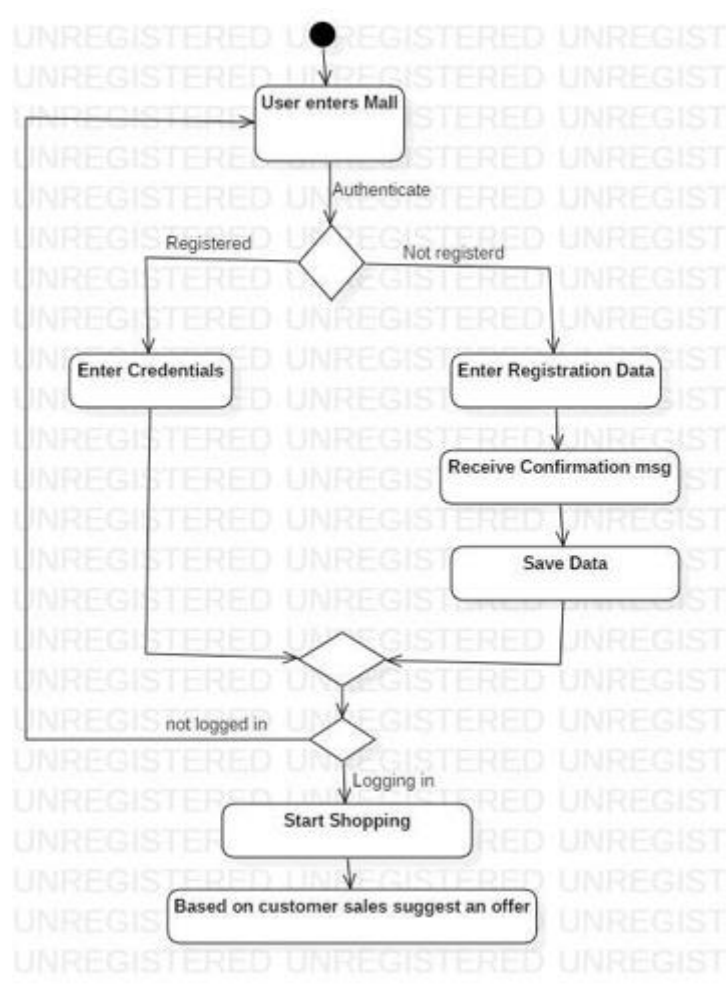


Fig 3 : Activity Diagram

VI. RESULTS

Entry Page :



Fig 4 : Entry Page



Fig 5 : Prediction Result

Login Page :

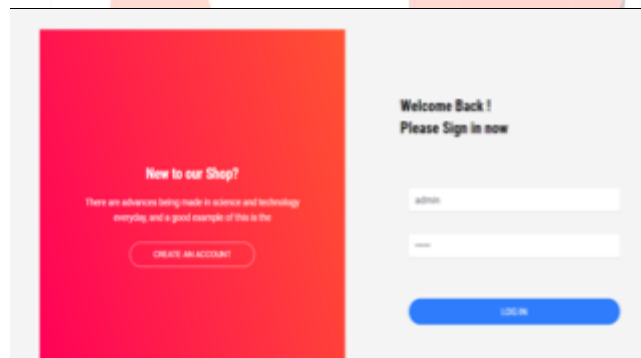


Fig 6 : Login Page

Admin Options:



Fig 7 : Manage Customer

Shop No.	Name	Contact No.	Email ID	Description	SR	Delete
1	Zara	9854171556	zara@gmail.com	Men & Women fashion Wears	SR	Delete
2	Lifestyle	9632587412	lifestyle@gmail.com	All kinds of fashion wears	SR	Delete
3	Zivame	9632587690	Zivame@gmail.com	Ladies Wears	SR	Delete
4	Andrew Collections	886471338	acollections@gmail.com	Gents and Ladies Collections	SR	Delete
5	Checkpoint	8768770660	checkpointcollection@gmail.com	SR shop	SR	Delete

Fig 8 : Manage Shops



Fig 9 : View Sales

Shops Options:

Offer ID	Category	Name	Offer Price	Description	Start Date	To Date	Image	SR	Delete
1	Shoes	get off	100	get off 10%	3/23/2020 10:00:00 AM	6/5/2020 12:00:00 AM		SR	Delete
2	Shoes	50% flat	200	get off 50%	18/06/2019 10:00:00 AM	1/05/2020 12:00:00 AM		SR	Delete
3	Shoes	sale	100	buy one get one	9/30/2020 12:00:00 AM	9/30/2020 12:00:00 AM		SR	Delete
4	Shoes	week promo	1000	week off	3/18/2020 12:00:00 AM	6/18/2020 12:00:00 AM		SR	Delete

Fig 10 : Manage Offer

User ID	Name	Email ID	Contact No.
1	Shradha Dhuri	shradhadhuri@gmail.com	9632587412
3	Ajay Rathod	ajay@gmail.com	794561230
4	Morgan Sakurba	morgan72@gmail.com	828664729
5	Rahul Yadav	yadavrahul734@gmail.com	9619613962
6	Deepan Paul	npaul72@gmail.com	8149682122

Fig 11 : View Customer

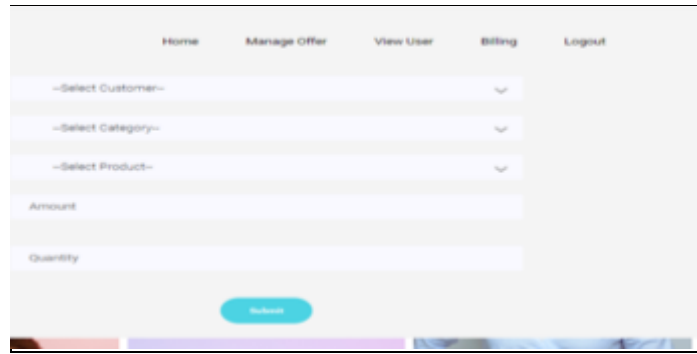


Fig 12 : Billing

VII. CONCLUSION

We will do research on the following area of advice also as pricing. We will attempt to consider both user and provider concerns of adjusting demand and its cost. This will ensure both provider and customers benefit. Apart from this we'll consider competitive prices and its result on pricing. We will study best fit auction based pricing to support optimized fine grained schemes. Also a partial waste issue may be an area of study which is going to be brainstormed to seek out an efficient solution.

VIII. REFERENCES

- [1] C. Jie, Y. Xiaobing, and Z. Zhifei, "Integrating OWA and data mining for analyzing customers churn in e-commerce," The Editorial Office of JSSC and Springer- Verlag Berlin Heidelberg, vol. 28, pp. 381-391 2015.
- [2] The science behind customer churn. [Online]. Available:<http://financeinbusinesslife.info/the-sciencebehind-customer-churn/>
- [3] S. Neslin, S. Gupta, W. Kamakura, L. Junxiang, and C. H. Mason, "Defection detection: Measuring and understanding the predictive accuracy of customer churn models," *Journal of Marketing Research*, vol. 43, pp. 204-211, 2006.
- [4] I[Online]Available:<https://www.ibm.com/developerworks/library/badata-mining-techniques/>, Developer works, Accessed: November 2016
- [5] E. Siegel, *Predictive Analytics, The power to Predict Who Will Click, Buy, Lie or Die*, Wiley, 2013.
- [6] V. Migueis, D. V. den Poel, A. Camanho, and J. Falcao, "Modelling partial customer churn: On the value of first product-category purchase sequences," *Expert Systems with Applications*, pp. 11250-11256, 2012.
- [7] M. Nielson, *Neural Networks and Deep Learning*, Online Book, 2016.
- [8] Zuo Y, Yada K, Ali AS. Prediction of Consumer Purchasing in a Grocery Store Using Machine Learning Techniques. In *Computer Science and Engineering (APWC on CSE), 2016 3rd Asia-Pacific World Congress on 2016 Dec 5* (pp. 18-25). IEEE.
- [9] Kaneko Y, Yada K. A deep learning approach for the prediction of retail store sales. In *Data Mining Workshops (ICDMW), 2016 IEEE 16th International Conference on 2016 Dec 12* (pp. 531-537). IEEE.
- [10] Bing L, Yuliang S. Prediction of User's Purchase Intention Based on Machine Learning. In *Soft Computing & Machine Intelligence (SCMI), 2016 3rd International Conference on 2016 Nov 23* (pp. 99-103). IEEE.
- [11] Jinggui Liaqat. Al. proposed A parallel algorithm adapted for mining big data I IEEE Workshop on Electronics and Computer Applications in year 2014
- [12] Sheela Gole and Bharat Tidke, proposed a system Frequent Item set Mining for Big Data in social media using the ClustBigFIM algorithm, International Conference on Pervasive Computing in 2012. *International Research Journal of Engineering and Technology (IRJET)* e-ISSN: 2395-0056 Volume: 05 Issue: 10 | Oct 2018 www.irjet.net p-ISSN: 2395-0072 © 2018, IRJET | Impact Factor value: 7.211 | ISO 9001:2008 Certified Journal | Page 935
- [13] Siddique Ibrahim at. Al. A Survey on Infrequent Weighted Itemset Mining Approaches, *IJAR CET, Vol.4*, pp. 199-203 in 2015.
- [14] Surendar Natarajan and Sountharajan Sehar proposed the idea of Distributed FP-ARMH Algorithm in Hadoop Map Reduce Framework for IEEE, in 2013.
- [15] Xiaoting Wei at. Al. define the, Incremental FP-Growth Mining Strategy for Dynamic Threshold Value and Database Based on Map reduce proposed a idea *Proceedings of the 18th IEEE International Conference on Computer Supported Cooperative Work in Design* in 2014