A MINOR PROJECT REPORT ON

**“ONLINE USER BEHAVIOUR ANALYSIS”**

**Submitted**

*In the partial fulfilment of the requirements for*

*The award of the degree of*

**BACHELOR OF TECHNOLOGY**

In

**COMPUTER SCIENCE & ENGINEERING**

By

G. Manasa (171FA04264)

Sk. Afrid (171FA04236)

Ch. Charitha (171FA04243)

Under the esteemed guidance of

Mr. S.V.V. Phani Kumar, Associate Professor

****

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**VIGNAN'S FOUNDATION FOR SCIENCE, TECHNOLOGY AND RESEARCH**

(**Accredited by NAAC “A” grade**)

**Vadlamudi, Guntur.**

**VIGNAN’S FOUNDATION FOR SCIENCE TECHNOLOGY AND RESEARCH**

(Accredited by NAAC “A” grade)

****

**CERTIFICATE**

This is to certify that the Minor project Report entitled “**ONLINE USER BEHAVIOUR ANALYSIS**” that is being submitted by G. Manasa (171FA04264), Sk. Afrid (171FA04236) and Ch. Charitha (171FA04243) in partial fulfilment for the award of B.Tech degree in Computer Science and Engineering to the Vignan’s Foundation for Science, Technology and Research, Deemed to be University, is a record of bonafide work carried out by them under my supervision.

Mr. S.V.V. Phani Kumar External Examiner Dr. D. Venkatesulu

Associate Professor HOD, CSE

## ACKNOWLEDGEMENT

We are very grateful to our beloved Chairman **Dr. Lavu Rathaiah**, and Vice Chairman **Mr. Lavu Krishna Devarayalu ,** for their love and care.

It is our pleasure to extend our sincere thanks to Vice-Chancellor **Dr. M.Y.S. Prasad** and Dean Engineering & Management, **Dr. M. Rukmini,** for providing an opportunity to do my academics in VFSTR.

It is a great pleasure for me to express my sincere thanks to **Dr. D. Venkatesulu HOD, CSE** of VFSTR**,** for providing me an opportunity to do my Mini Project.

We feel it our responsibility to thank **Mr. S.V.V. Phani Kumar** under whose valuable guidance that the project came out successfully after each stage.

We extend our whole hearted gratitude to all our faculty members of Department of Computer Science and Engineering who helped us in our academics throughout course.

Finally we wish to express thanks to our family members for the love and affection overseas and forbearance and cheerful depositions, which are vital for sustaining effort, required for completing this work.

With Sincere regards,

|  |  |
| --- | --- |
| G. Manasa | (171FA04264) |
| Sk. Afrid | (171FA04236) |
| Ch. Charitha | (171FA04243) |

**ABSTRACT**

Consumer Behaviour is a complex and challenging field to analyze by the marketer as preferences vary over a period of time. The traditional method of purchase is replaced with online model that facilitating the consumer anytime to purchase the goods by providing all the benefits under a single roof. Various E-Commerce models provide both the product and service sectors to utilize the facilities and opportunities at the right time. The study ononline user behaviour analysis will mainly focus upon the user purchasing or shopping behaviour via internet. It will deal with the transactions of the user and will try to find out whether there exists any association between the several items present in a transaction. If any strong association is observed between them with the help of some parameters like confidence, support, lift then it will help us to recommend those products to other customers who are exhibiting the similar kind of purchasing behaviour.

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| **CHAPTER NAME** | **Pg. no** |
| 1. INTRODUCTION | 1 |
| 2. LITERATURE SURVEY | 3 |
| 3. PROPOSED METHODOLOGY | 6 |
| 3.1 Apriori Algorithm | 8 |
| 3.2 Eclat Algorithm | 8 |
| 3.3 FP Growth(Frequent Pattern Growth) | 8 |
| 3.3.1 Frequent Pattern Algorithm Steps | 9 |
| 4. SYSTEM ARCHITECTURE | 10 |
| 5. EXPERIMENTAL RESULTS |  |
| 5.1 Dataset Description | 11 |
| 5.2 Apriori Algorithm Results | 13 |
| 5.3 FP-Growth Algorithm Results | 16 |
| 5.4 Apriori Vs FP Growth | 17 |
| 6. Conclusion | 18 |
|  |  |
| REFERENCES | 19 |

**LIST OF TABLES**

|  |  |
| --- | --- |
| **TABLE NAME** | **Pg. no** |
| 5.2.1 Frequent items with varying support count | 13 |

**LIST OF FIGURES**

|  |  |
| --- | --- |
| **FIGURE NAME** | **Pg. no** |
| 4.1 System Architecture | 10 |
| 5.1.1 Showing Items in Dataset | 11 |
| 5.1.2 Frequency of Itemsets | 12 |
| 5.1.3 Frequency of Most Popular Items | 12 |
| 5.2.1 Frequent Itemsets for Apriori Algorithm | 14 |
| 5.2.2 Association Rule Generation for Apriori | 15 |
| 5.3.1 Frequent Itemsets for FP Growth Algorithm | 16 |
| 5.3.2 Association Rule Generation for FP Growth | 16 |
| 5.4.1 Apriori Vs FP Growth | 17 |

**CHAPTER 1**

**INTRODUCTION**

Data mining can be defined as process of extracting useful data from a large collection of raw data to understand the patterns and do analysis. Data mining uses complicated mathematical algorithms to fragment the data and to estimate the probability of further coming events. For example, by applying data mining techniques automatic discovery the patterns, prediction of probable results, focus on large data sets and databases are carried out. In this paper[1] we explore the data mining technique for online user behaviour analysis is using association rule mining.

Market Basket analysis is a data mining technique that focusing on detecting the purchase patterns of the customers by extracting association rules from a store’s transactional datasets. For example, when the customer buys items in a store then all the item details about their purchased goes into the transaction database. After, this large data are analyzed to determine the purchasing pattern of customers. Also decisions like which item has more stock, which item is more selling cross selling are determined. Association Rule Mining (ARM) identifies the relationship between a large set of data items and forms the base for market basket analysis. Association rule mining has been widely used in various industries besides supermarkets, such as mail order, telemarketing production, food ordering through online, fraud detection of credit card and e-commerce, rainfall prediction. One of the challenges for companies that have invested more on customer data collection that is how to extract major information from their vast customer databases and product feature databases, in order to gain competitive advantage. Market basket analysis has been mostly used in many companies as a means to find product associations. A retailer must know the needs of customers and fit to them. Market basket analysis is one of the possible way to find out which items can be put together.

Market Basket Analysis 2 Market Basket Analysis[2,3] helps to identify the purchasing behavior of the customer. By mining the data from the huge transaction database shop managers can study the behaviour or buying habits of the customer to increase the sale. In Market Basket Analysis, you look to see if there are combinations of products that frequently occurred together in a transaction.

In this modern world whatever might be the field to which human beings will belong to like sales, marketing etc all those fields will revolve around the most popular word ”business” .

Now-a-days most of the business is happening via internet because online shopping made people’s life easier and simpler. Every website administrator will always dream to improve his sales and make his efforts in that way. So there arises the need of online user behavior analysis.

As online shopping is increasing tremendously day by day a large amount of data regarding user purchasing transactions is available on internet. Presence of such data will be very beneficial to a large no. of applications.

Many websites can use such transactions related data i.e, past purchase history. They will analyze the transactions keenly and identify the set of items that were brought more frequently as well as the products which were brought along with some other products. As a result of such analysis they will come to know which products have more demand and which products will be most likely to be purchased with other products. So that whenever a user is purchasing the similar set of items they can recommend those items to such customers. Thus they can improve their sales.

One of the mechanisms that were used in this project to analyze the transactions and derive the most frequently purchasing items is association rules using apriori algorithm, Eclat Algorithm, Fp-Growth algorithm.

**CHAPTER 2**

**LITERATURE SURVEY**

Data Mining[4] provides a lot of opportunities in the market sector. Decision making and understanding the behavior of the customer has become vital and challenging problem for the organization in order to sustain in this competitive world. The challenges that the organization faces is to extract the information from various customer databases, in order to gain competitive advantage.

Yanthy et al [5] in this paper author states about the important goal in data mining is to reveal hidden knowledge from data and various algorithms have been proposed for, but the problem is that typically not all rules are useful –only small amount of data of the generated rules would be of interest to any given users. Hence numerous methods such as confidence values, support count, and lift have been proposed to determine the good or most interesting rules. However some algorithms are good at generating rules high in one aspect but bad in other.

Rakesh Agarwal [6] proposed the Apriori algorithm. Apriori was the first associative rule mining algorithm which is proposed and future development is done in association, classification, associative classification algorithms which have used Apriori algorithm as part of the technique.

Apriori algorithm[7,8] is a level-wise, breadth-first algorithm which counts transactions Apriori algorithm uses prior knowledge of frequent item set properties. Apriori uses an iterative approach known as a level-wise search, in which n-item sets are used to explore (N+1) – itemsets.

To improve the efficiency of the level-wise generation of frequent item sets Apriori property is used. Apriori algorithm insists that all the non-empty subsets of a frequent item set must also be frequent. This is made possible because of the anti-monotone property of support count measure- the support count for an item set never exceeds the support of its subsets.

A two-step process contains of joining and pruning actions are done iteratively It is one of the Data Mining Algorithm which is used to find the frequent items/item set from a given data repository. The algorithm involves 2 steps

a. Pruning

b. Joining

The Apriori property is the important essential factor to be consider before proceeding with the algorithm. Apriori property states that If an item X is joined with item Y then,

Support (XUY) =min (Support(X), Support(Y))

Basically when we are finding the strength of an association rule i.e. how strong the relationship is between the transaction of the items that we measure through the use of the support and confidence.

The support of an item is the number of transaction containing the item. Those items that do not meet the minimum support are excluded from the further processing. Support tells how frequently a rule is applicable to a given data set.

Confidence is defined as a conditional probability that a transaction containing the LHS will also contain the RHS.

Confidence determines how frequently item in Right hand side that appears in the transaction that contain at left hand side also. While determining the rules we must measure these two components because it is very important to us for understanding. A rule that has very low support count may occur simply by chance.

Han [9, 10] presented a new association rule mining approach that does not use candidate rule generation called FP-growth that generates a highly condensed frequent pattern tree i.e, fptree representation of the transactional database. Each database transaction is represented in the tree form by at most one path. FP-tree is smaller in size than the original database. For this construction of tree it requires two database scans, where in the first scan, frequent item-sets along with their support in each transaction are collected and in the second scan, FP-tree is constructed.

Liu [11] proposed CBA the first Associative Classification (AC) algorithm.CBA implements the famous Apriori algorithm [12] in order to discover frequent rule items.

The Apriori algorithm contains 3 main steps:

a. Continuous attributes in the training dataset gets discredited.

b. Frequent Rule Item Discovery

c. Rule Generation

Phani Prasad J, Murlidher Mourya [13] in this paper author states that there are lots of case studies about the a ssociation Rules and existing data mining algorithms usage for market basket analysis but focuses on Apriori algorithm and concludes that the algorithm can be modified and it can be extended in the future work which also reduces the time complexity. Author also clearly told de-merits of the algorithm but claims that there is a way to improve the efficiency of the algorithm.

Owygs[14] in this paper author states that Association rule learning is used in machine learning for discovering more interesting relations between variables. Apriori algorithm is the most popular algorithm for association rules mining and extracting frequent itemsets in association rule learning. It has been designed to operate on databases containing transactions, such as purchases by customers in a store or through website. We can create data visualization for association rules in data mining by using MLxtend for getting association rules.

**CHAPTER-3**

**PROPOSED METHODOLOGY**

Association rules are produced using algorithms like:

* Apriori Algorithm
* Eclat Algorithm
* FP-growth Algorithm

A rule can be defined as an implication, **X⟶Y** where X and Y are subsets of I(X,Y⊆I), and they have no element in common. X and Y are the antecedent and the consequent of the rule, respectively.

**Eg:** {Bread,Egg}=> {Milk} ItemSet={Bread,Egg,Milk}

There are various metrics in place to help us understand the strength of assosciation between antecedent and consequent:

* Support
* Confidence
* Lift or Correlation or interest
* Leverage
* Conviction

**Support:** It gives an idea of how frequent an itemset is in all the transactions.To say in formal terms it's the fraction of total no. of transactions in which the itemset occurs.We refer to an itemset as a "frequent itemset" if you support is larger than a specified minimum-support threshold.

Supp(X=>Y)=(Transactions containing both X & Y)

(Total no.of transactions)

**Range:[0,1]** Value of support helps us identifying the rules worth for future analysis.

**Confidence:** It defines the likelihood of occurrence of consequent on the cart given that cart already has antecedent. It siginifies the likelihood of item Y being purchased when item X Xis purchased.

confidence(X−>Y)= support(X−>Y)

support(X)

* **Lift:** Lift gives the rise in the probability of having {Y} on the cart with the knowledge of {X} being present over the probability of having {Y} on the cart without knowledge about presence of {X}.

Lift(X−>Y) = confidence(X−>Y)

support(Y)

* Lift=1 means both are independent
* Lift > 1 means there is a strong relation between the items
* Lift < 1 means there is less chance that both items will present in the same transaction.
* **Levarage**: It computes the difference between the observed frequency of X & Y appearing together and the frequency that we would expect if A and C are independent.

Leverage(X−>Y) = support(X−>Y) − support(X) ∗ support(Y)

**Range: [-1, 1]**

* If X, Y are positively correlated then we get leverage>0 ,we need such type of rules.
* If X, Y are negatively correlated then we get leverage<0.
* If X, Y are independent, then we get leverage = 0.
* **Conviction:** It can be interpreted as the ratio of the expected frequency that X occurs without Y (that is to say, the frequency that the rule makes an incorrect prediction) if X and Y were independent divided by the observed frequency of incorrect predictions.

Conviction(X−>Y) = support(Y)

confidence(X−>Y)

**3.1 APRIORI ALGORITHM:**

Apriori Algorithm is a significant for mining frequent item sets. It is the most useful algorithm for association rule mining

**Steps :**

1. Input: {Transactional database, Minimum support count}

2. Output: {Interesting association rules}

3. Identify the frequent itemsets: Itemsets that satisfy the minimum support count

An itemset is said to be frequent iff all of its subsets are frequent. i.e., {I1, I2} is a frequent itemset, only if both {I1} and {I2} are frequent

Recursively find all frequent itemsets of level from 1 to k (k-itemset).

4. Generate Association rules based on identified frequent itemsets.[15,16]

**3.2 ECLAT ALGORITHM:**

BottleNecks of Apriori:

Candidate generation can result in huge candidate sets

Multiple Scans of Database--- needs (n+1) scans, n is the longest pattern

To solve some of the above problems, Eclat has been introduced.

The ECLAT algorithm stands for Equivalence Class Clustering and bottom-up Lattice Traversal. It is one of the popular methods of Association Rule mining. It is a more efficient and scalable version of the Apriori algorithm. While the Apriori algorithm works in a horizontal sense imitating the Breadth-First Search of a graph, the ECLAT algorithm works in a vertical manner just like the Depth-First Search of a graph. This vertical approach of the ECLAT algorithm makes it a faster algorithm than the Apriori algorithm.[17,18]

**3.3 FP GROWTH (Frequent Pattern Growth):**

Shortcomings of Apriori Algorithm

* Using Apriori needs a generation of candidate itemsets. These itemsets may be large in number if the itemset in the database is huge.
* Apriori needs multiple scans of the database to check the support of each itemset generated and this leads to high costs. These shortcomings can be overcome using the FP growth algorithm.[19]

This algorithm is an improvement to the Apriori method. A frequent pattern is generated without the need for candidate generation.[20] FP growth algorithm represents the database in the form of a tree called a frequent pattern tree or FP tree.[21]

**3.3.1 Frequent Pattern Algorithm Steps:**

1. Calculate minimum support count
2. Find frequency of occurrence
3. Prioritize the items.
4. Order the items according to priority.
5. Validation.

**CHAPTER – 4**

**SYSTEM ARCHITECTURE**

Database

Load transaction file

Data Preprocessing

Association rule mining

Data Mining Techniques

Generate frequent itemsets

Store strong association rules

Fig: 4.1 System Architecture

First we load the dataset from the database and apply data preprocessing techniques to remove noisy data, missing values and redundant data. Then apply association techniques like apriori algorithm Fp-growth algorithm to generate strong association rules.

**CHAPTER 5**

**EXPERIMENTAL RESULTS**

**5.1 Dataset Description**

The data was collected from:

<http://www.kagels.com/wp.content/uploads/2017/08/Market_Basket_Optimisation.csv> due to the unavailability of data from the online shopping sites.

The dataset is about the items purchased by the customers over a period of time. The data set contains 7501 transactions of different customers buying different items from the store.

We have to find correlations between the different items in the store. So that we can know if a customer is buying bread, honey and milk. What is the next item, the customer will buy from the store.

First we need to perform data preprocessing to remove redundant data and missing values. We use the TransactionEncoder() of mlxtend.preprocessing to do this work.TransactionEncoder() is an Encoder class for transaction data in Python list.It finds out what are all the different products in the transactions and will assign each transaction a list which contains a boolean array where each index represents the corresponding product whether purchased in the transaction or not i.e. True or False.

A sample of 5 transactions in a dataset is shown below:

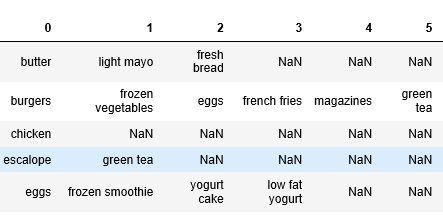


Fig: 5.1.1 Showing Items In Dataset

The below figure shows the total count of items and frequency of all items in the dataset along with their names.

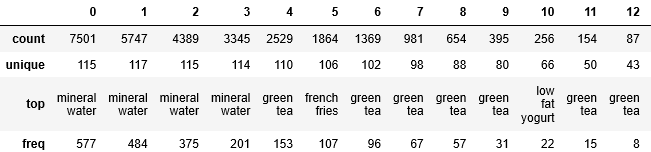


Fig: 5.1.2 Frequency of Itemsets

Generation of the frequency of most popular items can also be shown in graphical form.

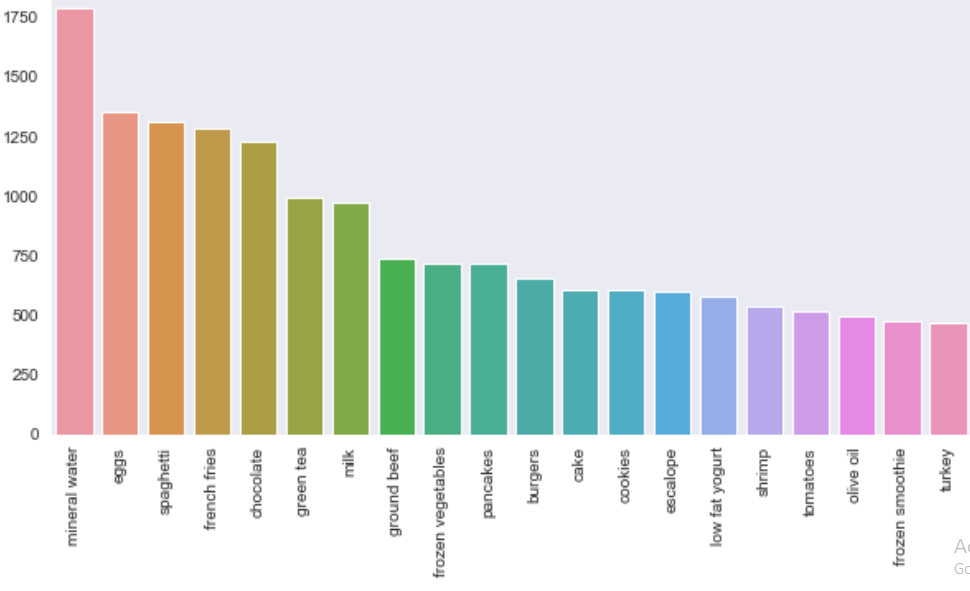


Fig: 5.1.3 Frequency of Most Popular Items

We can find that mineral water is the most purchased item from the store, we may advice that mineral water must be always in the stock not only that mostly we can see from the above graph what 20 items are being frequently purchased.

**5.2 Apriori Algorithm Results**

The Apriori algorithm generates association rules for a given data set.

In the test approach sample input was given to the system and different support. At first large sample of input dataset were given to the system with different support. Then few sample input were given to the system with different support count..The table 5.2.1 shows the support count and number of rules generated by changing the support value.

Table 5.2.1: Frequent items with varying support count

|  |  |  |
| --- | --- | --- |
| S. no | Support count | No. of frequent itemsets |
| 1 | 0.001 | 386487 |
| 2 | 0.005 | 526 |
| 3 | 0.01 | 209 |
| 4 | 0.02 | 87 |
| 5 | 0.05 | 27 |
| 6 | 0.1 | 6 |

We observed that when support count increases the number of frequent itemset generation decreases.

Let’s generate itemsets with atleast 5% support for applying apriori algorithm

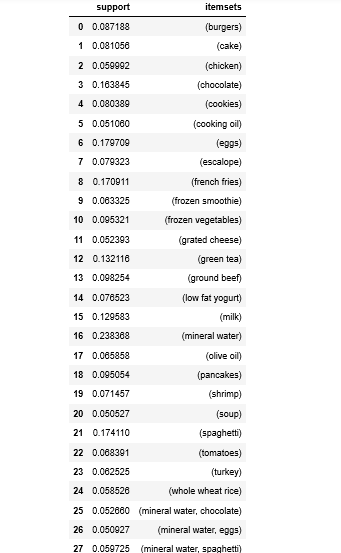


Fig 5.2.1 Frequent itemsets for Apriori algorithm

The above figure tells us that there are 27 frequent itemsets of different lengths, so the first step of our apriori algorithm is finished. Now move to next step i.e., generate association rules by using this frequent item sets.

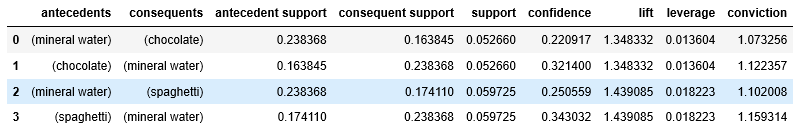


Fig 5.2.2 Association rule generation for Apriori

Above we can see the 4 rules generated with lift greater than 1.3

Intuition we can get is that:

* 22% of transactions containing mineral water also contain chocolate
* 32% of transactions containing chocolate also contain mineral water
* 34% of transactions containing spaghetti also contain mineral water
* 25% of transactions containing mineral water also contain spaghetti

There is more chance of the transaction {spaghetti,mineral water} than {chocolate,mineral water} as we can find the interesting nature of rule by comparing lift,leverage and conviction of {spaghetti,mineral water} and {chocolate,mineral water}.

Therefore, we can say that most of the customers buys mineral water. So, we can recommend mineral water to customers along with their products.

**5.3 FP-Growth Algorithm Results:**

Lets generate frequent itemsets with support count as 5%

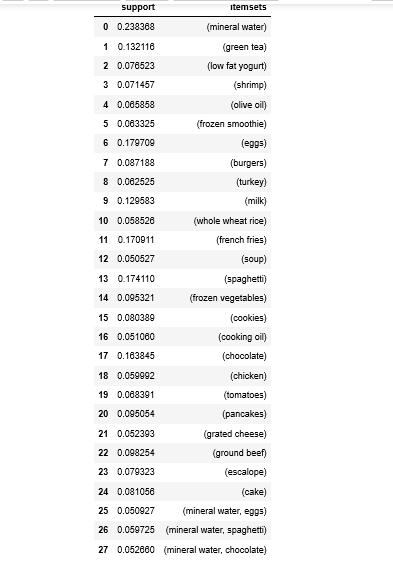


Fig 5.3.1: Frequent itemsets using Fp-Growth algorithm

The above figure tells us that there are 27 frequent itemsets of different lengths, so the first step of our Fp-Growth algorithm is finished. Now move to next step i.e., generate association rules by using this frequent item sets.

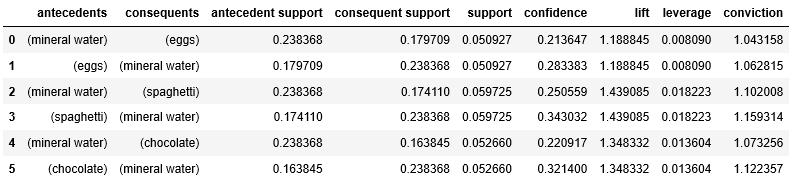


Fig 5.3.2: Association rule generation for FP-Growth

We could observe that {spaghetti}->{mineral water} is mostly like to occur as we can observe it from the lift. By using Fp-Growth we get the strong association rules than apriori algorithm

**5.4 Apriori Vs FP Growth:**

Since FP-Growth doesn't require creating candidate sets explicitly, it can be magnitudes faster than the alternative Apriori algorithm. FP-Growth is about 5 times faster. Let's look at it.

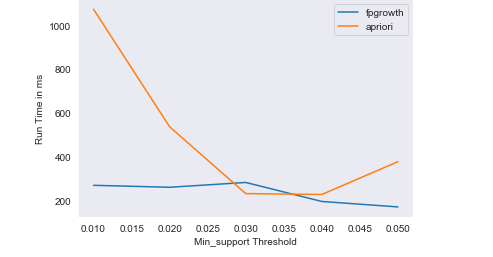


Fig 5.4.1 Apriori Vs FP Growth

We can gain the required insights from the above graph about the run time comparison between the apriori and Fp-Growth.

**CHAPTER 6**

**CONCLUSION**

The Fp-growth algorithm effectively generates highly informative frequent itemsets and association rules for the data of the online user shopping details than Apriori algorithm. Strong association rules are generated based on frequent items.

Thus association rules can be generated by using apriori algorithm and Fp-Growth which will identify the given the frequent item sets from which with the help of parameters like support, confidence, lift, convention it can generate strong association rules that will reflect the real world transactions. Using these rules we can perform recommendation of items to the users/customers.

**REFERENCES**

[1]B. Singh and H. K. Singh, “Web data mining research: a survey,” in Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on. IEEE, 2010, pp. 1–10. [20] W. M. P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, 1st ed. Springer Publishing Company, Incorporated, 2011.

[2] N. Poggi, V. Muthusamy, D. Carrera, and R. Khalaf, “Business process mining from e-commerce web logs,” in Proceedings of the 11th Inter- national Conference on Business Process Management, ser. BPM’13. Berlin, Heidelberg: Springer-Verlag, 2013, pp. 65–80.

[3] F. M. Maggi, R. P. J. C. Bose, and W. M. P. van der Aalst, Efficient Discovery of Understandable Declarative Process Models from Event Logs. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 270– 285.

[4] Pang-Ning Tan, Michael Steinbach, Vipin Kumar: Introduction to Data Mining.

[5] W. Yanthy, T. Sekiya, K. Yamaguchi,”Mining Interesting Rules by association and Classification Algorithms”, FCST 09.

[6] R. Agrawal and R. Srikant,” Fast Algorithms for Mining Association Rules in Large Databases”, Journal of Computer Science and Technology, vol. 15.

[7]Deepak.r.poojari,apriori-algorithm-in-python recommendation engine

[8] X. Yin, J. Han, “CPAR: Classification based on Predictive Association Rules”, Proceedings of the Third SIAM International Conference on Data Mining, pp 331-335, 2003. 9.Gourab Kundu , Sirajum Munir, Md. Faizul Bari, Md. Monirul Islam, and K. Murase, ”A Novel Algorithm for Associative Classification”,14thInternational Conference, ICONIP 2007, Kitakyushu, Japan, pp 453-459 , November 13-16, 2007.

[9] J. Han, J. Pei, and Y. Yin, “Mining frequent patterns without candidate generation”, Proc of the ACM SIGMOD International Conference on, vol. 1, , pp. 1-12, 2000.

[10] J. Han, J. Pei, Y. Yin, and R. Mao, ”Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach”

[11] Phai Prasad J, Murlidher Mourya, ”A Study On Market basket Analysis Using Data

[12 ]Sanjeev Market Basket Analysis, DWIT college,2016

[13] ”create-data-visualization-for-association-rules-in-data-mining”,

[14]<https://en.wikipedia.org/wiki/Association_rule_learning>.

[15]Yan, Xifeng, and Jiawei Han. "gspan: Graph-based substructure pattern mining,” Data Mining ICDM Proceedings,IEEE International Conference, 2002.

[16]Honest, Nirali, Atul Patel, and Bankim Patel, “A study of Path Completion Techniques in Web Usage Mining,” Computational Intelligence & Communication Technology (CICT), IEEE International Conference on. IEEE, 2015.

[17]<file:///C:/Users/HOME/Downloads/dharmaraajan2016(1).pdf>

[18]Agrawal, R., Imielinski, T., Swami, A. Data Mining: A performance Perspective. IEEE Trans. Knowledge and Data Engineering, vol, 5,6, 1993a, pp. 914-925.

[19] Farhana Seemi,Hania Aslam,Hamid Mukhtar, “Browsing Behaviour anlaysis using Data mining”, University of Engineering and Technology(UET),2019.

[20] Abhijit Raorane & R.V.Kulkarni, “Data Mining Techniques:consumer behaviour analysis”

[21] Data mining techniques and concepts Jiauei Han, Michele Kamber, simon Fraser University ISBN 1-55860-8-2001.

[22] Mining association with the collective strength approach- Aggarwal C.C.,Yu, P.S.2001.

[23] Leidys del Carmen,Kevin Andrey Roasales Ferreria,” Analysis of the behaviour of customers in social networks using data mining”, IEEE,2016.

[24] Agrawal, Rakesh, and Ramakrishnan Srikant, “Fast algorithms for mining association rules” Proc. 20th international conference very large data bases, VLDB. Vol. 1215, 1994.

DATASET LINK:

<https://www.kaggle.com/roshansharma/market-basket-analysis?select=Market_Basket_Optimisation.csv>

<https://www.kaggle.com/sajidcse/market-basket-analysis/notebook#FP-GROWTH(Frequent-Pattern-Growth)>