

An effort to study the ordering pattern of customers in a fast food outlet using FP growth algorithm

NUSRAT JABEEN.T
RESEARCH SCHOLAR
BHARATHIAR UNIVERSITY
Email : tn_jabeen@yahoo.co.in

V.RAMYA
RESEARCH SCHOLAR
BHARATHIAR UNIVERSITY

ABSTRACT

The task of finding frequent pattern in large databases is very essential and has been studied in large scale for the expansion of any product in a Company. This assignment is computationally costly when countless exist. Visit designs are designs that show up as often as possible in a dataset. The disclosure of intriguing connection relationship among gigantic measures of business exchange records can help in numerous business basic leadership process. The FP growth algorithm can be used to find frequent item sets without using candidate generations. The successive example tree is utilized to store packed and focal data about regular examples that can be utilized to mine continuous examples in extensive databases. This paper talks about the FP Tree idea and tries to examine the requesting example of clients in a fast food outlet. This approach decides the affiliation decides that happen in the dataset of a fast food outlet that can be utilized to outline business procedures to upsurge income for a specific outlet when the deal goes low. Being given a set of transactions of the customers the purpose of the association rule is to find correlations between the sold products so that the decisions can be made in terms of offering the combo offers or packages to customers to implement successful marketing techniques for an outlet. It is interesting to learn the association between the customer's order when we find that every customer orders a pizza or a burger along with a soft drink like Pepsi or a starter like French fries. The knowledge of frequent pattern increases the sale of the soft drinks or any snack that is ordered along with it. This pattern helps the Company to improve their sales and increase their revenue. The implementation of FP growth algorithm will be to a great degree helpful in advertise inquires about. We can discover concealed data and relationship from the information and further choices can be taken in view of the gained data.

Keywords: Association, Apriori, FP Growth Tree, Data Mining

1. INTRODUCTION

The aim of association rule is to find correlations between the sold products in a given set of transactions of the customers. This analyses customer eating habits by finding associations between the different items that customers place in their orders.

1.1.Data Mining

Information Mining is the way toward finding fascinating examples and learning from a lot of information. The

information source can incorporate databases, information stockrooms, the web, and other data archives or information that are gushed into the framework powerfully [1].

1.2.Uses of Data Mining

The most common uses of data mining are:

- Basket analysis
- Sales forecasting
- Database marketing

- Merchandise planning
- Card marketing

1.3. Benefits of Data Mining

The data mining process includes reporting and analysis. Data-driven marketers are wise to keep in mind that they have to use detailing and investigation as two separate parts of the entire information mining process, with the goal that they can settle on information driven promoting choices established in the bits of knowledge picked up from both revealing and examination.

- More cash – Learn which kinds of items clients have bought and benefit from that knowledge to customize involvement, increment client reliability, and lift client lifetime esteem
- Improve marking and showcasing – Get input and utilize information mining to recognize what is working and what isn't with marking and promoting
- Streamline outreach – Make the greater part of your effort all the more auspicious and pertinent with information mining

1.4 Association

An affiliation administer has two sections, a precursor (if) and an ensuing (at that point). A forerunner is a thing found in the information. An ensuing is a thing that is found in mix with the predecessor. Affiliation rules are made by breaking down information for visit if/at that point examples and utilizing the criteria support and certainty to recognize the most essential connections. Support means that how every now and again the things show up in the database. Certainty demonstrates the circumstances the if/at that point articulations have been observed to be valid.

2. FREQUENT ITEMSET MINING ALGORITHM

Fp-development receives isolate and vanquish procedure [3]. It packs the database speaking to visit things into a successive example tree or Fp-tree which holds the itemset affiliation data. It at that point isolates the compacted database into an arrangement of restrictive databases; each related with one regular thing or example piece and mines every database independently. For each "example section" just its related informational indexes should be inspected. In this way this approach may significantly decrease the measure of the dataset to be looked alongside the development of examples being inspected.

3.1 Construction of Fp-tree

FP-Tree is constructed using two steps over the data set: [1]

Step 1 :

- Check data and find support for each item, Eliminate infrequent items
- Arrange frequent items in decreasing order based on their support.

Let us take an example: u, v, w, x, y.

Step 2:

Construct the FP-Tree

- Read first transaction : {u,v}
Create two nodes u and v and the path null $\rightarrow u \rightarrow v$. Set count of u and v to 1.
- Read second transaction : {v, w, x }
}

Create three nodes for v, w and x and the path.null $\rightarrow v \rightarrow w \rightarrow x$. Count is set to 1.

- .Read third transaction: { u, w, x, y }
}

The transactions are continued until the paths are mapped in the FP-tree.

Table 1 : Transactions

TID	ITEMS
1	{u,v}
2	{v,w,x}
3	{u,w,x,y}
4	{u,x,y}
5	{u,v,w}
6	{u,v,w,x}
7	{u}
8	{u,v,w}
9	{u,v,x}
10	{v,w,y}

a. FP-Tree size

The size of FP-Tree depends on how the items are ordered.

b. Frequent Itemset Generation

- FP-Growth extracts frequent itemsets from the FP-tree.
- It proceeds from the leaves towards the root.
- Uses divide and conquer to scan for frequent itemsets ending in y, then xy, and so on.
- Extracts prefix path sub-trees ending in an itemset.
- Each prefix path sub-tree is processed recursively to extract the frequent itemsets. Finally merges the solutions.
- Example the prefix path sub tree for y will be used to extract frequent itemset ending in y then in xy, wy, vy, uy and so on.

4. IMPLEMENTATION

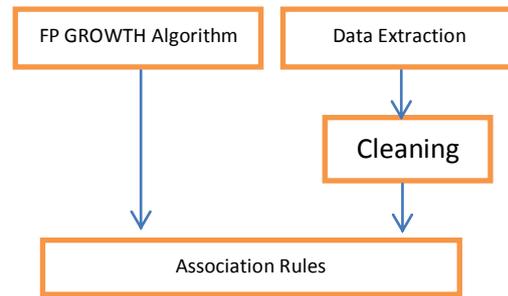


Fig1: Flowchart of stages during Implementation

Figure 1 demonstrates the execution procedure. The execution starts with the client arrange dataset acquired from the client and involves an arrangement of characteristics. This dataset is cleaned for finding the recurrence of requested thing [2]. The Fp development calculation is connected on the cleaned dataset that outcomes in the arrangement of affiliation rules required for investigation.

3.1.Data Collection

The data required for our analysis was taken from a restaurant in Chennai. The data which we utilized in our analysis consists of the customers ordered items in a menu. It comprises of a range of attributes: Starter, Main Course, Beverages, etc The parameters taken into consideration are :STARTER is classified as French fries or Nuggets, MAIN COURSE is divided into four groups. Chicken Burger, Vegetable Burger, Spicy Chicken Wrap, Grilled Fish Wrap. BEVERAGES as Pepsi and Cold Coffee. The important parameter CUSTOMER ORDERS A STARTER OR MAIN COURSE WITH A BEVERAGE IS set to Yes or No.

3.2.Data Preprocessing

Data preprocessing [3] involves data cleaning and data transformation. Information Cleaning cleans the information by filling in missing esteems,

distinguishing or evacuating anomalies and settling irregularities. The following measure was taken to clean our data: Remove all the orders of the customers who have bought the Main Course without Beverages or Starters. Removing these kinds of orders will give us the data that contains only those orders that have been placed with a Main Course along with Beverages or Starters. A numeric esteem was allotted to each of the info passages. All classifications had a remarkable numeric personality. Since four general classifications were chosen, four segments comprising of novel numeric assignments were utilized.

Table 2. Parameters of dataset and codes assigned to them

COLUMN	CODE ASSIGNED	ATTRIBUTES
CUSTOMER ORDERS A STARTER OR MAIN COURSE WITH A BEVERAGE	1	Yes
	0	No
STARTER	11	French fries
	12	Nuggets
MAIN COURSE	21	Chicken Burger
	22	Vegetable Burger
	23	Spicy Chicken Wrap
	24	Grilled Fish Wrap
BEVERAGES	31	Pepsi
	32	Cold Coffee

4. RESULT AND ANALYSIS

4.1. Generating Association Rules

Association rules are found after preparing the dataset for input. The default support and certainty levels are taken as 20% and 80% separately.

4.2 Analysis of Association Rules

One control is found, implying the principles {22,31} - > {1}, which proposes that when the component fields 'Vegetable Burger' and 'Pepsi' are discovered, the y are joined by the field 'No' , where 22 portrays the esteem 'Vegetable Burger' of MAIN COURSE, 31 delineates the esteem 'Pepsi' of BEVERAGES characteristic and 1 delineates the 'No', esteem of CUSTOMER ORDERS A STARTER OR MAIN COURSE WITH A BEVERAGE. Finally, when a customer orders a Vegetable Burger with Pepsi, then 85.04% times, the customer orders any main course with a beverage.

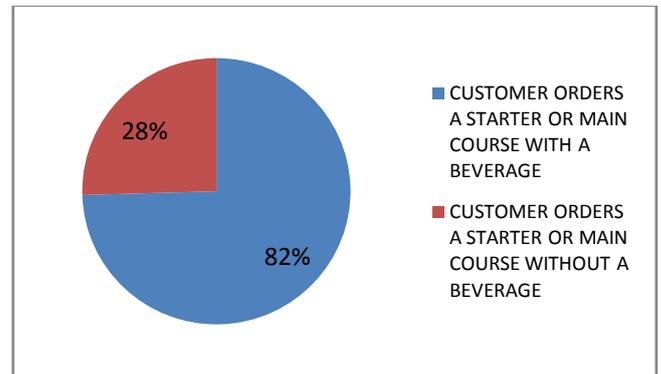


Fig 2. Frequency of occurrences for all records on the basis of 'CUSTOMER ORDERS A STARTER OR MAIN COURSE WITH A BEVERAGE'

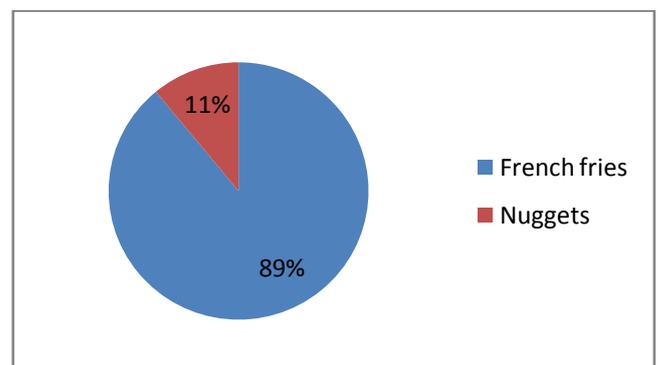


Fig. 3 Frequency of occurrences for all records on the basis of 'STARTERS'

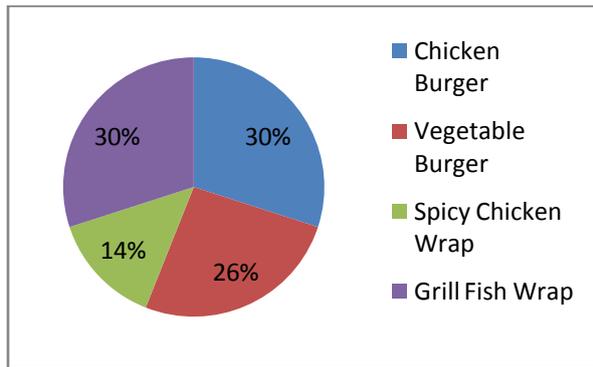


Fig.4. Frequency of occurrences for all records on the basis of 'STARTERS'

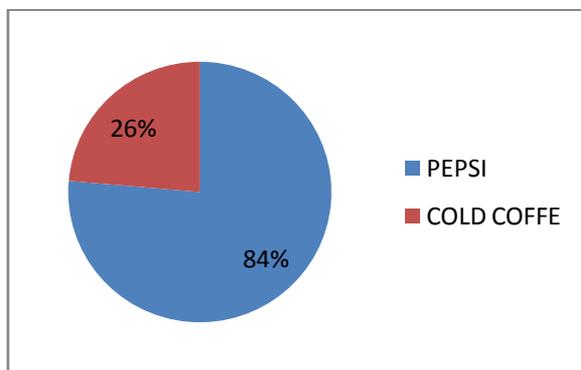


Fig. 5. Frequency of occurrences for all records on the basis of 'BEVERAGES'

We watch that a client orders principle course with drinks is more than those requesting the fundamental course alone. Discusses recurrence of occurrences based on STARTER, here we can presume that the client wants to arrange French fries instead of Nuggets. The recurrence of events depends on Main Course. Here it can be construed that about a large portion of the client's request chicken burger and more than a quarter arrange different things. We can watch that the clients tend to drink Pepsi when contrasted with Cold Coffee.

5. CONCLUSION

This paper displays the criticalness of the FP development calculation that can be utilized to discover visit thing sets without utilizing applicant ages. This approach analyzed the association rules that occur in

the dataset of a fast food outlet. It is interesting to learn the association between the customers order when we find that every customer orders a burger along with a soft drink like Pepsi or a Cold Coffee. The knowledge of frequent pattern increases the sale of the soft drinks or any snack that is ordered along with it. This pattern helps the Company to improve their sales and increase their revenue. The usage of FP development calculation will be to a great degree valuable in showcase examines. We can discover concealed data and relationship from the information and further choices can be taken in light of the obtained data.

REFERENCES

- [1] Jiawei Han, Micheline Kamber, Jian Pei, "Data Mining: Concepts and Techniques", June 2011, Elsevier.
- [2] Shivam Sidhu, Upendra Kumar Meena, "FP Growth Algorithm Implementation", International Journal of Computer Application, Vol 93, No.8, May 2014.
- [3] Dr. Varun Kumar, AnupamChadha, "Mining Association Rules in Students's Assessment Data", International Journal of Computer Science, Vol.9, Issue 5, No.3, September 2012.
- [4] S.O.Abdulsalam, K.S. Adewole, "Data Mining in Market Basket Transaction: An Association Rule Mining Approach", International Journal of Applied Information System, Volume 7-10, October 2014
- [5] Rajdeep Kaur Aulakh, "Association Rule Mining Using Effective Algorithm: A Review", International Journal of Advanced Research in Computer Science and Engineering, Vol.5, Issue 3, March 2015.