



Achieving Effective Customer Relationship using Frequent Pattern-Growth Algorithm Association Rule Learning Technique

N. V. Ugwu*, C. N. Udanor

Department of Computer Science, University of Nigeria Nsukka, Enugu State, NIGERIA.

Abstract

Customer relationship management (CRM) is a methodology and tool that possesses the plan and techniques that companies should follow in relating with their customers. In today's business world, the relationship between organizations and their customers is very important in engaging the customers' interest, which has the direct effect in increasing the business profit. Due to ineffective collaboration and interaction between organizations and their customers, identifying who the real customers are, along with what they need has failed. A breach of trust from the company, and lack of feedback from the customer could make an organization not to be able to compete with her rivals in a business environment and win her customers' loyalty. Therefore, the guarantee of the customer continuing transactions with the company may no longer be assured. This work deploys an association rule learning technique of data mining using frequent pattern growth algorithm to identify which items are regularly purchased together by customers and based on this result, analyzes and understands the customers' buying habits. Object-Oriented Analysis and Design methodology (OOAD) is used to analyze and design the system, whereas the implementation is carried out using Python programming language and My-SQL database management system. The contribution of this work is that it enables firms to gather and analyze customers' interests in a product so that the needs of their valued customers will be met in order to make them return for more business transactions, thereby achieving customer retention.

Keywords: customer relationship management, association rule, data mining, frequency pattern algorithm

1. INTRODUCTION

With the arrival of information and communication technology (ICT), many organizations have been promoting information systems to deliver products and services to their customers effectively. Due to the high increase in global competition in businesses [1], ICT has created ways for new market trends which enable the customers to have a wide array of choices, and be more informed about the products and services of interest. On the other hand, ICT has also made the service providers to have more knowledge, more power and more influence in enlarging their interests, which has a direct impact on increasing business profit. In today's business landscape, organizations must be more concerned about customer satisfaction and retention and must make efforts to gain and win their trust by providing an ICT platform that is always available to the customers. An organization's ineffective dealings with her customers may lead to a breach in the

relationship. And this would likely make the organization unable to compete with her rivals. The importance of insisting on close, firm and quality relationships with the customer cannot be compromised. Fig. 1 illustrates the customer retention program which enables the organization to retain her valued customers, making for a good relationship with their customers.

1.1. Customer Services

The customer is the focal point of any company or organization. Therefore, companies should make conscientious effort to retain any customer within their reach. Within the company, customer services are given precedence to promote customers retainership. There are two main approaches designed by companies to enhance customers' services;

- **Reactive services**

This is a situation where the customer reacts based on the problem they perceived from the service providers and makes contact with them to solve it.

- **Proactive services**

This is a situation where the company does

*Corresponding author (Tel: +234 806 8544 121)

Email addresses: emekaugwu13@gmail.com (N. V. Ugwu), collins.udanor@unn.edu.ng (C. N. Udanor)

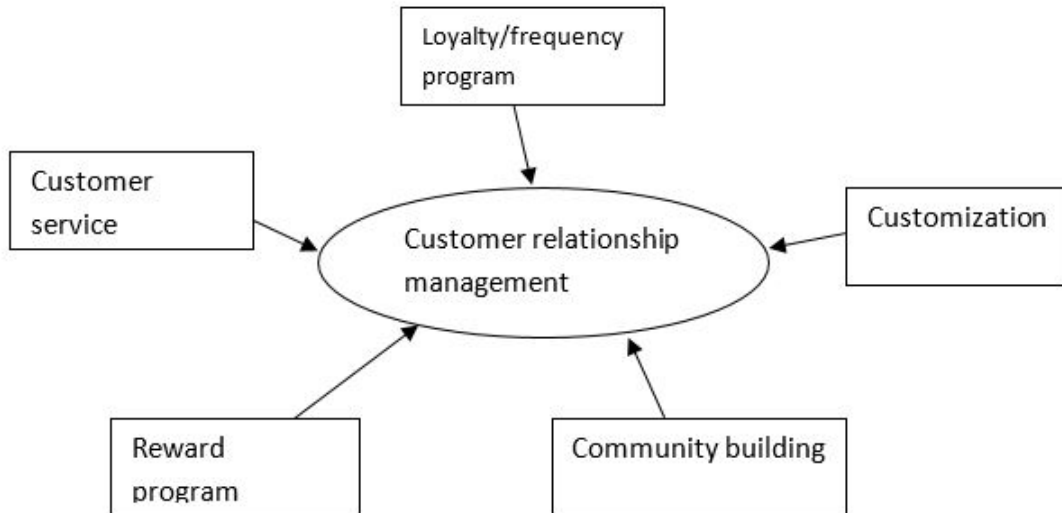


Figure 1: Customer retention program [2].

not wait for the customer to complain before the problem is fixed, but reaches out to the customer to have an exchange of ideas as soon as possible.

1.2. Loyalty/Frequency Program

This program is designed by a company to give a reward to any customer for a good turnover to the company. For example, some web-based companies give bonuses to customers for regular visits to the companies' websites.

1.3. Customization

Customization is a method of modifying products and services in order to meet up with, and satisfy an individual customer's need. The idea of mass customization goes further than one-to-one marketing as it insinuates the formation of products and services for individual customers. The information obtained from customers to modify products can be of advantage to the manufacturers to put a new look or appearance to the products.

1.4. Reward Program

To appreciate the patronage of a good customer, companies can decide to give rewards to its customers on successive purchases of merchandise from them. And the reward may be in form of reducing the actual price of goods for them by giving them a discount or giving special preferences to regular customers, etc.

1.5. Community Building

Building a network for easy exchange of information among the customers, and also, creating a medium of interactions between the customers and the company is very vital, and this is called community building. The goal is to collect all the suitable data and information about their customers and serve them the way they expect. And

with this, the manager can design an enabling platform that will be more accommodating to the customer to influence other people who may also wish to buy from the corporation [2]. For instance, Amazon.com provides a platform where anybody can sell or purchase whatever he/she desires using its platform and this makes shopping experience more satisfying. Such experience could be like recommendations based on past purchases, and a list of reviews and guides written by users that have purchased the exact products the prospective customer desires.

For an organization to retain their customers, a good customer relationship management system (CRM) is needed. It is a customer-centric business system which a corporation implements or rents to improve a customer experience and satisfaction by customizing products and services to the customers' needs [3].

In information systems data mining is described as a method that uses several types of data analysis and modelling strategies to determine patterns and associations in data which can be used to make an accurate forecast. It also helps companies to select the correct customer to focus on, offers the correct additional items to their existing clients and identifies good clients who may be about to leave [2]. Data mining has played an important role in industries and plays a central role in the field of data science. Association rules when applied to data mining helps immensely in discovering patterns of frequent sets of items that associate together. This is regularly recognized as "Market Basket Analysis", which is said to be the first application area of association mining. It is often used to help an organization to discover an association or relationship between attributes or items of interest in a large transactional database [4].

CRM applications that use data mining are called analytical customer relationship manage-

ment [2]. The analytical CRM is the focus of this work because, it is a process that will be used to collect registered customers' information secretly for analyses purposes in order to gain deep knowledge about their purchasing behaviour, and it also helps in discovering patterns of frequent sets of items that are habitually associated. For example, bread and butter are habitually associated. CRM is a well-determined approach to scrutinizing the customer's statistics that helps to improve several business processes in marketing and services. Organizations use analytical CRM to gather customers' information from several channels and also gain deep knowledge about customers' purchasing patterns or habits. Analytical CRM is one of the systematic ways that also helps an organization to build new marketing policy or strategy [5]. Data mining has a strong association with analytical CRM because they are two mutual functions used by the organization to mine and analyze data in a large transactional database such as customers' information, sales information, marketing information, service information, and channel information.

The contribution of this work is that it deploys an association rule learning technique of data mining using frequent pattern growth algorithm to identify which items are regularly purchased together by customers and based on this result, analyzes and understands the customers' buying habits. This will go a long way in helping both service providers and their customers. It helps the service providers to predict customers' buying habits and ensures that the products regularly purchased by these customers are readily available, thus achieving customer retention, and maximizing profit in the end. The customers also are assured of the ready availability of their supply of commodities.

The rest of the paper is organized as follows; Section 2 reviews similar works. Section 3 provides the methodology and approach to the research, while Section 4 presents the results of the research with its accompanying discussions. Section 5 summarizes and concludes the research findings while throwing light to further works.

2. LITERATURE REVIEW

In the research on the customer relationship management system, association rule learning (market basket analysis), is among the most notable research methods and also, the most frequently used technique in data mining. Association rule learning helps to find suitable conditions and measures for extracting remarkable and hidden relations and patterns among variables in huge transactional databases. It helps in giving an organization insight into defining business strategies [6]. The two notable algorithms are; apriori algorithm which utilizes breadth-first search technique and Frequent Pattern (FP) growth algorithm which operates with depth-first search method and also operates with divide and conquer technique. These are the most popular algorithms for mining association rules. But

on comparing both algorithms; Frequent Pattern growth was found to have numerous advantages over the apriori algorithm in terms of efficiency. The reason is that, for the apriori algorithm to scan any transaction the entire database would have to be scanned and this reduces the efficiency of the algorithm [7].

Suchacka and Chodak [8], proposed an approach that developed from association rule discovery in e-customer sessions in a manner to evaluate the buying likelihood in an online session. Two client clusters were differentiated based on client choices for types of items: innovative clients who see and once in a while purchase typical items, as well as innovative items like audio-books and multimedia items, and traditional clients who see and once in a while purchase only typical items like printed books. Then, association rule was applied to targeting and recognizing these sessions' feature for both client clusters which depict a high likelihood of making a purchase in session, and to examine chronological data gotten from a real on-line bookstore. The outcome provided an essential input to the identification of features characterizing the on-line conduct of customers in on-line stores.

Mak, et al. [9], proposed an intelligent Financial Data Mining Model (FDMM) for mining client behaviour in the financial industry so as to boost the availability of decision support data and therefore enhance a client's contentment. A combination of two notable techniques, clustering and association rules were applied to help the executives of a corporation to have more knowledge on improved analysis of client buying pattern and to devise a perfect investment convergence to their clients.

According to Beom-Soo and Yong-Moo [10], the authors proposed a customer relationship management (CRM) strategy for a small-sized on-line shopping mall using association rules and sequential pattern for the application. The authors stated the meaning of V.I.P customers in relation to Recency, Frequency, and Monetary (RFM) value; where Recency refers to the time interval between the last consuming and the current. Frequency refers to the number of transactions over a certain period of time and monetary refers to the amount of money spent on products or services. These four data mining techniques were used to classify customers as VIP or Non-VIP. Artificial neural network, decision tree, logistic regression, and bagging and their hit ratio were compared. Association rules and sequential hidden patterns were discovered for the V.I.P customers.

An algorithm described as rules induction algorithm for customer relationship management was proposed by Dhandayudam and Krishnamurthi [11]. This is one of the data mining methods that is used in generating rules for each cluster produced by a clustering algorithm. Rules can be written as IF-THEN condition and can only be generated to describe and predict the behaviour of the customer in a particular cluster or group using LEM2 (Learning from Examples Module, version

2) algorithm for induction of decision rules. The authors reported that after the experiment was conducted using LEM2 and the proposed algorithm, the results show that rules induction algorithm was used for describing customer behaviour in each cluster or group.

According to Isakki et al. [12], the K-means was applied to the segmentation of customers' details in which Apriori algorithm was applied to identifying customers' behaviours from transactional records. The aim of this method was to develop a very useful style for introducing products with configurations for customers of different gender based on past transactions. Clustering and association rules were introduced to predict and identify customer behaviour based on their previous transactions.

Changchien and ChuenLu [13] proposed a data mining procedure with the aim of developing a data mining association rule to assist in on-line recommendation. This proposed procedure for data mining was comprised of two important sections of operation. The first section was a clustering section that uses a neural network and self-organization map (SOM) that performs data grouping tasks in a large number of database records. The second section was the extraction section which makes use of rough set theory that would be able to mine association rules for a group of the same nature of data proceedings and the association among diverse groups. The developed system was proposed to be an illustration on a database using a snippet of sales records.

Njoni et al. [14], proposed a methodology for mining social media data using fuzzy Association Rule technique. However, the method or techniques like Fuzzy association rules can also be combined with a suitable algorithm to build tools that can be of advantage to address this challenge. Facebook was chosen for the hypothetical analysis and the main objective was to obtain the non-obvious relationships among people (for example potential employees of a company) to get more knowledge into the behaviour, character, and other properties that can be of a problem to the performance of the employee.

According to Pong Yang Wan [15] an association rules mining among items based on rough sets. Distinct feature vectors can also be extracted from different types of products. Reduction sets of products may be gotten hold of by making use of discernibility matrix and discernibility function, which is then used to obtain association rules. However, by introducing the correlation analysis, the result of the rules will then be verified.

A tourism recommender system model was proposed by Masoumeh and Mahdavi [16]. And this recommender system model was used to suggest tours to customers by means of two data mining methods: clustering and association rules. Initially, SOM algorithm was used to cluster tourists to ascertain the actual amount of groups. And the clusters were created using a k-means algorithm. However, by examining the tours in which the tourists have ordered previously, association rules

were generated for each cluster with Apriori- algorithm. A recommendation was performed for active customers on the website. And this was based on the location of the targeted customer in the cluster and the previous behavioural patterns of the tourists based on their shopping history.

Kaur and Singh [17] presented the important features of a frequent pattern (FP-Growth) algorithm in mining frequent pattern in a sports store using association rules. The FP-growth algorithm was used on the data gathered from the sports store to extract the frequent pattern between the products. The research was performed with the intention of evaluating the practical possibility of using samples for finding frequent item sets.

C. YS et al. [18], proposed a novel algorithm for an on-line business store which was described as incremental weighted mining association rules (IWMAR), utilizing FP- tree. The forth put algorithm was based on Recency, Frequency, Monetary value (RFM) analysis for recommending prediction for an on-line store. The algorithm was also to enhance the exactness of a recommendation of products with higher purchase-ability, utilizing the customer's score and the product score in line with the position of the customer. The weighted association rules were introduced to discover the hidden association/relations between products with a higher purchase-ability.

Narvekar and Syed [7], proposed and designed a new and an efficient FP-tree association rules mining method which extracts the entire frequent items sets without generating conditional FP-trees. The matter with the existing FP-growth algorithm was elaborated and was stated to have a setback in producing a huge amount of conditional FP-trees which influenced the effectiveness of the algorithm. This new method of association rule mining, scrutinize the database one time which decreases the effectiveness of the algorithm. It also discovered the frequency of the recurrent item sets to discover the needed association rules. However, after applying the new improved FP-tree algorithm, the outcome illustrates that a lot of memory was saved and the number scanning of the database was reduced and no generation of conditional FP-tree.

Rohan and Mishra [19], illustrated the importance of data mining within the customer relationship management by (1) creating an enabling environment between enterprises to understand their customer requirements. (2) trying to find ways to establish trust between the customers and enterprises, and (3) finding the various ways in which the enterprise can retain their valued customers. Data mining plays an essential role in a decision support system. The proposed system uses a kNN algorithm of data mining technique. And this was used to create an unbiased data set to provide less noisy and hazardous information by using the various techniques of classification.

From the study that was carried out by Liao et al. [20], the authors, suggest the consumers' means of consumption by integrating on-line shopping and Business-to-customer home deliv-

ery business model and attempts to use association rules. Association rules were introduced to ascertain the unknown bundling of fresh products and non-fresh products in a supermarket. From the study, two important data mining techniques were used, association rules (Apriori algorithm) which were to mine information from household customers, and k-means algorithm which was also used in cluster analysis to search the clustering of segmentation for probable promotion and sales blueprint in order to market to customer groups. Kaura and Kanga [21], they proposed a periodic algorithm with the intent to improve on the performance of the existing algorithm by not only mining the static data in the database but also providing a new way to capture changes that occur in data in stages. The proposed algorithm was designed to:

- Find out interesting patterns from a large quantity of data.
- Automatically trace the changes in facts from previous data.
- Forecasting future association rules as well as giving the exact methodology to find out outliers. The new suggested algorithm performed on change modeling concept.

An efficient algorithm was proposed by Agrawal et al. [22], for creating important association rules among groups of bought goods. Estimation and pruning methods were used along with the algorithm to circumvent measuring some sets of items whereas assuring fullness. Buffer management was integrated to manage with the problem of inadequate memory during measuring the enormous amount of item sets. Application of the algorithm to sales data obtained from a big retailing company depicted its high efficiency.

According to Chou et al. [23], the researchers put forth an online usage mining procedure to assess user's information and desires for definite product and also forecasts user's target of seeking and purchasing. The sequence mining procedure was introduced to study user's navigation ways on the e-commerce website also to find out the navigation molds. The Artificial Neural Network (ANN) was proposed to get clients' profiles according to the molds that were found out. The purpose was to classify e-customers into clusters with related behavioural molds and also the aptitude to forecast upcoming customers' behaviour related to their purchasing.

The research works so far reviewed did not contain works that explored how customer relationship management system using association rule mining technique was used in analyzing and predicting customer behaviour for an online shopping in conjunction with frequent pattern growth algorithm. However, the proposed system designs and implements a piece of software that automatically gathers and analyzes products which customers purchase together using frequent pattern growth algorithm without the knowledge of the

registered customers. More so, it predicts the purchasing behaviour of the valued customers in order to satisfy them by making sure that the product(s) which they always purchase is available. And also, guides the business owners on strategic decision-making. A chart was generated to identify product(s) that are regularly purchased by the customers and also identify product(s) which might benefit from the promotion.

3. MATERIALS AND METHOD

This research followed an experimental quantitative approach. In developing the system, the researcher made use of the popular software engineering model, the Object Oriented Analysis and Design Methodology (OOAD). It is a software means that represents a system as a set of inter-relating objects. Each object stands for some object in the system being modelled as is typified by its class, its states (data element), and its behaviour. Different replicas can be developed to depict the static or structural view, and dynamic or behavioural view, as well as the runtime implementation of these interactive objects. The unified modelling language (UML) is an example of the symbols that represent these models.

3.1. Method of Data Collection and Analysis

The research adopted the quantitative research methodology. Data were collected through the sales activity of on-line registered customers. First, the interested customers launched the website to view products that the system offers, completed the on-line registration, order for items, make payment and check out.

However, the large dataset were generated in the production environment through the recorded sales of products. A data size of 309 sales of products was collected and was analyzed using frequent pattern growth algorithm, as was shown in the Tables 1 and 2, to;

1. Identify which items that are regularly purchased together;
2. Identify the buying behaviour or pattern of the loyal customers.

3.2. Use Case Diagram of the Proposed System

The use case diagram in Fig. 2, shows the two actors of the system as well as the respective actions they perform on the system.

- The two actors of the system are the system admin and customer.
- The oval shapes on figure denote the use cases or operations, which the actors of the system act on.
- The system admin performs the following activities: logs into the system, creates/edits item categories, adds/edits items, views items, views orders, generates a report, sends item information to customers, changes the password, and logs out.

Table 1: Shows the fp-tree transactional table. Let the minimum support threshold = 3, and minimum confidence = 60%.

Transactional data set (T)	List of items in data base	Support count (frequency)	Frequent pattern (in a decreasing order)	Order of item set (in a decreasing order)
T ID	ITEM SETS			
1	{M,O,N,K,E,Y}	M	K	5
2	{D,O,N,K,E,Y}	O	E	4
3	{M,A,K,E}	N	M	3
4	{M,U,C,K,Y}	K	O	3
5	{C,O,K,I,E}	E	Y	3
	Y	3		
	D	1		
	A	1		
	U	1		
	C	2		
	I	1		

Table 2: Conditional pattern base and conditional frequent pattern tree.

Ordered item set	List of items in an increasing order of priority/ minimum support value	Conditional pattern base	Conditional FP tree	Frequent pattern generated
k,e,m,o,y	Y	{{k,e,m,o:1}, {k,e,o:1}, {k,m:1}}	{k:3}	y = <k,y:3>
k,e,o,y	O	{{k,e,m:1}, {k,e:2}}	{k,e:3}	o = <k,o:3>, o = <e,o:3>, o = <k,e,o:3>
k,e,m	M	{{k,e:2}, {k:1}}	{k:3}	m = <k,m:3>
k,m,y	E	{k:4}	{k:4}	e = <e,k:4>
k,e,o	K			

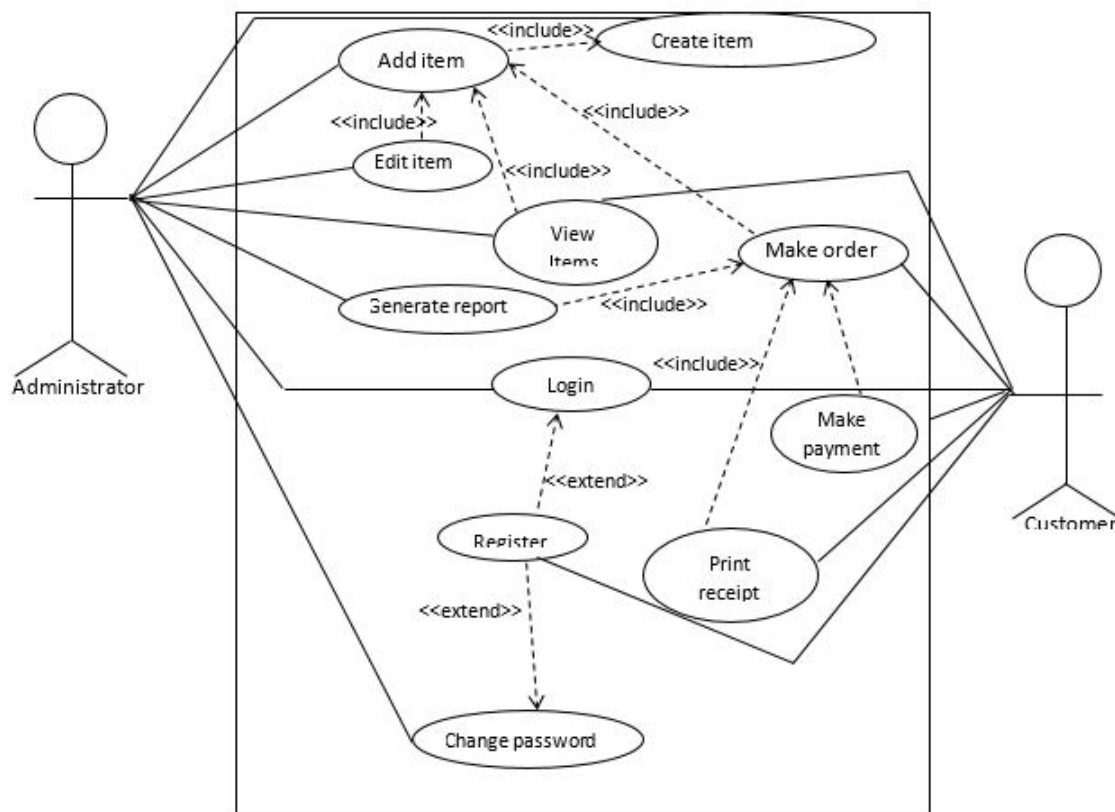


Figure 2: Use case boundary diagram of the new system.

- A customer when on the website views items, registers and/or logs in to the web application to order for items, checks out, makes payment, prints receipt, views information on items the site is promoting and logs out.

3.3. Algorithm Design for the Model

An algorithm is simply a process or a set of rules to be followed step by step when executing an operation. The FP-growth algorithm used in this work is based on the pseudo code in Section 3.3.

Frequent pattern- growth algorithm

The frequent pattern growth algorithm is simply a different way of discovering frequent item sets without using candidate generations, therefore enhancing the performance. However, a divide-and-conquer method was used to actualize the functionality of the algorithm. The foundation of this approach is the procedure of the particular data structure known as frequent pattern tree (FP- tree), which keeps hold of the item set association information. However, this algorithm works as follows [24]:

1. Scan the transaction database once so as to search for all the frequent items and their support.
2. Arrange the frequent items in decreasing order of their support.
3. At first, begin building the FP-tree with a root "null".
4. The primary transaction is obtained from the transaction database. Remove all non-frequent items and file the remaining items according to the order of sorted frequent items.
5. Use the transaction to build the primary branch of the tree with each node matching to a frequent item and showing that item's frequency, which is one for the primary transaction.
6. Obtain the subsequent transaction from the transaction database. Remove all non-frequent items and list the remaining items according to the order sorted frequent items.
7. Put in the transaction inside the tree using any frequent prefix that may come out. Augment the items counts.
8. Carry on with step 6 pending when all transactions along with the database are processed.

3.4. FP-Tree Algorithm Approach

The FP-Growth algorithmic program for mining frequent pattern with FP-tree by pattern fragment growth is [24]:

```

a. Input: a FP-tree formed with the above
    cited algorithm;
D-Transaction database;
S- Minimum support threshold.
Output: the complete set of frequent
patterns.
Method: decision FP (FP-tree, null).
Procedure FP (Tree, A)
{
  If Tree has a-one path P, then for each
  permutation (denoted as B) of the nodes
  among the trail P do generate mold B U A
  with sup = minimum support of nodes in B else
  For each ai among the header of the Tree do
  {
    Generate mold B= ai U A with sup =
    ai.support;
    Construct B's cond pattern base and B's
    cond FP- tree
    Tree B;
    then decision FP(Tree B, B)
  }
}

```

Figure 3 shows the FP-Tree algorithm flowchart as depicted in the pseudo code above.

4. RESULTS AND DISCUSSION

In the following section we present the results and discussion from the experiment.

4.1. Frequent Item Set Report

Figure 4 depicts the association rule. It shows how frequent a particular item is being purchased together with another item. It also shows the likelihood of affiliation between data items within large data sets in various types of databases.

4.2. Generated Chart

In the graph in Fig. 5, each bar represents the amount of sales for a particular product. The difference in height of the bars depicts or show how often customer(s) purchase these products. From this idea, it helps to identify the product(s) which might benefit from promotion and also helps the admin for decision making.

4.3. Illustrating the FP-tree Workings

Consider the following transactions when an online supermarket tracks sales data for eleven items denoted by 'M', 'O', 'N', 'K', 'E', 'Y', 'D', 'A', 'U', 'C', 'I'. The achieved results by the implementation of the fp-growth algorithm for association rule in online shopping is discussed below.

First scan

From the frequent pattern tree (Fig. 6), conditional pattern bases and conditional frequent pattern tree are generated. Conditions for generating conditional pattern base from the items:

1. Start from the mode with minimum support value
2. Exclude the mode with most support value

Table 3: Association rules generated from the conditional pattern tree.

Association rule
$K \Rightarrow Y$
$Y \Rightarrow K$
$EO \Rightarrow K$
$K \Rightarrow EO$
$KO \Rightarrow E$
$E \Rightarrow KO$
$KE \Rightarrow O$
$O \Rightarrow KE$
$M \Rightarrow K$
$K \Rightarrow M$
$E \Rightarrow K$
$K \Rightarrow E$

Table 4: Rating scales for assessing exposure.

Association rule	Support count	Confidence
$K \Rightarrow Y$	3	60%
$Y \Rightarrow K$	3	100%
$EO \Rightarrow K$	3	100%
$K \Rightarrow EO$	3	60%
$KO \Rightarrow E$	3	100%
$E \Rightarrow KO$	3	75%
$KE \Rightarrow O$	3	100%
$O \Rightarrow KE$	3	100%
$M \Rightarrow K$	3	100%
$K \Rightarrow M$	3	60%
$E \Rightarrow K$	4	100%
$K \Rightarrow E$	4	80%

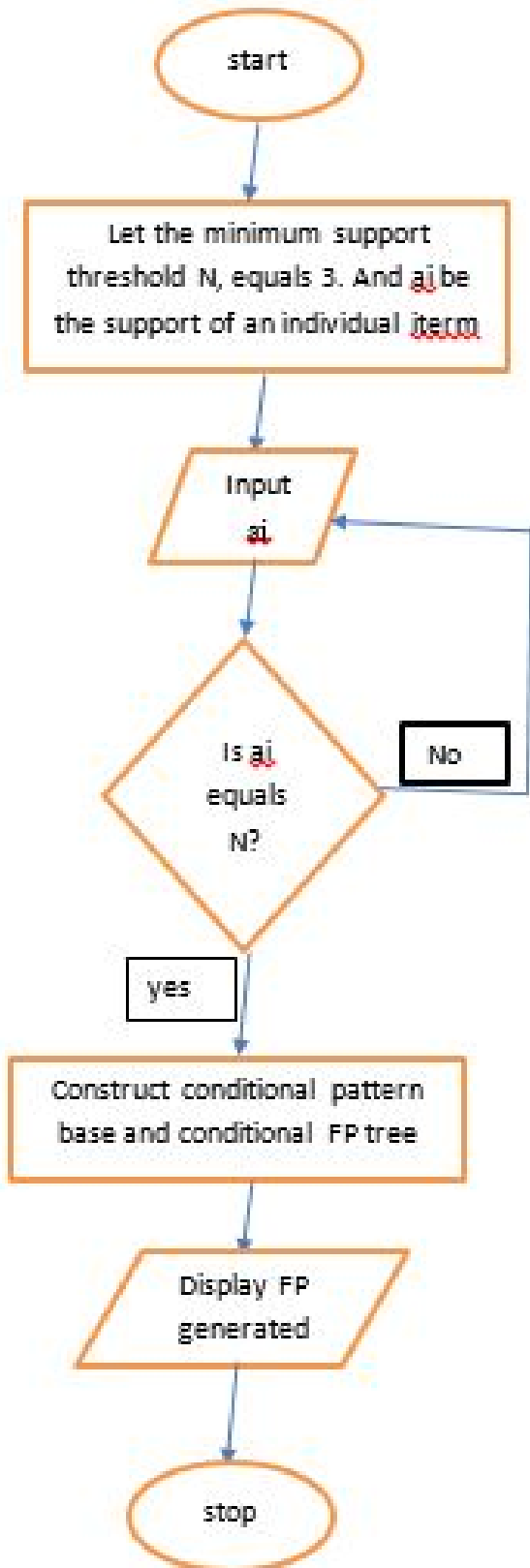


Figure 3: FP-Tree algorithm flowchart.

From the Table 3, the rule $(K \Rightarrow Y)$, portrays or expresses association rule which is an imperative technology in data mining that can be used to analyze and predict customer purchasing behaviour. It is also expressed as if/then statement which means if K then Y.

This shows that, there is a strong association that connects the purchasing of Item **K** and **Y**, where **K** depicts Antecedence, and **Y** depicts consequent. And they are disjoint item sets expressed as, $K \cap Y = \phi$. This can be helpful to organizations to understand the purchasing pattern of their valued customers because a customer, who purchases item **K**, can equally purchase item **Y**.

The potency of an association rule can be measured in terms of its support and confidence, as shown in Table 4. Support is the percentage of transaction in **T** that contains $K \cup Y$, [25]. It describes or denotes the probability that contains both **K** and **Y** in the whole records in the data store. It determines how frequently a rule is relevant to a given dataset. The count for an item increases by one for each transaction in the database during the scrutinizing process. Mathematically, the support count α , for an item set k can be declared as shown in Eq. (1) [26]:

$$\alpha(k) = |\{ti \mid k \subseteq ti, ti \in T\}| \tag{1}$$

where the symbol $|\cdot|$, indicate the number of elements in a set. And **T** indicates the set of the

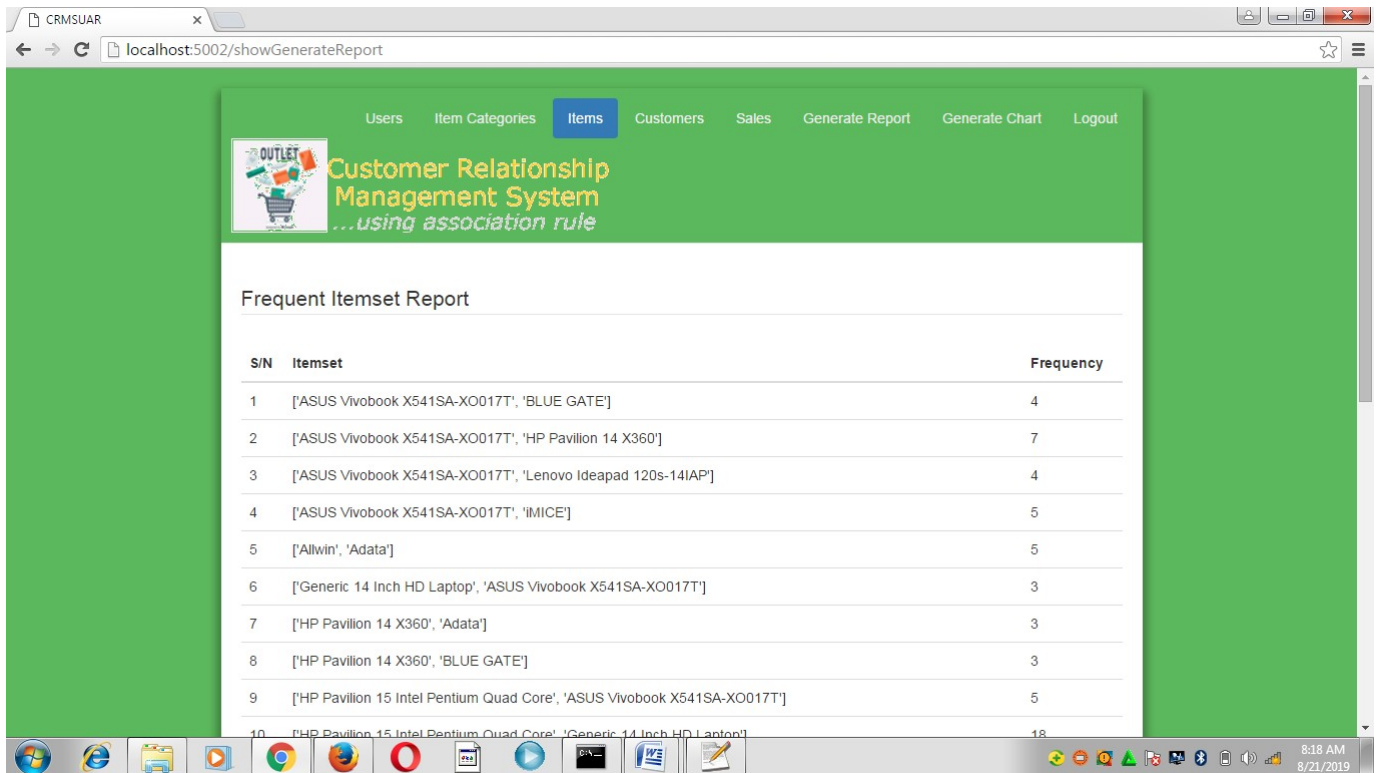


Figure 4: FP-Tree algorithm flowchart.

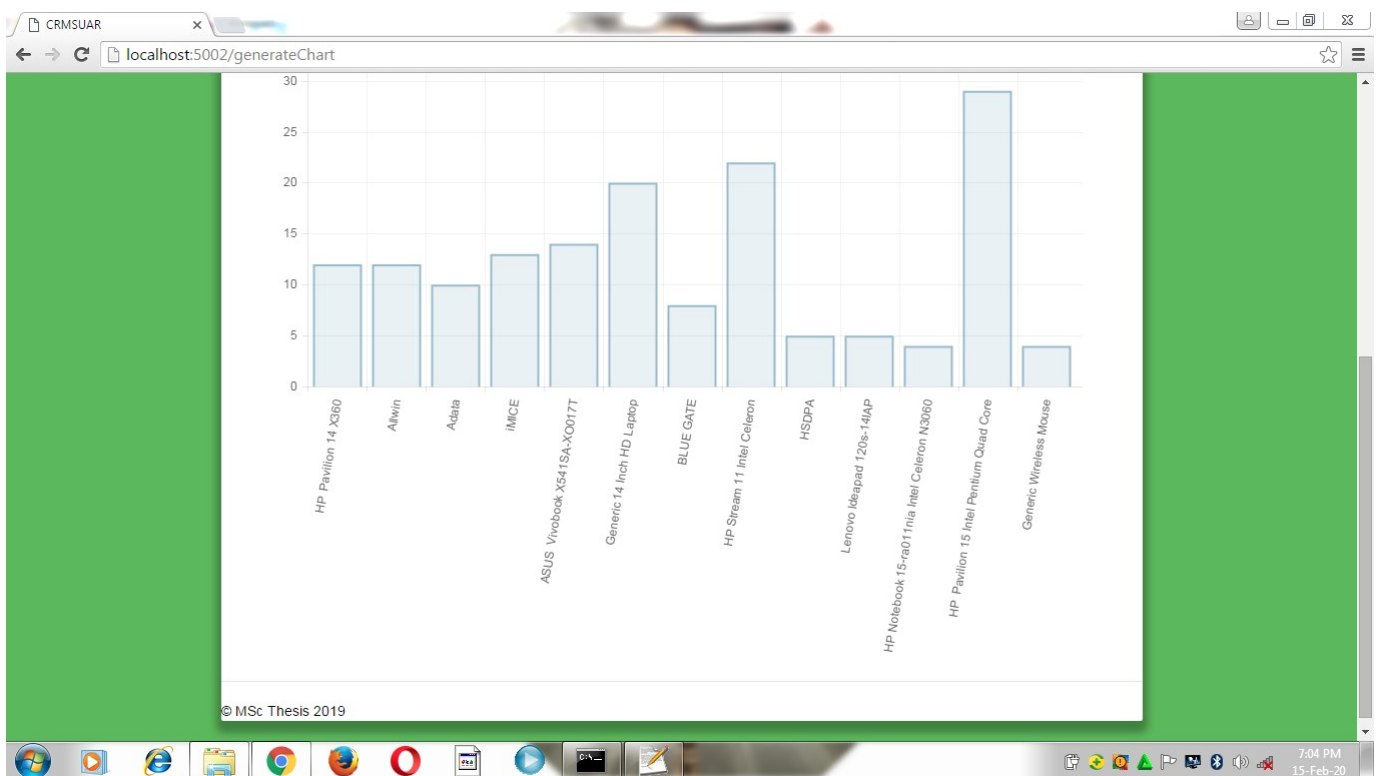


Figure 5: Chart showing the amount of items sold within a particular period.

Second scan.

Item ID	Order of item set	Support count	N. link
K	K,E,M,O,Y	5	
E	K,E,O,Y	4	
M	K,E,M	3	
O	K,M,Y	3	
Y	K,E,O	3	

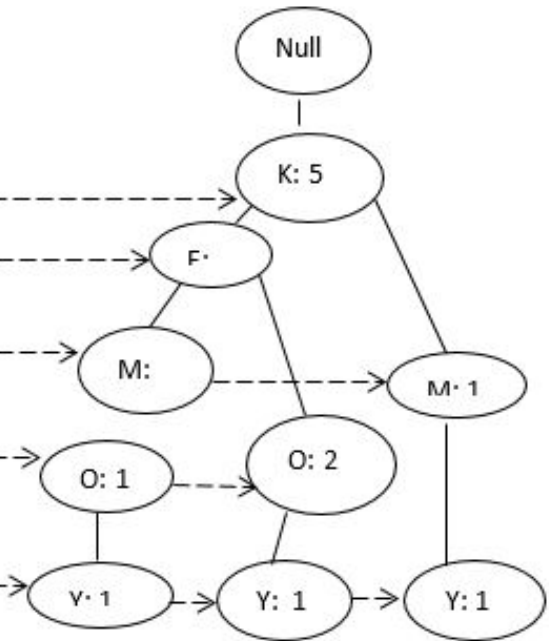


Figure 6: Frequency pattern tree.

entire transactions in the database, while ti indicates the individual transaction contains a subset of items chosen from the dataset.

Support for an item $\left(\frac{k}{y}\right)$ is evaluated by the following principle shown in Eq. (2): Let support count k/y be α therefore,

$$\text{Support} = \left(\frac{k}{y}\right) = \frac{\alpha}{t} \times 100 \tag{2}$$

Where t represents the whole transactions in the database.

However, support $\left(\frac{k}{y}\right) = \frac{3}{5} \times 100 = 60\%$

From the expression above (Eq. (2)), the support count of K and Y in the transactions is 3, i.e. (60%). This is because there were only three transactions that contain the two items in the total transactions. However, if the support of the items K and Y is below 3 or (60%), it shows that these items are self-standing in the database. The organization may not pay much attention to such goods that are not frequently bought and this may not be lucrative to encourage the business. Obviously, when the value of the support count is higher or equals 3 or (60%), it portrays that the items are joined and high support is preferred for more exciting association rules.

Confidence can say to be a percentage/fraction of the number of transactions that covers $K \cup y$ to the total amount of records that have K , where if the percentages are more than the limit of confidence, associations rule $K \Rightarrow Y$ can be formed as in Eq. (3). Let the support k be β therefore,

$$\text{Confidence} = \left(\frac{k}{y}\right) = \frac{\alpha}{t} \times 100 \tag{3}$$

β represent the number of transaction that contains k .

$$\text{Confidence} = \left(\frac{k}{y}\right) = \frac{3}{5} \times 100 = 60\%$$

Confidence determines or denotes the probability that if transactions containing K can also contain Y and however, it evaluates the strength of the association rules. From the expression in (Eq. (3)), it shows that the confidence of the association rule $K \Rightarrow Y$ is 60 which depicts that the probability of getting item Y when the item K has already been gotten is 60%.

Association rules using frequent pattern algorithms for the analysis is an essential tool that helps the organization upsurge the efficiency of marketing and sales policies by mining products data collected from past sales activities of on-line registered customers. The outcome of the above-calculated association rules depicts which items are positioned on the shelves together which would upsurge the proceeds with the products which are frequently purchased. Therefore, it is so helpful to organizations to understand their esteemed customers based on the preceding data collected from the past sales activities, predicts and understand their purchasing pattern that will help them for a good decision making which has the direct link to maximizing business profit.

5. CONCLUSION

The importance of companies managing the association or interactions with their customers is very critical in the expansion of clientele base which has a direct impact on increasing business turnover. The purpose of this study, was

to investigate and identify which items are regularly purchased together by customers and based on this result, it analyzes and understands the customer's buying patterns. Frequent pattern growth algorithm which was very helpful in obtaining the most excellent association rule was also incorporated in the system. It also helps to extract frequent item sets directly from the transactional database. The result of this study shows that data mining methods are capable of providing effective tools for understanding the hidden pattern between items in a huge transactional database.

6. RECOMMENDATION

Implementing a customer relationship management system using association rule software is highly recommended for every online shopping mall in Nigeria. This software, will promote growth and increases efficiency and profitability across the entire business value chain. However, the system will also enable the organization to have the opportunities of guaranteeing better relations between companies and their customers, that can lead to increased sales through better timing due to anticipating needs based on historic trends of customers' orders.

It will reduce product cost substantially as long as the product specs are according to customer's needs so that changes are limited. However, it will enable firms to gain a good reputation and also helping companies to boost profitability by establishing and maintaining strong relationships with their clients. In retaining customers, there will be less need to spend money on creating awareness to find and acquire new customers.

References

- [1] (2012) Get app & thinkaboutcrm. [Online]. Available: <https://www.getapp.com/download/customer-relationship-management-small-business.pdf>
- [2] G. Gupta and H. Aggarwal, "Improving customer relationship management using data mining," *International Journal of Machine Learning and Computing*, vol. 2, no. 6, p. 876, 2012.
- [3] G. Krishna and R. Vadlamani, "Evolutionary computing applied to customer relationship management," A survey article in engineering application of AI, 2016.
- [4] A. Rai. (2018, May) Association rule mining: An overview of association Rule mining and its application. [Online]. Available: <https://www.upgrad.com/blog/association-rule-mining-an-overview-and-its-applications/>
- [5] Techonestop. (2018, May) What is analytical crm. [Online]. Available: <http://techonestep.com/what-is-analytical-crm>
- [6] M. Žekić-Sušacand and A. Has, "Discovering market basket patterns using hierarchical association rules," *Croatian Operational Research Review*, vol. 6, pp. 475–487, 2015.
- [7] M. Narvekar and S. Syed, "An optimized algorithm for association rule mining using FP tree," in *International Conference on Advanced Computing Technologies and Applications (ICACTA)*, 2015.
- [8] G. Suchacka and G. Chodak, "Using association rules to assess purchase probability in online stores," vol. 15, no. 3, pp. 751–780, 2017.
- [9] M. Mak, T. George, and S. Ting, "A financial data mining model for extracting customer behavior," *International journal of engineering and business Management*, vol. 3, no. 3, pp. 59–72, 2011.
- [10] S. Beom-Soo and S. Yong-Moo, "CRM Strategies for A Small-Sized Online Shopping Mall Based on Association Rules and Sequential Patterns," in *Association for Information Systems AIS Electronic Library (AISeL) PACIS 2010 Proceedings Pacific Asia Conference on Information Systems*, vol. 39, no. 9. Expert System Appl, 2012, pp. 7736–7742.
- [11] P. Dhandayudamand and I. Krishnamurthi, "Enhanced rule induction algorithm for customer relationship management," *An International Journal, Applied Mathematics & Information Sciences*, vol. 7, no. 4, pp. 1471–1478, 2013.
- [12] P. Isakki, A. Devi, and S. Rajagopalan, "Analysis of customer behavior using clustering and association rules," *International Journal of Computer Applications*, vol. 43, no. 23, pp. 0975–8887, 2012.
- [13] S. Changchien and T. ChuenLu, "Mining association rules procedure to support on-line recommendation by customers and products fragmentation," *Expert Systems*, vol. 20, no. 4, pp. 325–335, 2001.
- [14] B. Njoni, J. Kenyatta, and A. Kibe, "Social media analytics for human resource management using association rule," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 6, no. 10, 2016.
- [15] Y. Peng, "An algorithm of commodities association rules mining in e-commerce based on rough set," in *International Conference on Internet Technology and Applications*. New York: IEEE iTAP'10, 2018, pp. 1–3.
- [16] M. Masoumeh and M. Mahdavi, "Providing a model for predicting tour sale in mobile e-tourism recommender systems," *International Journal of Information Technology Con**, vol. 2, no. 1, 2012.
- [17] H. Kaur and K. Singh, "Market basket analysis of sports store using association rules," *International Journal of Recent Trends in Electrical & Electronics Eng.*, vol. 3, no. 1, pp. 81–85, 2013.
- [18] "Market basket analysis of sports store using association rules," *Incremental weighted mining based on RFM analysis for recommending prediction in u-commerce*, vol. 7, no. 6, pp. 133–144, 2013.
- [19] A. Rohan and K. Mishra, "Improving customer relationship management using data mining," *International Journal of Scientific & Engineering Research*, vol. 7, no. 2, 2016.
- [20] S. Liao, Y. Chen, and Y. Lin, "Mining customer knowledge to implement online shopping and home delivery for hypermarkets," *Expert Systems with Applications*, vol. 38, pp. 3982–3991, 2011.
- [21] M. Kaura and S. Kanga, "Market basket analysis: Identify the changing trends of market data using association rule mining," in *International Conference on Computational Modeling and Security*, vol. 85. Procedia Computer Science, 2016, pp. 78–85.
- [22] R. Agrawal, T. Imielin'ski, and A. Swami, "Mining association rules between sets of items in large databases," in *International Conference on Management of Data*, vol. 22, no. 2, New York, 1993, pp. 207–216, SIGMOD'93.
- [23] P. Chou, P. Li, K. Chen, and M. Wu, "Integrating web mining and neural network for personalized e-commerce automatic service," *Expert System with Application*, vol. 37, no. 4, p. 2898–2910, 2010.
- [24] M. SMythili and S. Mohamed, "Integrating web mining and neural network for personalized e-commerce automatic service," *Performance evaluation of apriori and FP-Growth algorithm*, vol. 79, no. 10, 2013.
- [25] W. Samaraweera, C. Waduge, and R. Meththananda, "Market basket analysis: A profit based approach to apriori algorithm," in *9th International research conference-KDU*, Sri Lanka, 2016.
- [26] M. Tan, P. Steinbach, A. Karpatne, and V. Kumar, *Introduction to data mining*, 2nd ed., ser. 978-0-13-312890-1. Pearson Education, 2018.