

Automated Market Basket Analysis System

Izang A. A.

Department of Computer
Science and Information
Technology, School of
Computing and Engineering
Sciences, Babcock University,
IlishanRemo, Ogun State

Okoro U. R.

Department of Computer
Science and Information
Technology, School of
Computing and Engineering
Sciences, Babcock University,
IlishanRemo, Ogun State

Olarewaju T. B.

Department of Computer
Science and Information
Technology, School of
Computing and Engineering
Sciences, Babcock University,
IlishanRemo, Ogun State

Fasanu T. D.

Department of Computer Science and Information
Technology, School of Computing and
Engineering
Sciences, Babcock University, IlishanRemo, Ogun
State

Adeyinka A.

Department of Computer Science,
School of Computing and Engineering
Sciences, Babcock
University, IlishanRemo, Ogun State

ABSTRACT

Market Basket Analysis (MBA) is a widely used technique among marketers to identify the best possible combination of products or services frequently bought by customers. Market Basket Analysis is one of the data mining techniques used in recent times to know the correlation between one item to another purchased. The problem of determining customers preference in terms of items purchased was focused on the traditional and heuristics algorithms with limited factors in the past. However in recent times through this study, building an automated basket analysis system, will help shop owners identify customers purchasing behavior, patterns and identify the relationship between products and item purchased in order to maximize profit through the use of association rule mining. Association rules is one of the data mining techniques which is used for identifying the relationship between one item to another. Association rule is the bedrock of a market basket analysis system as it helps to determine the correlation that exist between the items purchased. An automated MBA system was implemented through the use of the spiral model development model and a combination of HTML, PHP and MySQL as the programming environment to make this system web-based. The Automated Market Basket Analysis System would improve on search methodologies that can also be of help in generating recommendations for consumers though the association rule mining algorithm embedded in the system. The provided results reveal that the obtained solutions seem to be more realistic and applicable.

Keywords

Market Basket Analysis, Association Rule Mining, Data Mining, and Recommender systems.

1. INTRODUCTION

Market Basket Analysis (MBA) was coined from a supermarket scene where a customer enters and picks a basket to start shopping. The goal of a MBA system is to predict the customers buying pattern and possibly the goods in the superstore within the categories that the customer will like to pick and useful information to support retail marketing decisions [1].

Market Basket Analysis ultimately results in a better understanding of the customers and their purchasing behavior, allowing retailers to explore associations, predict the

likelihood of a customer response based on associations to maximize profit for the retailers by providing better services to the consumers and ultimately optimizing marketing and sales operations for results.

MBA in the end effects a better understanding of your clients and their buying behavior, allowing you to explore institutions, predict the probability of a purchaser reaction based on associations to maximize earnings for the shops by means of supplying better offerings to purchasers and ultimately optimize your marketing and income operations for better profit making [2].

This study focuses on building an automated system that will predict customers buying pattern. Also, the system should be able to make recommendations the next time the customer comes online to make purchases.

2. WHAT IS MARKET BASKET ANALYSIS

Market Basket Analysis System is a widely used technique among the Marketers to identify the best possible combinatory of the products or services which are frequently bought by the customers. This is also called product association analysis. Association analysis is mostly done based on an algorithm named “Apriori Algorithm” [3]. The Outcome of this analysis is called association rules. Marketers use these rules to strategize their recommendations. Market Basket Analysis (MBA) was coined from a supermarket scene where a customer enters and picks a basket to start shopping. The goal of a Market Basket Analysis System is to predict the customers buying pattern and possibly the goods in the superstore within the categories that the customer will like to pick and useful information to support retail marketing decisions [4]. For example, a customer who buys bread is likely to buy butter. Therefore, it is like predicting what items a customer will pick alongside. This study focuses on building an automated system that will predict a customer’s buying pattern. Also, the system should be able to make recommendations the next time the customer comes online to make purchases. Before an automated system can be built, an algorithm is needed that will help with indicating the level of association between the items picked and that algorithm is called Association Rule Mining Algorithm.

3. RELATED WORKS

Adoptive association rule is a method of market basket analysis based on faster rule generation algorithm. This method proposed by [5] was used in mining multi-measurement rule, depending on Adaptive Genetic Algorithm (AGA) with crossover matrix and mutation matrix. Dynamic adaptive support Apriori (DAS Apriori) was also developed as a technique to compute the minimal help for obtaining class association rules.

K-apriori algorithm, which is a modified version of Apriori algorithm was implemented in Market Basket analysis for a grocery store primarily based on frequent Itemset Mining [6]. It addresses customer business by the use of clustering belongings, while generating informative frequent itemsets and association rules for the store. This algorithm was efficient because it was able to classify clients into specific clusters which include the use of purchaser hobby profiles and past items purchased by the client. In K-apriori algorithm set of rules, the records are clustered such that the customers are categorized and then the clusters common itemsets are generated.

This paper focuses on a survey of several data mining algorithms for MBA carried out by [7]; the evaluations of these algorithms showed that using association rule mining algorithm will help in better classification of large quantity of dataset while Apriori set of rules is determined to be better for association rule mining. It is recommended for revealing

relationships in a large database with very large amount of data which would help managers to recognize the set of items client are possibly going to buy.

Associative classification as an emerging trend in data mining proposed by [8] was discussed in this paper, it proves to be more efficient than traditional classification and association algorithm in terms of accuracy and efficiency. Association classification is the integration of association and classification where association rules are used in training step of classification, all associative type algorithms use an assist threshold to generate association regulations

4. METHODOLOGY

The software development model that was used is the spiral model. This model was used because it allows changes and additional functionality at a later time which makes development fast and more features will be easy to add and it creates space for customer feedback.

WAMP was used for the development of the automated market basket analysis system which has the following:

- Windows operating system
- Apache web server.
- MySQL to create the database for the application using PHPMyAdmin

HTML, CSS, PHP was used to design the web pages.

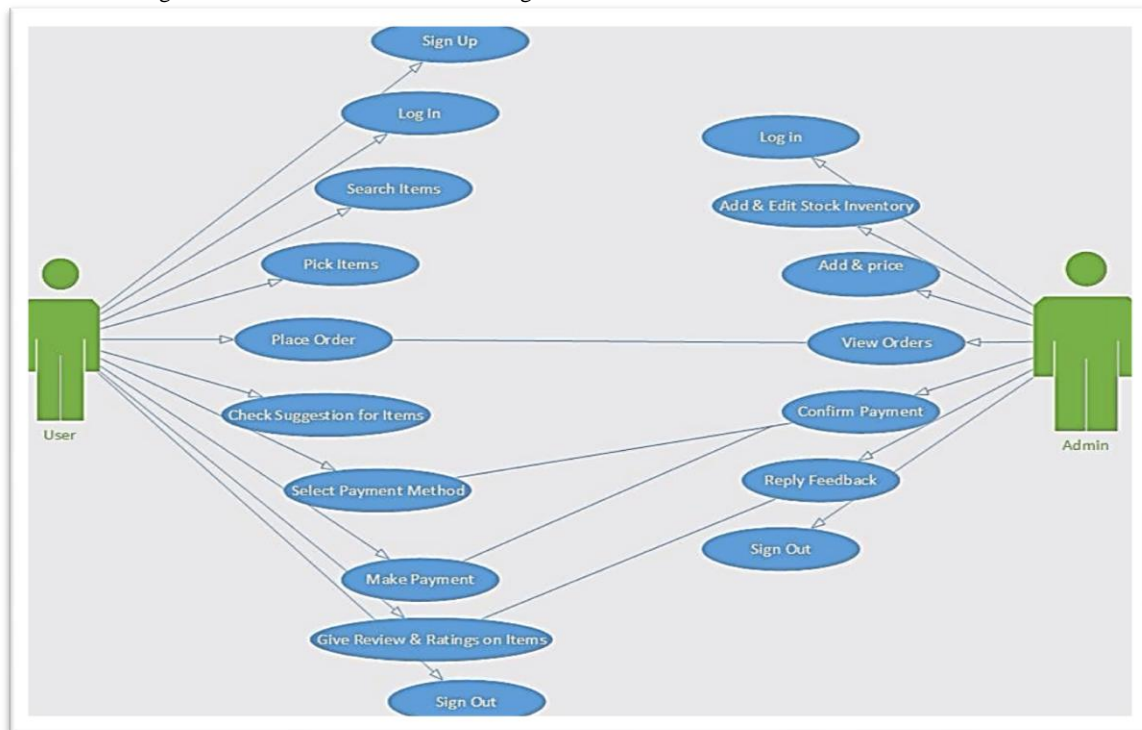


Fig.1 The use case diagram of the system showing the actors and their functions

There are two actors in the figure 1 above, which is the user and admin and their functions includes:

Users:

This are the customers that will use the system or the platform. They will be able to register and see products on the platform. And they will able to the following:

- i. Register and login
- ii. Search for goods available

- iii. View goods
- iv. See previous recommendation by other customers that have used the platform
- v. Select goods
- vi. Make transaction through the required means.
- vii. Print receipt
- viii. Send feedback

Admin:

This actor monitors all transactions and events happening on the platform. He is registered by default and can create cashiers who will monitor and see all orders coming in, and he is also able to the following

- i. Create new cashiers
- ii. Login
- iii. View all customers orders
- iv. Edit and add stock inventory
- v. Edit prices of goods
- vi. See all reviews and feedback
- vii. See approval of all payment made.

5. IMPLEMENTATION OF THE MARKET BASKET ANALYSIS SYSTEM

The system designed comprises of the front-end and the back-end. The front-end of the system includes the interface of the system, i.e. what the users can see. The front-end components include the various pages of the system i.e. Dashboard, Administrator’s module page, the Product page, the Stock edit page, Order page, Shopping basket page. Furthermore, these pages (front-end) are linked to the back-end component which contains the database of the system. Using PHP scripting language, the database was connected to the front-end.

5.1. Dashboard

A Dashboard is normally the primary web page a tourist navigating to a website from an internet browser will see, and it could also function as a landing page to draw visitors .The dashboard is used to facilitate navigation to other pages on the website.

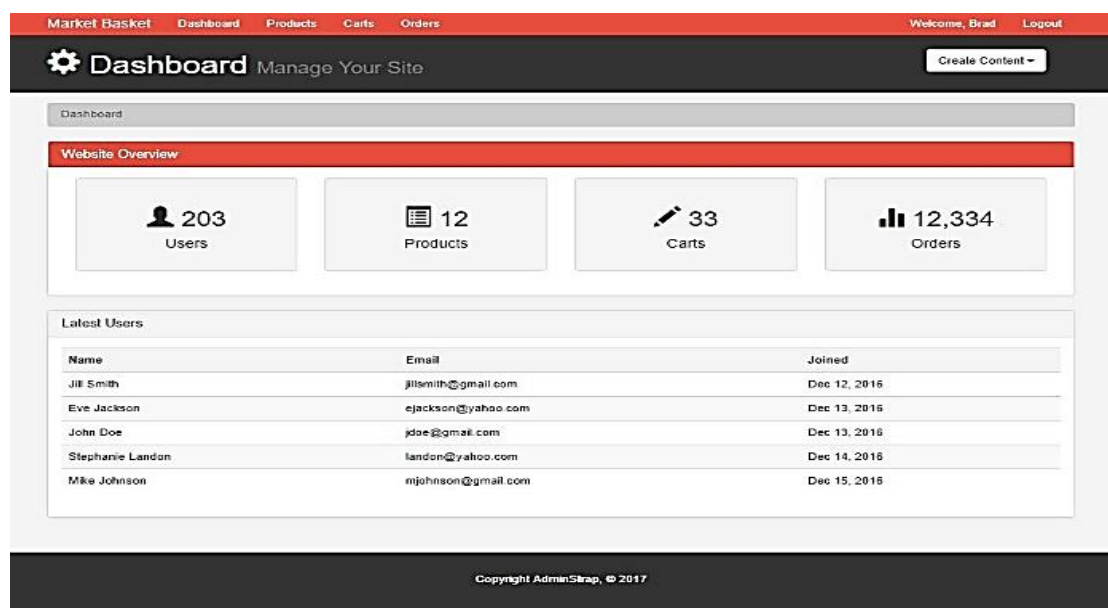


Fig.2: Dashboard.

5.2 Administrator’s Module

An administrator is a special user who has the authority to make changes on the software and effects that will affect

customers as an update when they login. Several or various administrators can be employed in the system with the task of login, manage orders

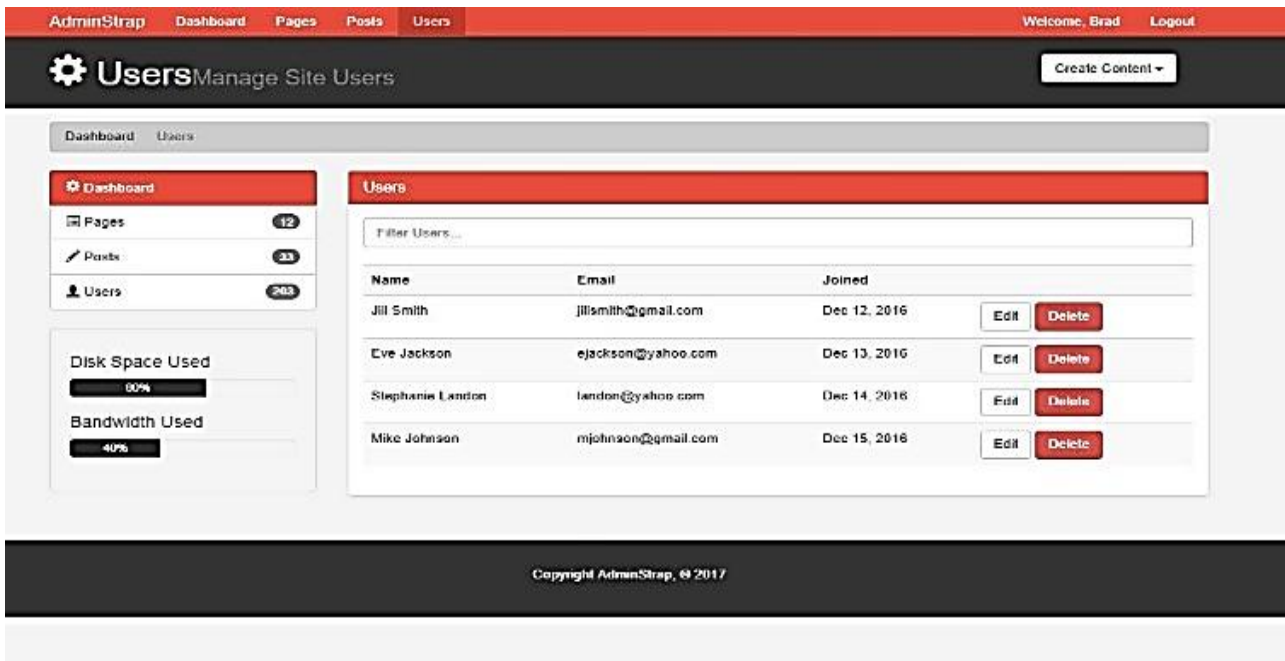


Fig.3 Administrator Module page

5.3 Product Page

This product page allows the admin to add and view items in stock. It also allows the admin to add them to the product page

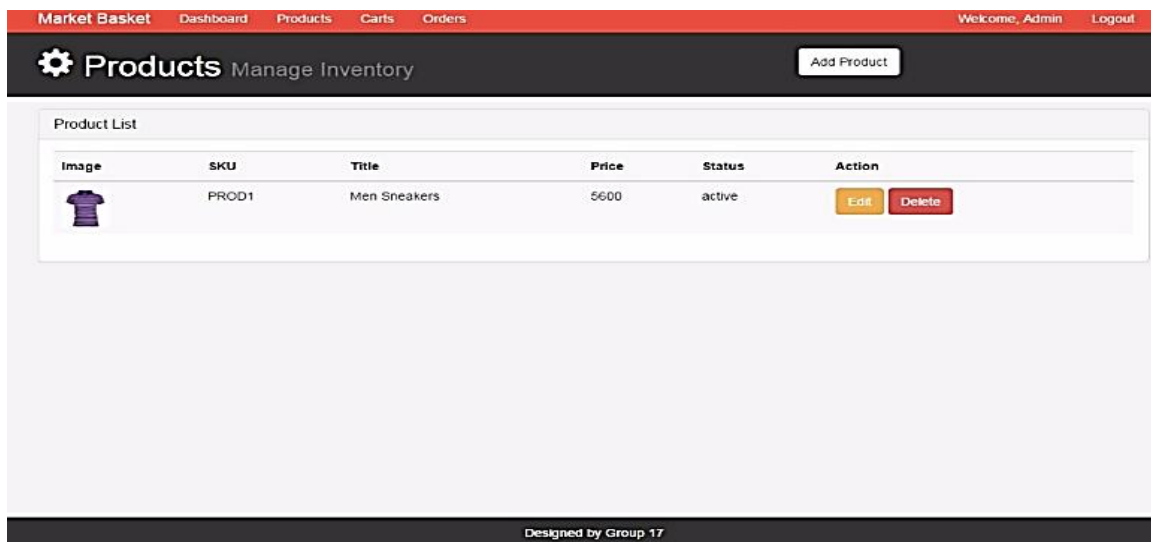


Fig.4 Product page

5.4. Stock Edit Page

This page allows the user to edit and add description for an item in stock.

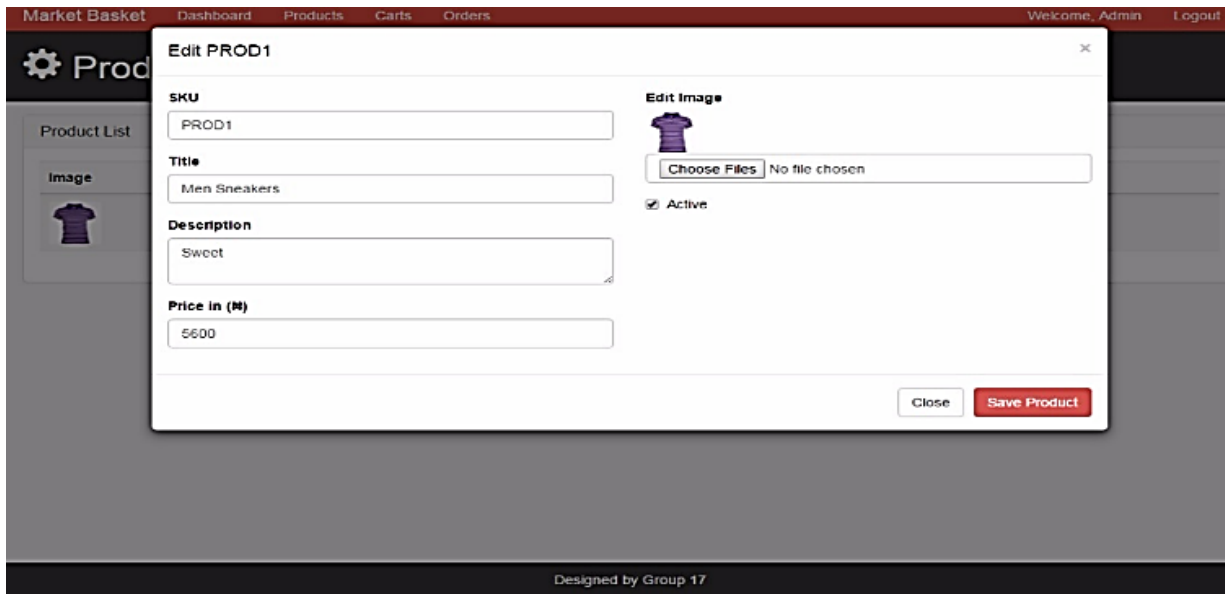


Fig.5 Stock edit page

5.5 Order Page

This page allows the user to see the order details of customers and provides a link to the order confirmation/summary page

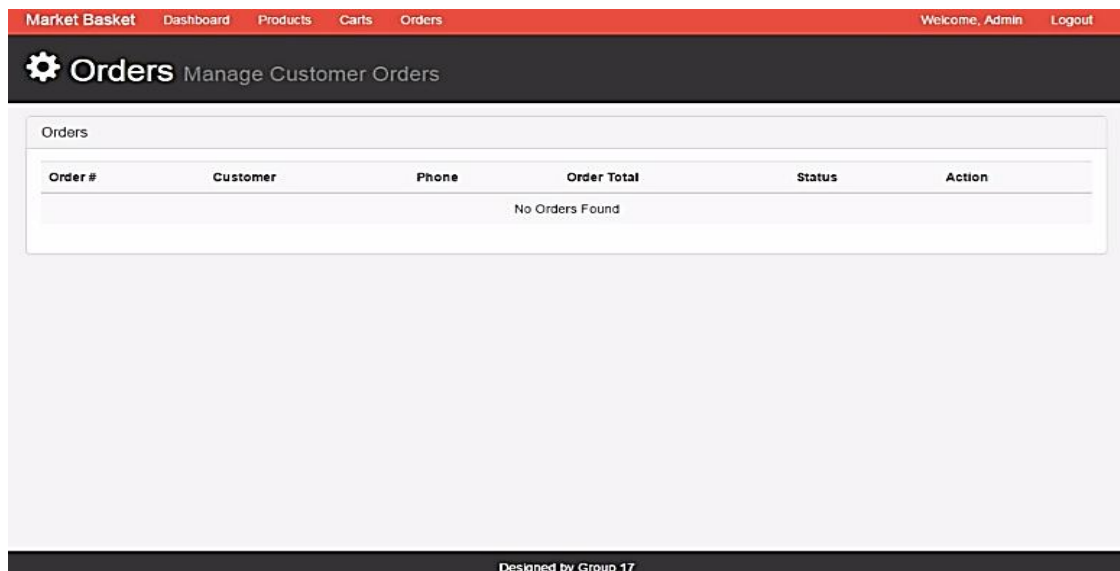


Fig.6 Order page

5.6 Shopping Baskets

This page contains the shopping basket and allows the user to manage customer order process

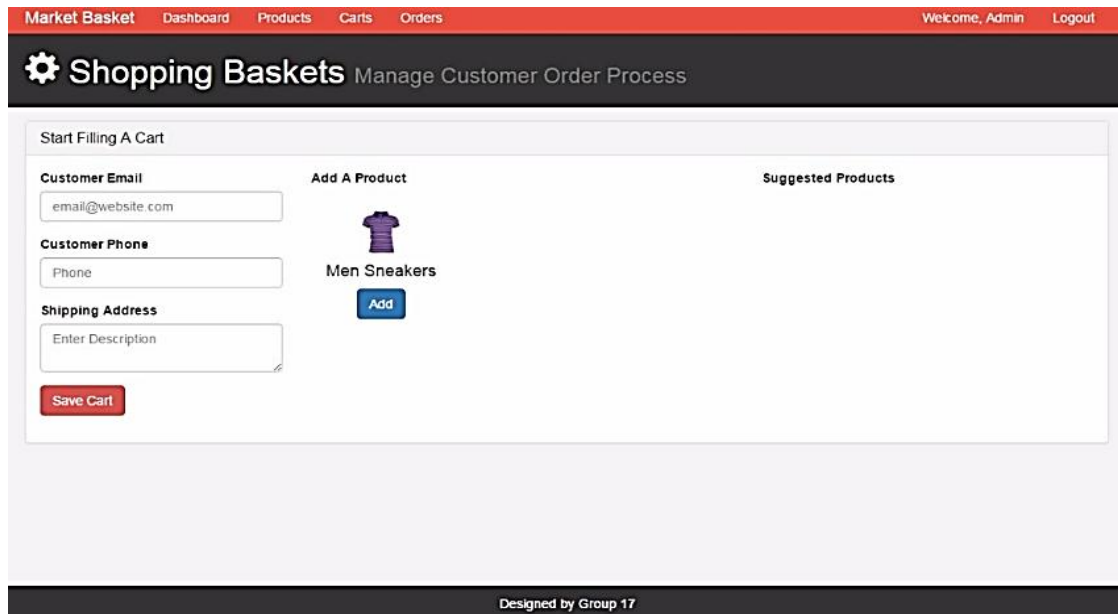


Fig.7 shopping basket page

6. OPERATION OF THE AUTOMATED MARKET BASKET ANALYSIS SYSTEM

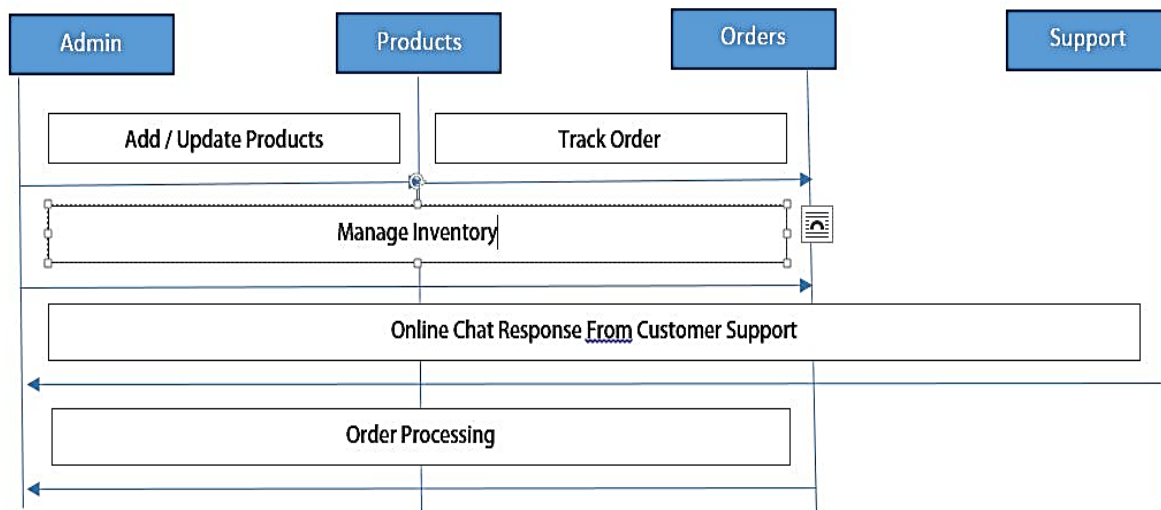


Fig 8. Admin sequence diagram

The diagram above explains a workflow of the admin sequence of the automated market basket analysis systems, this is also the backend of the system, that specify the operation of this system.

Use Case: The actor in this use case is the Admin and he is able to perform actions which include

- i. Add/update products: This allows the admin to add new items to the inventory and also update or change attributes of items in the inventory, like price, quantity available, size e.t.c
- ii. Track orders: This enables the admin to be see orders and individual shopping carts of customers
- iii. Manage inventory: admin can manage inventory of stocks in a store by knowing the amount and frequency in which a particular item in the inventory is bought, this also enables the decision making process in the product recommendation.

iv. Chat with customers: A platform for customer to admin interaction is also built into the system.

There are also several entities which are attributed to this admin, which include:

- i. Products: these the objects that we are identifying associations between, an item being purchased is treated like a binary variable and is given the value of 1 for purchase in a transaction and 0 otherwise. The binary representation, however, does not account for the quantity of item in the transactions, only its presence/ absence.
- ii. Orders: These are carts that have been checked out by the customers, in market basket analysis we can also refer to it as itemset, this is the item set in which the recommendation algorithm works on.
- iii. Support: This includes everything from item reviews to customer-seller interaction on the application.

The automated analysis is done by the system during the Order processing. First a training set is inputted to the recommendation algorithm which produces a recommendation model that can be used to generate new predictions. To evaluate the model a held out test set is fed to the learned model where predictions are generated for each user-item pair. Predictions with known labels (true value) are then used as an input to the evaluation algorithm to produce evaluation results.

The recommendation algorithm is used to identify item associations with high confidence or lift to generate rules. Finding rules with high confidence or lift is less computationally taxing once high-support itemsets have been identified, as confidence and lift values are calculated using support values of the remaining frequent itemsets

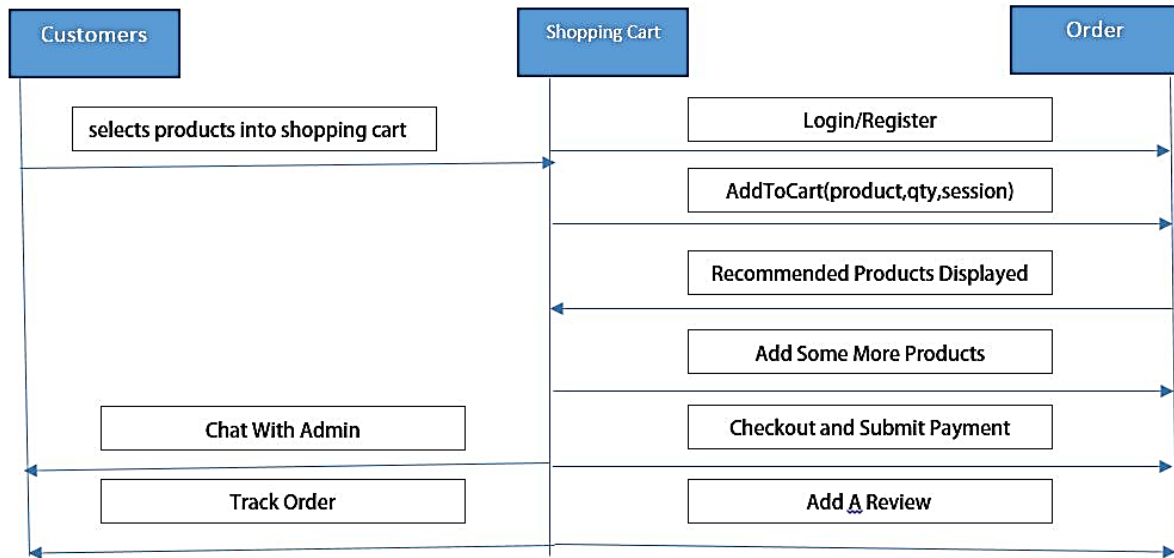


Fig 9. Customers sequence diagram

The diagram above explains a workflow of the customer sequence of the automated market basket analysis systems, this includes all webpages that are available to the customer, and also serves as shopping platform where customers can buy items.

Use Case: The actor in this use case is the Customer and he is able to perform actions which include

- i. Login/Register : there is a page that allows users to sign up, and subsequently gives users access to the system, this action allows users to have full access to pages allowed by the admin on the system.
- ii. Add products to shopping cart: There is a personalized cart for temporarily holding items customers are willing to buy, this action allows users to add as much items as possible to the cart based on the limit set by the admin in the inventory.
- iii. Select recommended items: customers can select and add items suggested by the recommendation engine to their shopping cart.
- iv. Checkout/Payment: The payment system allows the system to accept payment for items in the shopping cart, and automatically logs receipt for confirmed payments.
- v. Add Reviews: This action allows users to rate items and write reviews for confirmed items that a customer has purchased. This allows the system to better utilize its recommendation algorithm.
- vi. Chat with admin: This action allows users to interact with the seller through messages.

- vii. Track Orders: customers can check previous orders placed on the page, the recommendation engine also looks at previous orders so as to more accurately generate recommendations

There are also several entities which are attributed to this user, they include:

- i. Items: These are stocks available in the inventory.
- ii. Shopping Cart: This is temporary store of items which the customer is attempting to buy, items in the cart are limited to the stock in the inventory created by the admin
- iii. Orders: This comprises of a list of checked out items of a customer that their payments have been confirmed.
- iv. Support: This is compromises of the review/rating system for items bought by a customer, and also the platform for interacting with the seller.

The recommended items generated for the customer is based on the recommendation algorithm used on order processing in the admin use case, Since, distribution of items differs across /stores, it is expected that recommended items will change according to items being sold in a store. Thus, it is required for organisations to experiment with items in their stores.

7. CONCLUSION AND RECOMMENDATIONS

An ideal automated market basket analysis system is the one which there is a perfect analysis system working behind the scene to give users recommendations that will help their

shopping experience. In this paper, the introduction of various internet technologies and new algorithms has been able to provide solutions to the issue of analyzing a customer buying behavior. This software when deployed into use by retailers, would give giving retailers the edge they need in running business transactions with increased profit. Further works will implement certain features the project was not able to examine through the use of a more robust algorithm in the system, which would enable the system to work faster and more efficiently. Improvement on search methodologies can also be of help in generating recommendations for consumers.

8. REFERENCES

- [1] Yen-Liang Chena, Kwei Tangb, Ren-Jie Shena , & Ya-Han Hua (2004). Market basket analysis in a multiple store environment. *Decision Support Systems: Elsevier 40 (2005) 339 – 354* available online at www.sciencedirect.com.
- [2] Trupti , A. K., & Prof. Santosh, V. C. (2014). An Overview of Association Rule Mining Algorithms. *International Journal of Computer Science and Information Technologies vol 5(1)*. 927-930
- [3] Kavita Mittal, Dr.Gaurav Aggarwal, Dr.Pruna Mahajan (2017). A comparative study of association rule mining techniques and predictive mining approaches for association classification. *International Journal of Advanced Research in Computer Science Vol 8 (9)*, 365-372.
- [4] Savi, G., & Roopal, M. (2014). A Survey on Association Rule Mining in Market Basket Analysis. *International Journal of Information and Computation Technology*. ISSN 0974-2239 Volume 4, Number 4, pp. 409-414
- [5] M.Dhanabhakym, M., & Dr.M.Punithavalli, D. (2012). A Novel Market Basket Analysis Using Adaptive Association Rule Mining Algorithm. *International Journal of Scientific Research, 1(4)*, 25-28. doi:10.15373/22778179/sep2012/9
- [6] Kumar, D. A., Annie, M. C., & Begum, T. U. (2012). Computational time factor analysis of K-means algorithm on actual and transformed data clustering. *International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012)*. doi:10.1109/icprime.2012.6208286.
- [7] Manisha, G., Kanika, N., Saloni, i., & Neha, S. (2013). Comparative Survey on Association Rule Mining Algorithms. *International Journal of Computer Application vol 84 (10)* 1-3.
- [8] C.H. Lee, C. L. (2001). On mining general temporal association rules in a publication database. *Proceedings of the 2001 IEEE International Conference on Data Mining*, 337-344.
- [9] Association Rule Mining – Not Your Typical Data Science Algorithm | MapR. (n.d.). Retrieved from <https://mapr.com/blog/association-rule-mining-not-your-typical-data-science-algorithm/>
- [10] Crossley, N. (n.d.). Using Social Network Analysis: Researching Relational Structure. *Understanding Social Research: Thinking Creatively About Method*, 75-89. doi:10.4135/9781446287972.n5
- [11] Springer-Verlag, Troy Raeder, Nitesh V. Chawla Augus (2010) "Market basket analysis with networks"
- [12] Charles, S. Wasson, & John, Wiley & Sons, Inc. (2006). *System Analysis, Design, and Development*, by Charles S. Wasson 2006 by John Wiley & Sons, Inc.