

Shopping Recommender Technology Using Data Mining

Ramesh Babu B¹, Prathik Pai K², K Raghuv³

¹Student, Dept of Information Science Engineering, The National Institute Of Engineering, Mysore.

³Professor, Dept of Information Science Engineering, The National Institute Of Engineering, Mysore.

Abstract - The mode that people consume on the shopping gets more and more popular, it contributes the economic value in modern society. But it carries a series of problems to the consumers, such that the users can choose the expected product hardly. The paper mainly presents the fuzzy theory to deal with the users' behavior data, and the identification of customers taste for the product recommendations. A whole personalized recommendation model is built, and proposed evaluation and optimization function are applied to improve the accuracy of recommendation system with case study. Finally, the customer satisfaction function is better verified by the proposed system.

Key Words: Data Mining, Apriori algorithm, Association rules.

1. INTRODUCTION

With the rapid development of the technology, the shopping trend seems to be more and more necessary in peoples' daily life, and it is now changing peoples' consumption pattern, especially in some developed regions. The money which people spend through the shopping gets more and more, so it contributes the economic value in modern society, but it carries a series of problems to the consumers, such that the users can find and choose the expected product hardly. The product offers to the customers is universality. In this case, the personalized recommendation system was created, but the recommendation system has the universality, it causes the reducing accuracy of the recommendation. In recent years, there are many scholars who are starting to optimize the personalized recommendation system, and using this way to satisfy the different consumers' personal needs.

[4]The research of decision methodology and optimization based on customer purchase behavior is a hot point in shopping malls. Returns logistics method is used to determine whether the customer's returns aspiration are accepted or not, considering the customer's behavior. And the sensitivity of the time and satisfactions are considered as well.

Personalized recommendation of products for the customers is one of the factor to gain business profits as well as customers satisfaction. Today's products recommendation has universality but the customer buying behaviors differs based on their type and their interest. Advertising carries with it the ability to persuade, power to influence one's mind and mould the future. Advertising can be described as a way of communicating to the market in order to promote or sell

something. It can be a firm product, business service. In order to advertise, the messages are paid for by various agencies and sponsors which then appear in various forms of media like newspaper, television advertisement, direct mails, radio advertisement, magazines, emails, text messages, blogs and websites. Advertising is not only used to promote and sell but it has the power to persuade the mind and shape density of the market. The company's profit margins and economy can be taken to new heights with the power of advertising. Advertising meets the short term goals of a product like its information, generating awareness, improving credibility and productivity and also long terms goals which include maintaining a brand image of the product, adding emotional value to the brand and reaching a positive reputation in the market.

[4]The process of finding and analyzing useful patterns in a large amount of data is what is called Data Mining. Due to growing generation and increased work in development of Information Technology, vast amount of data is collected in various areas. This data is analyzed and maintained in a database in each sector. Hence, data mining can also be defined as an analytic operation design to analyze the data, mainly the so called "big data", and find consistent patterns and relationship that are systematic between the set of data variables, and then validating those findings by applying them on new subsets of databases. The patterns and the relationship so obtained is used to store and manipulate this data and use them in further decision making.

Data mining can also be defined as a logical process that is used to explore and search through large amount of big data so as to find more useful data in it. Previously, the pattern finding technique was not used and the data collected was just a collection of databases but with the technique of finding patterns in the data, more utilization of the data is being obtained which helps to make better decisions for the development of the business. Usually data mining is processed in three steps, mainly Data Exploration, Pattern identification and Deployment.

[5]During the first step of data exploration, the big data is processed and transformed into another form with only important variables left. This helps to find the nature of data and finally the problem is determined. After going through the first step of data exploration, the data gets refined and only specific variables are left which makes it go through the second step of pattern identification. This step involves the identification and choosing of patterns to apply them on new databases. After the second step, a specific pattern is obtained which can be now deployed on the new data sets, completing the process of data mining.

2. System Analysis

2.1 Existing System

The current system which is being used in the market is suggesting the offers to the customers universally, there is no technique to find the field of interest of a particular customer, whenever the offer is introduced it is sent to the customers universally whether the customer is interested in that field or no. this system may be useful in alerting the customers about the offers, but may fail to impress the customer because there will be chance that the customer is not interested with the offers that are suggested because the categories in which the offers are may not be among the area of interests of that customer. This method will be of some use if the offers are sent according to the customer's area of interest, so that it is some use for him rather than being a complete waste.

2.2 Proposed System

The system is proposed keeping in mind the growing rate in shopping and technology. The databases of the stores which contains the purchase details of the customer can be put to some use by applying new technologies to it. Data Mining is the process of finding and analyzing useful patterns in a large amount of data. Using Data Mining the pattern regarding the interest of a customer can be found out by analyzing the purchase database of the customer. Using the association rules under the Data Mining concept the pattern can be generated which is possible to give the area of interest of the customer. Whenever there is an offer in the store, if there is offer on the category in which the customer has area of interest then only those offers will be sent to the customer rather than sending the offers universally to all customers.

Under the association rules the algorithm called Apriori algorithm is used which helps to find the patterns which can be used to find the area of interest. The description of this algorithm is given below

Step 1: Scan the opinion data set and determine the support(s) of each item.

Step 2: Generate L1 (Frequent one item set).

Step 3: Use L_{k-1}, join L_{k-1} to generate the set of candidate k-item set.

Step 4: Scan the candidate k item set and generate the support of each candidate k – item set.

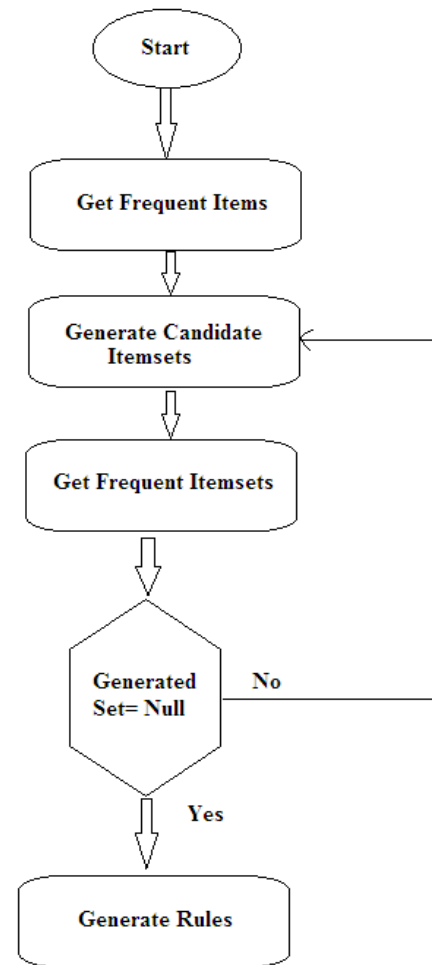
Step 5: Add to frequent item set, until C=Null Set.

Step 6: For each item in the frequent item set generate all non-empty subsets.

Step 7: For each non empty subset determine the confidence. If confidence is greater than or equal to this specified confidence .Then add to Strong Association Rule.

Step 8: Consider the items with more number of occurrences and those items are used for recommendations

The flow chart for this algorithm is shown in the Fig1



Flow of the Algorithm

Figure 1

By using this algorithm it is possible to get the pattern regarding the interests of the customers which is the main aim of the proposed model. The input and output of the algorithm is given below

Input and Output:

Input –Customers' trade data and product offers

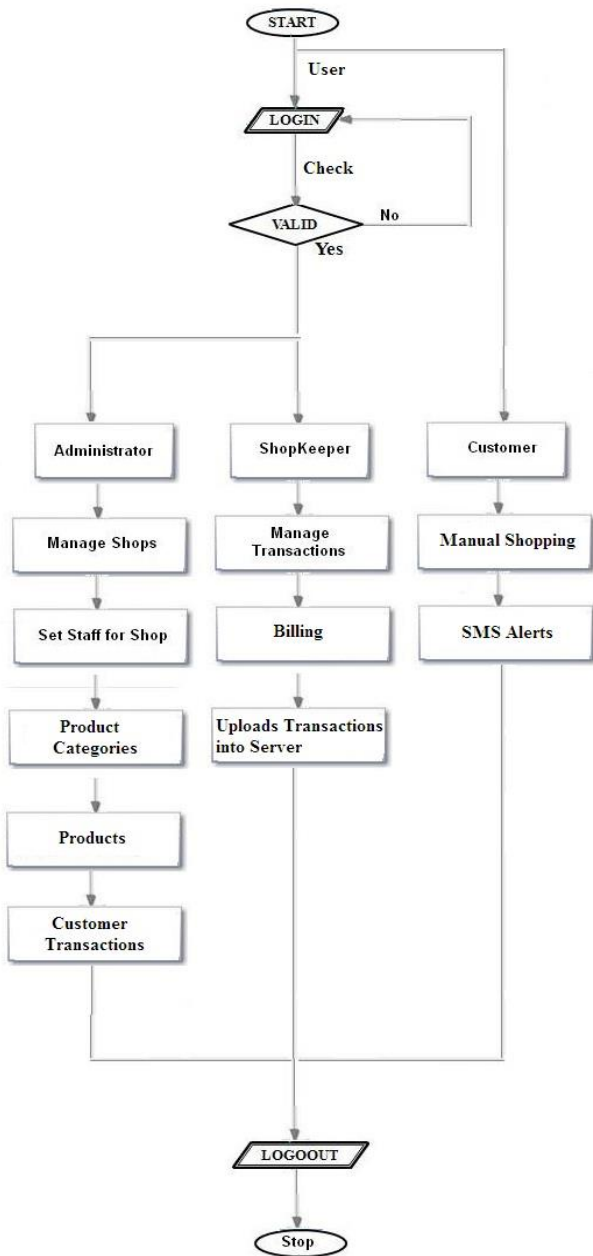
Output –Personalized recommendation of products and related offers.

There are three actors in this proposed model namely Admin, Shopkeeper and Customer

The Admin is the one who allocates the user id to the shops and also manages the databases, he is the one having all the privileges to the databases. The role of admin is shown below in Fig2

The shopkeeper is the one who looks after the store and the billing transactions in the store, he should collect the details

of the customer while billing and upload the billing transactions to the database and should keep the database updated. The role of shopkeeper is shown below in Fig2
 The customer is the one who goes to the store for shopping and he will receive the offers through SMS whenever there is offer for the categories which is under his area of interests. The role of customer is shown below in Fig2



**Flow Chart
Manual Shopping System**

Figure 2

The use case diagram for the admin is shown below in Fig 3 which shows what the admin looks after.

USE CASE DIAGRAM FOR ADMINISTRATOR

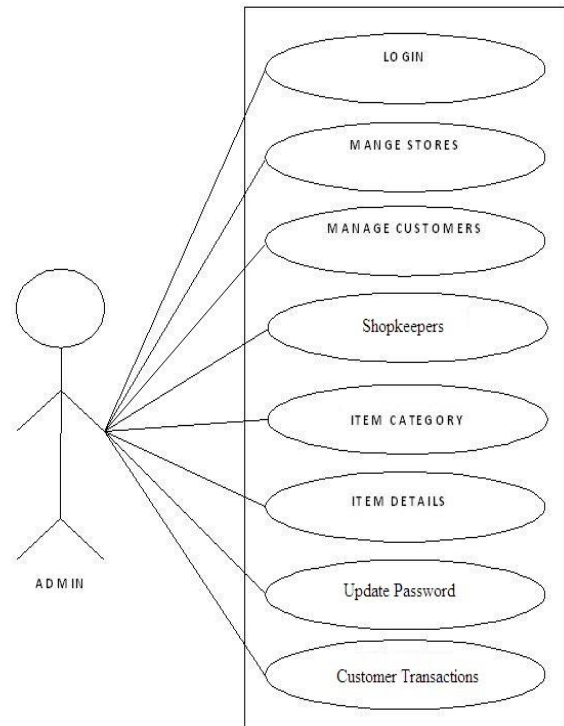
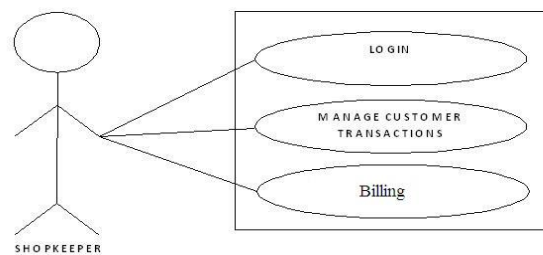


Figure 3

The use case diagram for the shopkeeper and the customer is shown below in Fig 4

USE CASE DIAGRAM FOR SHOPKEEPER



USE CASE DIAGRAM FOR CUSTOMER

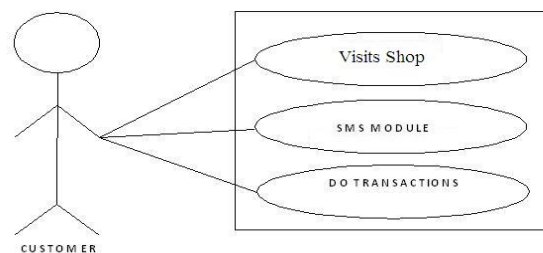


Figure 4

3. CONCLUSIONS

Predicting future actions of individual users is becoming possible given the multitude of data available in SNSs. In this we demonstrate the supervised approach for predicting purchasing activities of users. We also validate hypotheses behind the choice of features using time series of purchases. We then report satisfactory results for detecting the future purchase of two product types and we demonstrate that the inclusion of product-related data from within user's context can improve the accuracy to small degree. Future work embraces extensive investigation of the effectiveness of purchase prediction of diverse product types. This system provides competitive deals and offers to customers, thus giving them better options to choose from. It also provides the sellers better insights into customer needs. Using these results they can formulate their promotion strategy leading to increased sales. Also it will help them to manage their inventory efficiently based on the needs and requirements of the customer. All this combined will result in them being at a better position to compete with e-commerce sites, who are eating a major chunk of their market.

4. REFERENCES

- [1] [https://domino.fov.uni-mb.si/proceedings.nsf/0/1ae52d3e68074686c1256e9f00326308/\\$file/49_prassas.pdf](https://domino.fov.uni-mb.si/proceedings.nsf/0/1ae52d3e68074686c1256e9f00326308/$file/49_prassas.pdf)
- [2] Google images/manual shopping
- [3] https://en.wikipedia.org/wiki/Data_Mining
- [4] <http://ieeexplore.ieee.org/document/6574567/>
- [5] <http://ieeexplore.ieee.org/document/1524034/>
- [6] <http://smejoinup.com/blog/4-steps-for-successful-promotional-and-transactional-sms-marketing-of-your-business/>.
- [7] <http://ieeexplore.ieee.org/document/6574545/>
- [8] https://en.wikipedia.org/wiki/Apriori_algorithm
- [9] <https://en.wikipedia.org/wiki/SMS>