International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

E-commerce Shopping System with Semantic Search and Recommender System

Swastik Sonkusare[1], Mohit Singhvi[2], Kedar Atkar[3], Neha Kudu[4]

[1,2,3]*Student, Dept. of Information Technology, Vidyalankar Institute of Technology, Mumbai*
[4]*Professor, Dept. of Information Technology, Vidyalankar Institute of Technology, Mumbai*

Abstract - *In this paper, the proposed system is an E-commerce shopping system with advanced functionalities like semantic search and a recommendation system. The semantic search gives the optimum results to the search query without manually applying filters on it. The recommendation system recommends the similar products to the user. The recommendation system uses Pearson's correlation to find the relation between the two products. Other than these functionalities it also has functionalities like adding products to the cart, order a product, get a confirmation email, sorting of products based on various factors, several payment options, rating the products, etc. This paper encapsulates the idea of an E-commerce system with advanced functionalities.*

Key Words: Semantic search, Recommender, E commerce, Correlation.

1.INTRODUCTION

E commerce or Electronic commerce refers to different online business activities for products and services using the internet, and the transfer of money and data to execute these transactions.

Ecommerce has evolved to make products easier to discover and purchase through online retailers and marketplaces. Independent freelancers, small businesses, and large corporations have all benefited from ecommerce, which enables them to sell their goods and services at a scale that was not possible with traditional offline retail.

One of the most important features of any shopping system is search. With hundreds of products to choose from, a customer might not be able to find the exact product they wish for. This is what too much information does. What is the meaning of having tons of products if the right product is not discovered by the right customer at the right time?

Another considerable challenge is helping customers through the recommendation of a wide variety of product categories to efficiently find the one they will like the most. One of the techniques to handle the above challenge is a recommendation system.

2. Literature Survey

2.1 Improving Collaborative Filtering Recommender Systems Using Semantic Information

A Recommender System (RS) is a filtering system that tries to predict the rating or preferences a user would give to a digital item or to a social element (e.g. people) that he/she has not yet considered. Content-based (CB) and Collaborative Filtering (CF) are the two traditional and widely employed recommendation techniques. CB filtering techniques generate their prediction using a model that analyses the description of the item and the user's preferences. In CB filtering techniques, the recommended items are usually similar to items the user liked in the past. CF methods on the other hand generate their prediction using a model that analyses large amounts of information relating to user's historical rates. In CF methods, the items recommended to a user are usually similar to what similar users have liked without any understanding of the item itself. In other words, if user X and user Y gave rates to n items in a similar way then there is a high probability that they will give similar rates to other items. CF can be classified as: memory-based CF or model-based CF. Memory-based CF uses the historical users rating data in order to calculate similarity values that are used to generate the recommendation. The model-based CF uses the historical users rating data to create a model such as Bayesian classifier, fuzzy algorithm, clustering algorithm and genetic algorithm in order to generate the recommendation. [1].

2.2 Constructing Hierarchical Product Categories for E-commerce by Word Embedding and Clustering

When customers browse over the products, they have to choose one item out of thousands of similar items. This will cause a panic for customers since there are too many choices. A detail hierarchy is necessary to help the business for both the e-commerce websites and the customers; however, it is costly to build and update the hierarchy manually. We find that there are two problems in the current websites, firstly, the hierarchy is shallow; there are often too many products in the same category, and it is hard for a consumer to browse them. Secondly, the hierarchy is constructed manually when new products come, it is hard to update the hierarchy. Based on the

product description analysis, it is possible to solve the problems. We suggest using clustering algorithms to form natural hierarchy amount products in the same end category. Since the product description is in text, a word embedding method must be used to transform the text into vectors so that clustering algorithm can be applied. [3]

3. Methodology

3.1 The Website

Our goal was to develop a web application that would be attractive enough, have a professional look and user friendly. So that people of all age groups would be its end users.

The website allows the users to browse various products even without logging in. The user can sign up for the website which will allow the user to log in on the website, thus authenticate themselves.

Following are the features provided by the system:

• The user can search for any product available on the website. To make the search more effective the user is also allowed to describe the product directly instead of applying various filters.

• For every product customer see, they will be provided with recommendations of similar products which are made on the basis of ratings of the items.

• A very user-friendly user interface to enhance the shopping experience of the customers which will display all the relevant description of each product, images of the product.

• Once the user selects a product it can be added to the cart and then the customer can continue shopping further. After adding all the products to a cart the customer can view the cart, make changes if any, and then proceed to checkout.

• Users will be provided with several payment options to choose from, which will be beneficial for the customer.

• Once the order is confirmed, the user will be getting a confirmation mail on the e-mail ID provided while signing in.

• The user can rate the product by rating in terms of stars.

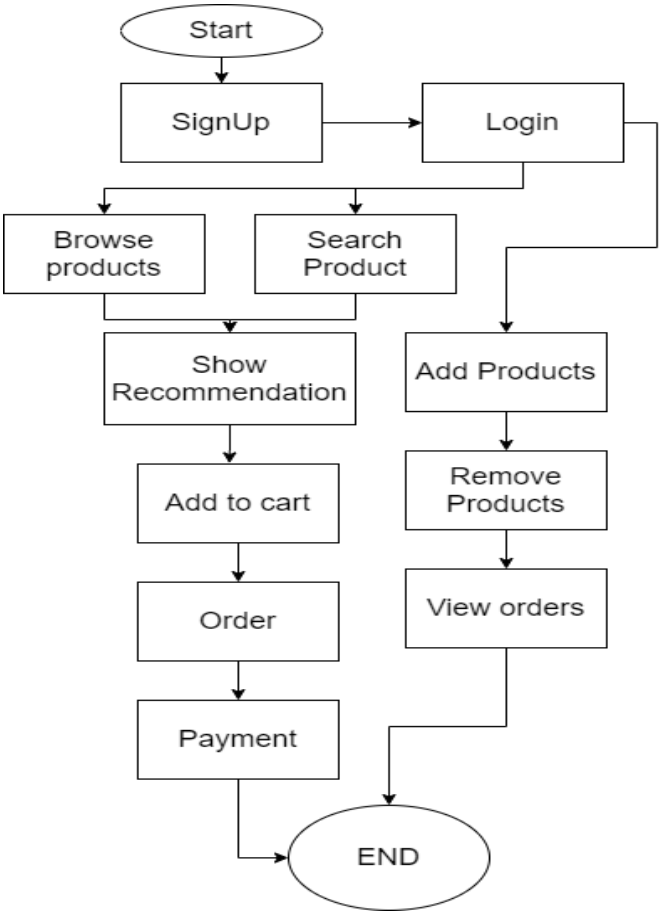

Fig -1: Proposed website work-flow

3.2 Search

When Customers are often required to apply various sorts and filters while searching for any product. While this can work well, it is time consuming. The more time it takes for a customer to find a right product the worse is the user experience. It would be better if the user can instead type in the exact description of the product they wish to see and get them all displayed.

Understanding what a customer actually wants is the key. Instead of pure keyword based searching, A searching system that actually can understand what a customer is trying to search by categorizing different keywords present in the sentence, understanding what each keyword implies and from the phrase entered by the customer present the most appropriate products by forming the SQL queries internally in a dynamic way, would be more effective and efficient.

This is what the proposed search system is based upon. The Sentence or phrase entered by the customer undergoes tokenization. After Tokenization, the stop words (not so important words like is, it, etc.) are removed. Then each of these words undergo lemmatization, which bring all the words to their basic

forms. What's left is a list of words which are then categorized based on what they mean, different dictionaries are used for categorization. Categories are general information about the product like color, gender, price range, etc. of the product. Once it is identified which word represents what, an SQL query can be formed which will be a combination of all the identified categories and the result can be obtained.

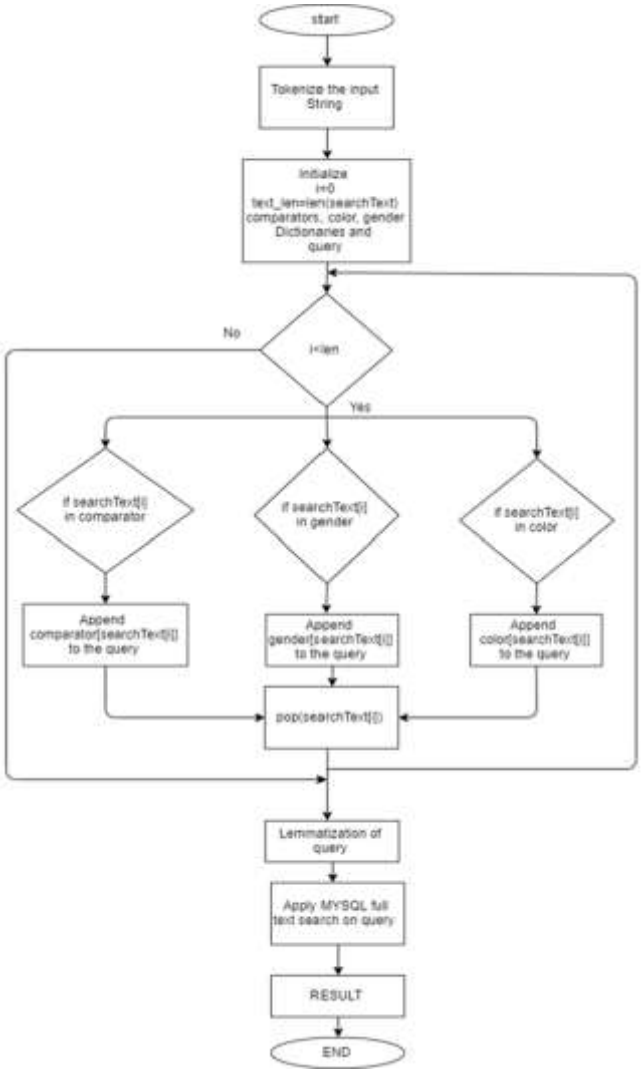

Fig -2: Proposed Semantic search work-flow

3.3 Recommender System

We used Amazon rating data for the recommender system using content-based filtering. It is a recommender system based on the similarity between items calculated using people's rating of those items. Collaborative filtering performs poorly for sparse data and is computationally expensive. Content based filtering models resolve these problems in systems that have more users than items. It uses rating distributions per item, not per user. With more users than items, each item tends to have more ratings than each user, so an item's average rating usually doesn't change quickly. This leads to more stable rating distributions in the model, so the model doesn't have to be rebuilt as often. When users consume and then rate an item, that item's similar items are picked from the existing system model and added to the user's recommendations.

Pearson's correlation is used to find correlation between all the products. The products which have the highest correlation values to the products recently purchased by the customer are good recommendations. Also, for every product a customer views, products which are highly correlated to them are displayed. Pearson's correlation formula is as follows:

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\,\sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

n is the sample size.

xi,yi are the individual sample points indexed with i

rxy is correlation value.

3.4 Dataset

Dataset used for products is Amazon product dataset. This dataset contains meta data for products as

Products (prod_id, prod_name, prod_brand, prod_image, prod_description, prod_category, product rating)

Other tables of the database are:

Users (user_id, username, email, password, address)

Orders (order_id, user_id, order_date, paymentdone)

Order_contents (order_id, product_id, quantity)

Usercart (user_id, prod_id, quantity)

Rating (user_id, prod_id, rating)

4. Proposed Model

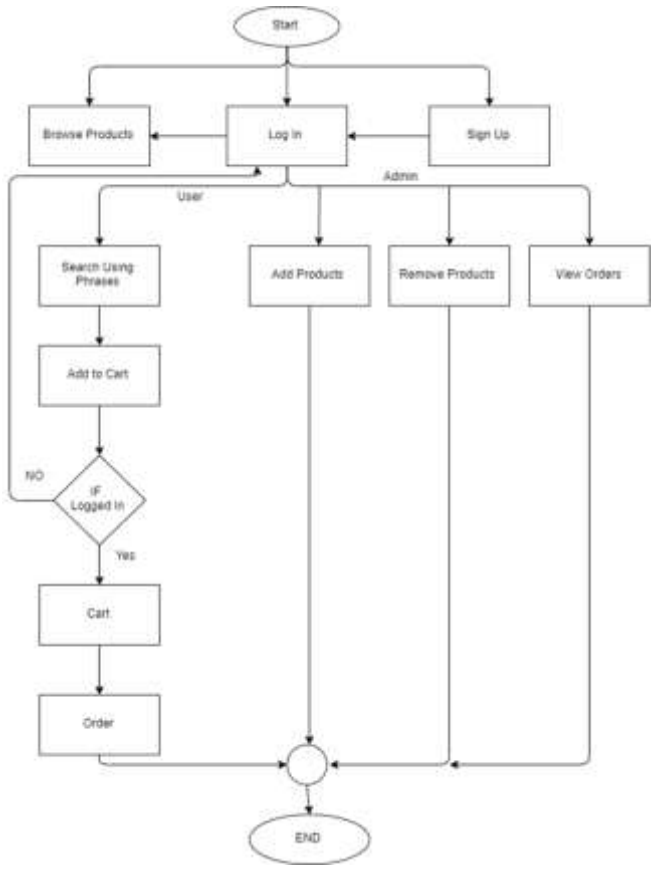

Fig 3: Block diagram of proposed system

The linear working of our project can be explained in the following manner:

1. User visits the website.

2. Sign up to the website or if the user already has account then he/ she will log with his/ her account.

3. The user can search the products using phrases and natural language.

4. The user can add the product to the cart.

5. System will recommend the products to the user based on user behaviour analysis.

6. The user will add the recommended product if wants to buy otherwise go to the next step

7. The user now can make changes in his/ her cart by adding or removing the items or can move further and order the products

8. The user is supposed to enter the delivery address and place the order.

9. After placing the order, the user will get confirmation Email from website.

5. Results

Fig 4 : Recommendations based on customer purchase history

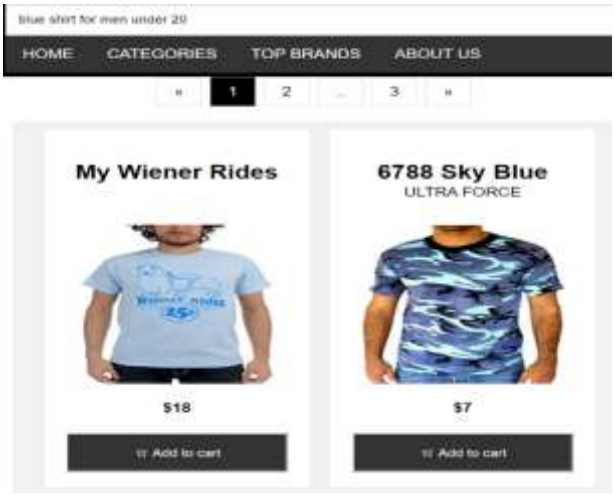

Fig 5: Search result for the user entered phrase "blue shirt for men under 20"

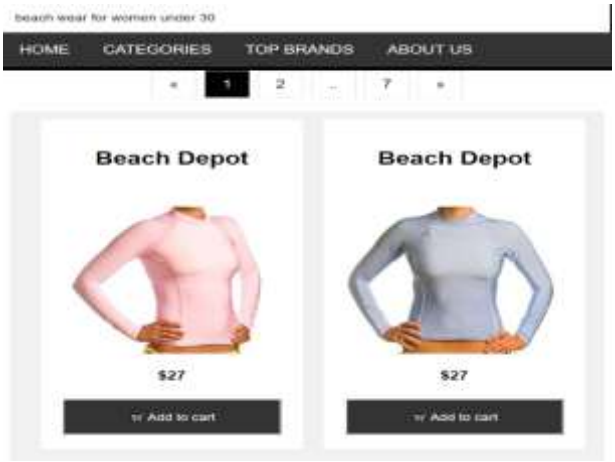

Fig 6: Search result for the user entered phrase "beach wear for women under 30 "

6. CONCLUSION

E-Commerce has simplified our lifestyle so much by reducing the time and money we spend on travelling to the shops. Searching system that understands what a customer is trying to search by categorizing keywords present in the sentence, understanding what they imply and present the most appropriate products by forming the SQL queries is very efficient.

REFERENCES

[1] Bushra Alhijawi "Improving Collaborative Filtering Recommender Systems Using Semantic Information" 2018 IEEE 9th International Conference on Information and Communication Systems

[2] Tessy Badriyah "A Hybrid Recommendation System for E-Commerce based on Product Description and User Profile" 2017 IEEE the Seventh International Conference on Innovative Computing Technology

[3] Yi-Hsiang Hsieh, Shih-Hung Wu, "Constructing Hierarchical Product Categories for E-commerce by Word Embedding and Clustering" 2017 EE International Conference on Information Reuse and Integration (IRI)

[4] Hui-ke Rao, Ai-ping Liu, Zhi Zeng, "Research on Personalized Referral Service and Big Data Mining for E-commerce with Machine Learning" IEEE 2018 4th International Conference on Computer and Technology Applications