Product Recommendation System based on User Trustworthiness & Sentiment Analysis

Gunjeet Kaur Soor [1*], *Amey* Morje [1*], *Rohit* Dalal [1*], and *Dr. Deepali* Vora[2]

[1]B.E., Information Technology, Vidyalankar Institute of Technology, India
[2]Professor, Information Technology, Vidyalankar Institute of Technology, India

Abstract: The current online product recommendation system based on reviews has many limitations due to randomness in the review patterns. The data which is used are the reviews and ratings from the e-commerce websites. This data might contain fake reviews that make the data uncertain. Due to this, the currently existing systems produce ambiguous results on this present data. Instead of this, the new system uses only genuine reviews, considering the trustworthiness of the user and generates the results in a more significant manner. The proposed system scrapes reviews from different online websites and performs opinion mining and sentiment analysis on it. Other factors like star ratings, the buyer's profile and previous purchases and whether the review has been given after purchasing or not are included. Based on these factors & user trustworthiness, the website from which the user should buy the product will be recommended.

1 Introduction

A product recommendation system predicts and shows the items that a user would like to purchase. It is basically a filtering system. The results might not be entirely accurate. It shows results according to your likes, and recent searches. The idea was generated on the paper which was published in 2017 in IEEE Transactions on Information Forensics and Security which was based on spam detection and fake social media reviews [1].

Recommender Systems have become more useful in recent years. They are utilized in various fields which include movies, music, news, books, research articles, search queries, social tags, and products in general. Mostly used in the digital domain, the majority of today's E-Commerce sites like eBay, Amazon, Flipkart, Alibaba, etc. make use of their proprietary recommendation algorithms in order to better serve the customers with the products they are bound to like [2].

Along with this, the product reviews are also used in the recommendation. Review helpfulness is gaining the attention of practitioners and academics. It helps in reducing risks and uncertainty faced by users in online shopping.

With the growing amount of information on the internet and with a significant rise in the number of users, it has become important for companies to search, map and according to the user's preferences and taste, it provides them with the relevant chunk of information. Chatbots also work on the same working, but they are a little smarter and learn from each product the user views or buys.

2 Literature Survey

There are a number of journals and conference papers that share information about the recommendation systems. The following table lists the key findings and algorithms used after reviewing some of the important papers. [3-11]

Table 1. Summary of Literature Survey

Sr. No.	Authors	Key Findings	Algorithm
1	K L Santhosh Kumar, Jayanti Desai, Jharna	Sentiment Analysis of reviews fetched from Amazon. Naïve Bayes algorithm shows comparatively better results than Logistic	Naïve Bayes classification, Logistic

* Corresponding author: gunjeetks@gmail.com, amey.morje@gmail.com, dalalrohit102@gmail.com

	Majumdar	Regression and SentiWordNet technique. The performance of these algorithms is measured by using quality metrics like Recall, Precision, and F-measure. [3]	Regression, SentiWordNet algorithm
2	Kai Yang, Yi Cai, Dongping Huang, et. al	A sentiment dictionary using external textual data is created and different classification models are compared along with a hybrid model. A hybrid model is made by combining Support Vector Machine (SVM) and Gradient Boosting Decision Tree (GBDT) based on Stacking approach, which is an ensemble learning approach usually used to combine different learning algorithms to get better performance. The baseline model made for verification compares between SVM, GBDT and the hybrid model, in which hybrid model is found to be the most effective. [4]	Support Vector Machine (SVM) and Gradient Boosting Decision Tree (GBDT)
3	S.A.Sadhana, L.SaiRamesh, et. al.	The proposed method based on a monolingual word alignment model (WAM) and for additionally best alignment quality, this approach uses Semi-supervised Word Alignment Model (SWAM) with supervision on alignment. The results show that the true positive rate gradually increases in prediction by the proposed SWAM technique. Experimental results also show that using the sentimental analysis the system provides the result with higher accuracy than the system without sentimental analysis. [5]	Semi- supervised Word Alignment Model
4	Prashast Kumar, Arjit Sachdeva	The process of gathering online end-user reviews for the products or services is automated. These reviews are then analyzed based on the sentiments expressed about specific features. Online product reviews from Flipkart are the inputs to this project, which the system analyzes to generate results (for reviews) which have been sorted according to various geographic regions. These results are also shared with the manufacturers to look at the sentiments and judge the improvisation and deterioration of the Product. [6]	Sentiment Analysis
5	Shaozhong Zhang and Haidong Zhong	For addressing the problem of mining users trust in an E-commerce system, the trust is divided into two categories, direct trust and propagation of trust, which represents a trust relationship between two individuals. [7]	NLProcessor linguistic parser
6	Paitoon Porntrakoon, Chayapol Moemeng	The trustworthiness of an e-vendor is measured in multiple dimensions like the product, price, and shipping. This helps the customers in making a good purchase decision from the user with a high trust score. This is analysed by using multidimensional lexicon and sentiment compensation technique. The accuracy of the proposed method (SenseComp) is compared with manual, sentiment to dimension (S2D) and dimension to sentiment (D2S) methods. The results of the experiment show that the sentiment compensation technique increases the accuracy of SenseComp in all dimensions with an overall accuracy of 93.60% as compared to S2D and D2S methods. [8]	Sentiment Compensation Technique
7	A. Angelpreethi, S. Britto Ramesh Kumar	The aim is to build a feature opinion miner which finds the important feature of the products by analyzing the review or opinion and create an opinion profile for each product which can be used by the customer. For this, the model uses a supervised syntactic approach for opinion mining. The proposed system here uses SentiWordNet for calculating the polarity of the opinion and dependency parsing is used for extracting feature matched opinion word. [9]	SentiWordNet, dependency parsing
8	Evangelos Triantaphyllou and Stuart H. Mann	The Analytic Hierarchy Process (AHP) is an effective approach in dealing with decision problems in the industry. The AHP is a decision support tool. It is used to solve complex decision problems. It works with a multi-level hierarchical structure of objectives,	Analytic Hierarchy Process (AHP)

		criteria, subcriteria, and alternatives. By using a set of pairwise comparisons, the pertinent data are derived. The reuslt of the comparisons is used to obtain the importance of the decision criteria weights, and the relative performance measures of alternative which is in terms of individual decision criterion. If the comparisons are not perfectly consistent, then it provides a mechanism for improving consistency. [10]	
9	Bilal Abu-Salih, Pornpit Wongthongtham, Kit Yan Chan, Dengya Zhu	A fine-grained users' credibility analysis framework for big social data, named CredSaT (Credibility incorporating Semantic analysis and Temporal factor) whose credibility core module is based on three main dimensions (i) distinguishing users of the various domains of knowledge; (ii) a novel metric incorporating a list of fine-grained key attributes is harnessed to create the feature-based ranking model; and (iii) the temporal factor is used to study the users' behaviour over time and reflect this behaviour by means of their domain-based credibility values. Experiments conducted to evaluate this approach validates the applicability and effectiveness of determining highly domain-based trustworthy users, as well as capturing spammers and other low trustworthy users. [11]	AlchemyAPI

2.1 Inferences from Literature Survey

From the above literature survey, it is inferred that sentiment analysis can be performed by using many algorithms but, Naïve Bayes Classifier and Logistic Regression gives optimum results. Out of the two, Logistic Regression proves to be better and works efficiently even on small datasets. For user trustworthiness, AHP is an effective approach in dealing with decision problems using a number of parameters to generate the ranking which is not possible in our case because of limited parameters. Hence, a custom scale would be used for user trustworthiness.

3 Proposed System

The following block diagram shows the flow of how the system will work. It starts with the user searching for a particular product and ends with the recommendation of the website to the user.

Fig. 1. Block Diagram

The above block diagram can be subdivided into five major blocks as follows:

i. **Data Scraping**
When the user enters the product, the data is collected from various E-commerce websites which are shown, viz. Amazon and Flipkart. This includes the images of the product, features of the product, product ID, star ratings and reviews from each website. All this data is obtained with scrapers which are different for different websites.

ii. **User Trustworthiness**
This block will determine whether the review given by the specific user is trusted or not. For this, various factors of the user like previous reviews, previous purchases, visits, etc. will be taken into consideration. This block is still in progress as it is facing difficulties to obtain the personal information. Rather we are trying to find a new solution.

iii. **Sentiment Analysis**
This block will perform sentiment analysis on the trusted reviews only. For this operation, we will use various python libraries out of which NLTK is an important one for Natural Language Processing. This analysis will categorize the reviews as the positive and negative using the Logistic Regression Model to produce results.
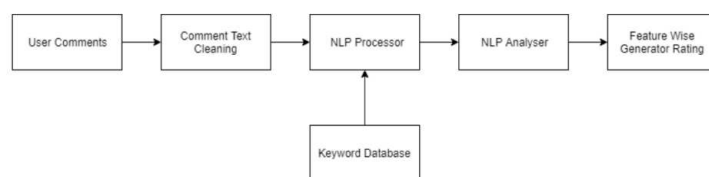

Fig. 2. Sentiment Analysis

iv. **Feature Comparison**
This block compares various other features like cost, etc. which are also collected during the scraping process. These individual run through the sentiment analysis and produce much better results.

v. **Recommendation of Website**
This block finally recommends the user the website from which the user should buy the product on the previous analysis done.
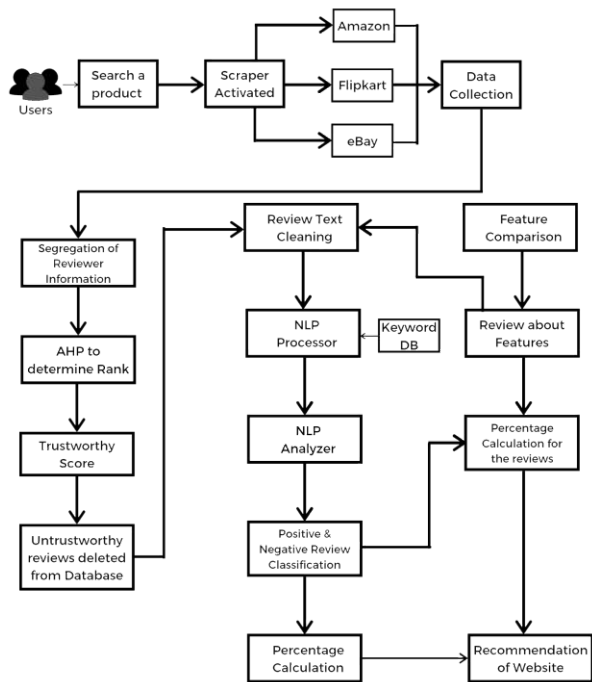
4 Design

Fig. 3. Flowchart

When a user searches for a product on the web page, the reviews from the e-commerce websites are scraped. The data collected includes the reviewer's details and the review. A rank is provided to the reviewer, which is used to calculate the trustworthy score. Untrustworthy reviews/ fake reviews are removed from the database. Further, the sentiment analysis is carried on the cleaned data using NLP. Features provided by the search item are also considered to calculate the percentage for trustworthiness. All these factors combine and the website from which the user should buy is recommended.

5 Implementation

The proposed design has been implemented block-wise as scrapers for different websites, the trustworthiness model and the sentiment analysis model and in addition, an user interface is developed.

Whenever the user enters the product in the search bar, the initial scraper which is in JavaScript is activated, scrapes the relevant links of Flipkart and Amazon which are stored onto a Firebase database for further use. Now, the exclusive scrapers of Amazon and Flipkart, which are designed in Scrapy, scrape the product reviews which also include the review heading, review text, rating, author, review helpfulness count and whether the purchase is verified or not. Scrapy is a framework in Python, which is generally used for scraping. It helps in scraping large amount of data in lesser time with various features like seamless extraction with IP jumping.

The trustworthiness of the review depends on the user and certain properties of its own which could be used to decide whether they should be included for the sentiment analysis part or not. For a review to be trustworthy, some parameters like how many reviews the

user has posted, how useful the reviews have been and determining the genuineness of the review i.e. whether the review is posted after purchasing or not are considered. For Amazon, only the reviews of verified purchase by the reviewer are considered and scraped. The reviews which have more than 10 helpful votes are considered. The review text length must be greater than or equal to 10. For Flipkart, the reviews with certified buyer badge are considered. The importance of the review could be assigned on the basis of likes (upvotes) and dislikes (downvotes). The review text length must be greater than or equal to 10.

The sentiment analysis on the selected reviews is performed by using python and its machine learning libraries and Logistic regression algorithm is applied on the reviews and a percentage of positive and negative reviews are obtained.

The User Interface is designed using the React.js framework and the backend is implemented using Express.js. The connection of this UI with the other elements is done with the help of Flask which is a Python framework. Flask integrates all the things into the UI and the final result is displayed which shows the product with its availability and information and the review sentiment analysis comparison on both the websites.
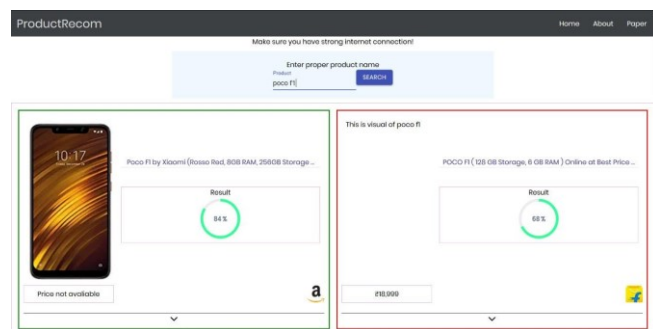
6 Results

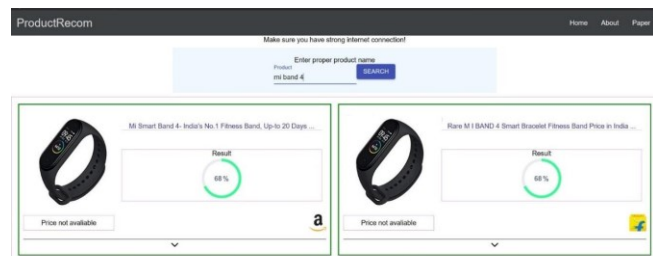**Fig. 4.** Recommendation with different percentage

Fig. 5. Recommendation with same percentage

Fig. 6. Fetched Reviews (CSV)

Fig. 7. Accuracy

Fig. 8. Confusion Matrix

This is quite obvious that the product whose reviews are more positive will have a higher percentage and hence will be recommended to the user. The complete procedure takes about a minute or so to display the result.

In figure 4, the product Poco F1 is searched, to which the photo and other specifications from both the sites are made available. Then, the reviews from both the sites are scrapped and sentiment analysis is performed. It is seen that the Amazon reviews are 84% positive while the Flipkart reviews are 68% positive. Thus, the site recommends the Amazon website to the user which gets highlighted with a green border. It is seen that the price is not available for the Amazon site while the photo is not available for the Flipkart one. This is because sometimes the product may not be available at the moment or there might not be a fixed price of the product. While in case of Flipkart, it is observed that the website blocks the attempts to scrape after a few continuous scrape attempts are made.

It is observed that the percentage of the sentiment analysis result might be the same. At such times, if only a single website is recommended, then it could be considered that the application favours a certain website which should not be the case. So, an example of this is shown in the figure 5. The product MI Band 4 has same result percentage of 68%. Hence, it depends on the user's personal choice to buy from a specific website. For the application's recommendations, both the websites are highlighted as green and recommended.

Whenever there is huge difference between Training and Testing set, it is observed that there might be some inaccuracy in the model. The results show that the Training set accuracy is 90.4% while that of Testing set is 86.3% which is a reasonable difference.

7 Conclusion

In our work, we have covered the mining of reviews from e-commerce websites on which sentiment analysis is performed along with checking of the reviewer's trustworthiness. Many alterations and experiments were performed for Sentiment Analysis and the over-fitting difference was reduced to approximately 4%. Another option would be to adopt an approach based on neural networks to further reduce over-fitting. A dedicated user interface is developed for the execution of the tasks and to display the recommended e-commerce website to the user.

8 References

1. S. Shehnepoor, M. Salehi, R. Farahbakhsh, N. Crespi, IEEE Transactions on Information Forensics and Security, **12** (2017)

2. "Towards Data Science Blog related on Recommendation Engines", Medium Blog (2017)

3. K.L.S. Kumar, J. Desai, J. Majumdar, IEEE ICCIC, (2016)

4. K. Yang, Y. Cai, D. Huang, J. Li, Z. Zhou, X. Lei, IEEE International Conference on Big Data and Smart Computing (BigComp), (2017)

5. S.A. Sadhana, L. SaiRamesh, A. Kannan, S. Sabena, S. Ganapathy, 2nd International Conference on Recent Trends and Challenges in Computational Models (2017)

6. P. Kumar, A. Sachdeva, 5th International Conference Confluence The Next Generation Information Technology Summit (2014)

7. S. Zhang, H. Zhong, IEEE Access Volume **7** (2019)

8. P. Porntrakoon, C. Moemeng, ECTI-CON, **15** (2018)

9. A. Angelpreethi, S.B.R. Kumar, WCCCT (2017)

10. E. Triantaphyllou, S.H. Mann, International Journal of Industrial Engineering: Applications and Practice, **2**, No. 1, pp. 35-44 (1995)

11. B. Abu-Salih, P. Wongthongtham, K.Y. Chan, D. Zhu, Journal of Information Sciences (2018)