

Smart phone Recommendation using User Reviews and Product Specifications

Mayur Peche, IT dept., SIT, Lonavla
Nitin Kashyap, IT dept., SIT, Lonavla
Chetan Pujari, IT dept., SIT, Lonavla
Prashik Thool, IT dept., SIT, Lonavla

Guide
Prof. Pravin Latane
SIT, Lonavla

Abstract—E-commerce platforms have been getting increasingly popular with a large number of people utilizing this platform increasingly. This is due to the increased convenience and ease of use offered by these platforms in shopping for essential items and other products at the comfort of your home. This is allowed a lot of disabled individuals and individuals with mobility issues to effectively stay at home and order various items that are needed easily. This has also led to a large number of e-commerce platforms that are vying for the attention of the users and trying to increase the user base to be able to provide effective solutions with ease. One such implementation is the realization of product recommendations to the user based on the various product reviews that are posted on the products. This is a highly complex task as a large number of reviews are posted for a product with varying degrees of positivity or negativity. Therefore to achieve effective product recommendation a number of related researches has been outlined in this research article which utilize the machine learning approaches.

Keywords: *Natural Language Processing, Fuzzy Artificial Neural Network, TF-IDF, Collaborative Filtering.*

I INTRODUCTION

The E-Commerce platform has seen enormous growth in online platforms in recent years. This can be attributed to the improvements in the internet platform which has led to a significant increase in the user base which has made most of the implementation highly cost-effective for a variety of Web services. This is crucial for the development of the internet ecosystem that has been largely devoted to increasing the usability and convenience of the users on this platform. The significant increase in the affordability of internet-enabled

devices has further catalyzed the growth which has led to an enormous explosion in the E-Commerce category.

They have been countless research on the approach to improve this platform even further and provide the users with innovative and effective features that separate the E-Commerce platform from their competitors. This has led to an increased amount of research in improving the experience of the user and adding novel features that can attract a larger number of users. This allows the platform to grow much more and provide a lot more services which in turn will be highly convenient for the user shopping from the comfort of their homes.

One such useful implementation is the realization of product recommendations based on the user's taste and previous shopping history. This is a highly useful and effective approach that can provide a personal touch to the customer service provided on the E-Commerce platform to every user. Product recommendations are extremely complex and require an in-depth understanding of the shopping patterns of the user along with the various products that have been previously bought which will be specific for a particular user. This leads to a large number of combinations that can be overwhelming and extremely difficult to calculate recommendations.

The paradigm of machine learning and natural language processing comes to the rescue in achieving this goal of product recommendation. Through the implementation of these approaches, the products can be effectively reviewed and realized for their potential for recommendation to a particular user. The natural language processing approach is particularly important as it can analyze the product reviews for the purpose of identifying useful products with positive reviews that can provide quality recommendations to the user. The machine learning approaches can help determine sarcasm that can be

highly effective in detecting negative and positive reviews with a high degree of certainty.

Therefore in this approach, natural language processing has been realized through the implementation of the bag of words approach. The bag of words is nothing but a dictionary that this customized according to the implementation. In our approach, the two bags of words are used to segregate the positive and negative words present in the reviews. This allows us to determine if the review is entirely positive or negative to ascertain the quality of the product being examined. This approach is further improved through the realization of term frequency and inverse document frequency that can determine the importance of a particular word that can be useful in the effective elimination of sarcasm from the product reviews.

The implementation of a Fuzzy artificial neural network allows the segregation of the product reviews into five different levels of positivity or negativity. This allows a range of labels that can be assigned to a product which can be helpful in the recommendation procedure depending on the user's preference. This is further enhanced through the realization of collaborative filtering which effectively reminds the cosine similarities between two factors, the first one being previously recommended items to similar users, and the other being the recommendation being provided by our system which is effectively combined to improve the accuracy of the product recommendation significantly. The methodology has been effectively elaborated in this research article along with appropriate evaluation of the performance metrics which has been fruitful and effectively useful along with the future directions of research.

This literature survey paper segregates the section 2 for the evaluation of the past work in the configuration of a literature survey, and finally, section 3 provides the conclusion and the future work.

II RELATED WORKS

M. Ibrahim [1] in recent time's online harassment and cyber-bullying are some serious issues in online communities. The proposed paper uses the Wikipedia dataset for different data augmentation techniques. As proposed in the paper they have used three algorithms as a solution there is long short-term memory, convolutional neural network, and bidirectional gated recurrent units. The very first type is to analyze the statement to verify the toxic content from the statement. Thus the proposed paper achieves the highest accuracy among all considered algorithms.

K. Sintoris [2] the proposed paper implements the NLP on Business Process Management (BPM). There are certain stages to be covered in the proposed paper Semantic Patterns to analyze the different patterns to map the text of modeling fragments. The next step is to transformation rules

derivation to detect the various issues from the transformation heuristics. And the very last stage is improvement, customization, and modification thus upgrading the tool. The result of approach by implementing BPMN is successfully upgraded.

W. Jitsakul [3] states that the internet is known as an important factor for communication. There are different types of classifiers are Probability-based, Learning-based, Rule-based, and Tree structure-based. There are many applications for sentiment analysis such as Random Forest, Bayesian Logistic Regression, and Support Vector Machine. The proposed paper proposes the text classification method by implementing Random Forest Classifier with a combination of different features. The results of the proposed paper divide the customer's review into negative or positive feedback

M. Andriansyah [4] classifies the comments from Instagram the people who are cyberbullying. The comments data is collected from Instagram celebrities/Telegram total of 1053 comments were taken as input for review classification. Review classification is used for Support Vector Machine (SVM). Support Vector Machine (SVM) is used for classifying review two parts such as namely -1 and 1. They have used Integrated Development Environment (IDE) R language using RStudio's. The proposed paper results accuracy percentage of 79, 412% by using the Indonesian program using Support Vector Machine (SVM) method.

J. Savigny [5] narrates in Indonesia YouTube is known as one of the most popular video-sharing tools many people watch videos and pass comments. The comments on the video carry the emotions such as happy, sad, angry, surprised, disgust and fear. In past there has been lots of research made in NLP for classification they have used word embedding. In the proposed paper they have Convolutional Neural Network (CNN) algorithm on the different attributes such as average word vector with TF-IDF, namely average word vector and paragraph vector. Thus one of the best reviews results by using a word embedding with CNN method.

Siswanto [6] explains that there has been a significant increase in the number of users on various online platforms whether a large number of users try to get the relevant information that they need. The exponential increase in information leads to a lot of irrelevant information that can be highly problematic to deal with. Therefore the authors in this approach proposed the utilization of effective naive Bayes algorithm and support vector machines for the analysis and classification of social media comments. The effective algorithm has led to a significant improvement in the results wherein naive Bayes scores are low accuracy than the support vector machine implementation.

N. Chandra elaborates on the topic of classification and recommendation algorithms that are highly useful in the

current world nowadays. This is since the information on online platforms has been increasing exponentially in recent years. This large amount of information is highly troubling and could lead to inadequate information being extracted for the users [7]. This also completely changes the approach leads to improper information being extracted. Therefore the authors have proposed an effective comment classification approach that utilizes the k-nearest neighbor algorithm to achieve effective and useful comment classification of antisocial comments on popular social media networks.

A. Ikeda [8] expresses that the online platform has been an indispensable source of information for a large number of individuals across the globe. These individuals utilize these platforms for or various social media and other information such as images and videos. These individuals provide their opinions on these platform's videos by the utilization of the commenting approach. The comments on these videos need to be effectively classified to achieve personalized advertisements and other approaches to the user. This research article outlines an effective technique for the referred contents on the classification of comments based on annotations.

F. Prabowo [9] introduces the concept of social media platforms and online social networks that have been thriving in recent years. This is because most of these approaches provide effective and useful realization of communication between individuals and the dissemination of information on a large scale. Most of these social media platforms have also hosted their e-commerce platform for effective sale and purchase of various items on the social networks. There is a lack of an effective technique that can utilize the various comments on online shops on these social media networks to provide effective insight into the quality. Therefore the earth was a proposed an effective approach that utilizes a support vector machine along with a convolutional neural network to achieve effective learning and classification of the comments in a statistical manner.

M. Takeda [10] discuss the growth of various web applications and Web services that have been increasing in volume every day. These approaches are significant as they allow for effective communication between individuals across a large amount of distance. These Web services can be accessed through any internet-capable device which allows more and more users to get connected to these platforms. There have been effective approaches that are allowed users to share the opinions in the form of comments on various posts and other reviews online. The researchers have proposed an effective use of hierarchy in tree structures and kernels for accurate classification of comments that can be useful for effective information retrieval.

III CONCLUSION AND FUTURESCOPE

The methodology for an effective recommender system for product recommendation in e-commerce platforms through the utilization of product reviews has been effectively outlined in this research article. In order to achieve this aim of product suggestion, the paradigms of deep learning and language processing come to the defense. The goods may be properly examined and realized for their possibilities for recommendations to a specific user by implementing these procedures. The natural language processing technique is exceptionally essential since it can examine product reviews in order to locate helpful goods with favorable ratings that may give the customer with excellent suggestions. For this purpose a number of related works have been studied in this survey article. These relevant studies have provided insight and assisted in the development of an effective product recommendation approach, which will be detailed in future editions. The devised approach implements the Bag of Words approach, Term Frequency, and Inverse Document Frequency along with Fuzzy Artificial Neural Network and Collaborative Filtering for product recommendations which will be further detailed in the future.

REFERENCES

- [1] Mai Ibrahim, Marwan Torki and Nagwa El-Makky, "Imbalanced Toxic Comments Classification using Data Augmentation and Deep Learning" 17th IEEE International Conference on Machine Learning and Applications 2018
- [2] Konstantinos Sintoris Kostas Vergidis, "Extracting Business Process Models using Natural Language Processing (NLP) Techniques" 2017 IEEE 19th Conference on Business Informatics
- [3] Watchreewan Jitsakul, Phayung Meesad, Sunantha Sodsee, "Enhancing Comment Feedback Classification using Text Classifiers with Word Centrality Measures" 2nd International Conference on Information Technology (INCIT) 2017.
- [4] Miftah Andriansyah, Afina Ahwan, Nico Ariesto Gilani, Remi Senjaya, Ali Akbar Ardiono Roma Nugraha, Rizki Nofita Sari, "Cyberbullying Comment Classification on Indonesian Selebgram Using Support Vector Machine Method", Conference: 2017 Second International Conference on Informatics and Computing (ICIC.)
- [5] Julio Savigny and Ayu Purwarianti, "Emotion Classification on Youtube Comments using Word Embedding" 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)
- [6] Siswanto, Y. P. Wibawa, W. Gata, G. Gata, and N. Kusumawardhani, "Classification Analysis of MotoGP Comments on Media Social Twitter Using Algorithm Support

Vector Machine and Naive Bayes," 2018 International Conference on Applied Information Technology and Innovation (ICAITI), 2018, pp. 96-101, doi: 10.1109/ICAITI.2018.8686751.

[7] N. Chandra, S. K. Khatri and S. Som, "Anti-social comment classification based on kNN algorithm," 2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2017, pp. 348-354, doi: 10.1109/ICRITO.2017.8342450.

[8] A. Ikeda, A. Kobayashi, H. Sakaji and S. Masuyama, "Classification of Comments on Nico Douga for Annotation Based on Referred Contents," 2015 18th International Conference on Network-Based Information Systems, 2015, pp. 673-678, doi: 10.1109/NBiS.2015.101.

[9] F. Prabowo and A. Purwarianti, "Instagram online shop's comment classification using statistical approach," 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2017, pp. 282-287, doi: 10.1109/ICITISEE.2017.8285512.

[10] M. Takeda, N. Kobayashi, F. Kitagawa and H. Shiina, "Classification of Comments by Tree Kernels Using the Hierarchy of Wikipedia for Tree Structures," 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 2016, pp. 123-127, doi: 10.1109/IIAI-AAI.2016.62.
