Product Recommendation using User Preference and Online Opinion Reviews

Sankalp Jain
M. Tech Scholar
Information Technology
NIT Raipur
sjain.mtech2019.it@nitrr.ac.in

Naresh Kumar Nagwani
Associate Professor
Computer Science & Technology
NIT Raipur
nknagwani.cs@nitrr.ac.in

***Abstract*—This paper focuses on product recommendation using user preference and online opinion mining. Multiple features of a product are analyzed to get user specific product. In this paper we have analyzed user review data of mobiles sold in online markets and then predicted the most suitable mobile for the user based on preference list provided by the user. To analyze the review, the textual data is processed in multiple stages like preprocessing, feature extraction, categorizing reviews based on features, and then predicting the polarity of the review data using several machines learning model like Logistic Regression, Support Vector Machines, Naïve Bayes Classifier and Multi-Layer Perceptron Classifier. After this user entered preferences are analyzed to give user specific products.**

Keywords—Reviews, Natural Language Processing, Opinion Mining, Sentiment Analysis, Logistic Regression, SVM, Naïve Bayes, MLP classifier, Preference List.

I. INTRODUCTION

Opinion mining (OM) or sentiment analysis is data analysis technique which uses the concept of natural language processing (NLP) and computational linguistics to identify and extract opinions or sentiments from the data. Opinions can be positive, negative or neutral. According to Bin Liu for a set of documents D which contains sentiments or opinion about some object, the opinion mining technique aims to find certain components and attributes which are commented on every document d belonging to D and then further output those comments as negative, positive or neutral [1]. Now as it is clear what opinion mining is it is important to understand where to apply it. OM finds its application for both customers and organizations. In the era of online shopping consumers do not know how the actual purchase material looks and feels like, this problem can be solved by reading the reviews on the merchant site. Also, almost every production company want to know what their customers like, what they want and their expectations this problem can be solved by using concept of OM for analyzing user given reviews.

To apply opinion mining in any textual data it is important to have proper understanding of text mining and natural language processing. Text mining is the process of extracting non-trivial knowledge from several texts and documents [2]. NLP is a field which combines cognitive science, computing science, computational linguistics and artificial intelligence and uses all of them to process and understand human language to perform different tasks [3]. Both text mining and NLP are used to perform sentiment analysis. There are multiple steps involved in NLP like converting text to corpus, converting corpus to lowercase, removing of punctuation marks, common words and stop words, POS tagging etc. [4]. Further this processed data is used for feature extraction.

Feature extraction is performed to extract meaningful attributes form the text data. As review data is large file of textual data with no attributes feature extraction is done to generate these attributes. Example in the sentence "This phones camera has got good image quality, but battery is bad" camera and battery are the attributes on which comment is made. Our job in this paper is to extract features form bulk of reviews for several mobile phones and then apply opinion mining techniques to analyze consumers opinion. To complete this job different feature extraction techniques like word count vector, Term Frequency-Inverse document frequency (TF-IDF) or Bi-grams can be used. After extraction of features and gaining knowledge about the opinions associated with those features it is important to understand user preferences. User preferences are what consumers like the most or prefer the most in the product. For example, in the review "When choosing handsets, it is important for me to a have good camera. Other features like screen or battery can be normal". Here camera, screen and battery are features which can be extracted but camera is a feature which user prefers the most. This is called user preference-based sentiment analysis.

II. RELATED WORK

A. Opinion Mining and Opinion Orientation

Opinion mining can be categorized into three types, document level, sentence level and aspect level [5]. Document level opinion mining analyzes the overall document to find its polarity whether negative or positive [6]. In sentence level mining every single sentence in the document is analyzed for its polarity. This helps in fined-grained analysis of the document and for feature extraction [7]. In aspect level opinion mining classification of text data is done based on features [8]. Aspect level opinion mining consists of two parts. In the first part identification of words is done, here words refer to opinionated words. And in the second part orientation detection is performed. For example, in the mobile review "This phones camera has got good image quality, but battery is bad" there are two opinionated words camera and battery. For the camera opinion is expressed as positive but for batter the opinion is expressed as negative this is opinion orientation. Opinion orientation is finding the polarity for the opinionated words [9].

B. Natuarl Language Processing and Feature Extraction

The process of feature extraction consists of many NLP steps [10]. It starts with crawling of reviews to tokenizing them and then Parts of Speech tagging (POS tag) is performed. After POS tag is done feature identification is performed [11]. The very first work of feature extraction and text summarization in the domain of sentiment analysis was performed by Hu et al [12]. Hu et al used POS tags to identify NN (noun-noun) phrases in text. These NN phrases are the possible features for text. But This NN pair many times led to identification on many unwanted features. So pruning was done using technique of association mining to extract relevant features. Further research by Giuseppe et al introduced an unsupervised feature extraction method which used user specific information of the evaluated entity [13]. Another method which can be used for feature extraction is TF-IDF [14]. TF-IDF uses the concept of word frequency with inverse document frequency to extract valid features from the document [15].

C. User Preference Mining

After feature extraction it is important to extract individual feature opinion from the review. Based on the individual feature opinion, preferences of users can be drawn. M. Al-khiza'ay in his research used Personalized Aspect analysis model (PAAM) to get individual user preferences [16]. Few other tools are also there for summarizing reviews. For example, Latent Dirichlet Allocation (LDA), this was used by Chen et a in his research he first used classification techniques and then applied LDA to summarize the reviews [17]. S. Hu used credibility interest and sentiment enhanced recommendation (CISER) model for online product recommendation [18]. Ashima Yadav used a weighted text representation framework to analyze opinions of online medical drug reviews [19].

III. METHODOLOGY

A high-level component block diagram of the steps involved in paper is shown below.

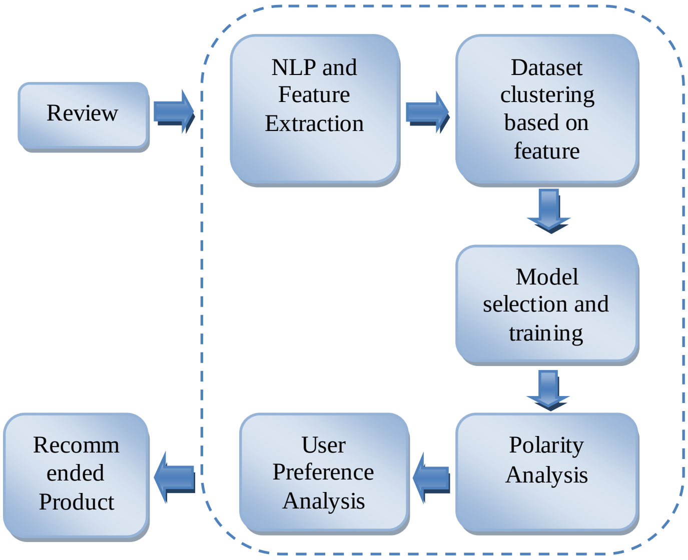

Fig. 1. High level component block diagram.

A. Review Collection

In the age of electronic commerce millions of opinions are shared by the consumers every day. These opinions are in the form of reviews available on almost every merchant site. Individuals before purchasing any product can go through these reviews to understand the opinion of different people about the product. The reviews are both helpful to consumers as well as organizations. For analyzing these reviews an individual can easily crawl these reviews from the merchant site by using any web crawling software available. After he has these reviews then natural language processing need to be performed.

B. Natural Language Processing

This step involves various sub steps to process the data. In the first step reviews are entered to NLP model. Here reviews are converted to lower case. These lowercase reviews are further processed to remove URLs, punctuations and references. After this stop word removal is done. To do this we import different packages consisting of many different stop words. These packages are then compared to the actual review data and words which are common are removed. For further processing stemming and lemmatizing operation is performed.

C. Feature Extraction

The Preprocessed data is used for feature extraction. Here TF-IDF (Term Frequency-Inverse Document Frequency) is used. TF-IDF is a measure which is used to find out the importance of word in a document. This is achieved by evaluating the product of two metrics: First is the count of occurrences of word in a particular document, and second is its inverse document frequency in a set of documents. The formula to evaluate this is given below in Eq. 1.

$$\text{Tf-idf}(t, d) = \text{tf}(t, d) * \text{idf}(t, d) \quad (1)$$

Here tf (t, d) is term frequency for a term t in document d and idf (t, d) of a term is the total number of d (documents) in the corpus divided by the df (document frequency) of a term.

D. Clustering dataset based on features

After feature extraction is performed, dataset needs to be clustered based on the extracted features. Here a single dataset is divided into multiple classes. The features extracted are known as direct attributes and can also have many derived attributes. Example: review1: “This camera captures nice photos” review2: “The images captured are awesome”. In both the reviews, consumer is talking about the quality of camera but in the review2 he referred camera as image. So, the data clustering algorithm must be intelligent enough to distribute direct and derived attributes in a common set.

E. Model Selection and Training

The problem in this paper is a classification problem. So, different machine learning classification algorithms will be used to predict the output. Algorithms to be used are mentioned below:

1) Logistic Regression: It is a stastical model which uses logistic function to model some binary dependent variable [20].

2) Naïve Bayes: Naïve Bayes is referred as probabilistic classifier which makes classification using the maximum a-posteriori decision rule as in Bayesian setting [21].

3) Support Vector Machine: The objective of SVM is to find a hyperplane in an n-dimensional space to distinctly classify the data [22].

4) Multilayer Perceptron: MLP is a class of ANN (Artificial Neural Network) which consist of multiple layers and uses back propogation. It is used to classify datasets which are not linearly seperable [23].

F. Polarity Analysis of Individual Feature

The dataset is provided as input to the above explained machine learning models. The outputs of the model are the polarity of the individual feature. Note: all the review with the rating greater than three and less than three is only used for polarity analysis. Review with rating equal to three is considered as neutral. The model outputs rating as positive or negative.

Based on polarity analysis a table is generated. The table consists of polarity and percentage polarity of each feature. This table is constructed forever single product added in dataset.

G. User Preference Analysis

User Preference is an ordered list of individual user's choice. These features are placed in decreasing order of preferences. Example: < A1, A2, A3> here A1, A2 and A3 are features of a product. The list <A1, A2, A3> says that feature A1 is the most preferred feature by user, then his second preference is feature A2, and the last one is feature A3. This list is used to search for user specific products. If a user enters above given list, then product with the maximum polarity for A1 then A2 and then A3 will be suggest.

IV. EXPERIMENTS, RESULTS AND DISCUSSION

The reviews for four different mobiles have been collected. These four mobiles are OnePlus Nord, Apple iPhone11, Mi 10 and Samsung Galaxy Note9. These mobiles have similar specifications. All these four gadgets have 100 reviews for each. These reviews are first preprocessed and then feature extraction is done. In this paper only three features are considered. These features are Camera, Display and Battery. All the reviews not consisting of any of the above-mentioned features and their derived features are removed from the database. The count for valid and invalid reviews is shown in Table I.

TABLE I. TOTAL VALID REVIEW

	Nord	iPhone11	Mi 10	Galaxy note9
Total reviews	100	100	100	100
Reviews with no features	5	6	10	2
Total valid reviews	95	94	90	98

As the list of valid reviews is generated removal of reviews which are neutral is done. Here, reviews which have rating column value as 3 are removed from the dataset. The count of remaining reviews for polarity analysis is depicted in Table II.

TABLE II. TOTAL VALID REVIEWS AFTER REMOVING NEUTRAL REVIEWS

	Nord	iPhone11	Mi 10	Galaxy note9
Valid Reviews	95	94	90	98
Neutral Reviews	9	12	12	14
Total Reviews	86	82	78	84

After removing neutral reviews, reviews are distributed in different clusters. Each cluster is represented by a particular feature. Reviews which consist of a particular feature are grouped into that feature's cluster. Example: Reviews which represent opinion for camera/image/photo are grouped in camera feature. Number of reviews which contains opinion for a feature are mentioned in Table III.

TABLE III. DISTRIBUTION OF REVIEWS BASED ON FEATURE

	Camera	Display	Battery
Nord	41	24	21
iPhone11	47	14	21
Mi 10	40	24	14
Galaxy note9	41	29	17

Now these reviews are used as input for machine learning models. Here, Logistic Regression, Naïve Bayes, Support Vector Machine and Multilayer Perceptron is used to train and test data. Train-Test splitting is done in 80-20 ratio. These models classify the input review into two polarity classes positive and negative. Polarity of each review is shown in Table IV.

TABLE IV. OUTPUT TABLE WITH POLARITY VALUES

Mobiles	Camera		Display		Battery	
	P	N	P	N	P	N
Nord	25	16	21	3	15	6
iPhone11	42	5	12	2	14	7
Mi 10	27	5	19	5	6	8
Galaxy Note9	31	10	26	3	15	2

In Table IV 'P' and 'N' are used to represent number of positive and negative reviews. After this polarity table has been generated, a positive polarity percentage table is required for user preference. Percentage of positive review of the total reviews are mentioned in Table V.

TABLE V. PERCENTAGE OF POSITIVE REVIEWS

	Camera	Display	Battery
Nord	60.9	87.5	71.4
iPhone11	89.4	85.7	70
Mi 10	67.5	79.1	42.8
Galaxy note9	75.6	89.65	88.23

As this table construction work is done it is important to analyze the models used for predicting these values. The below table shows the values obtained for Precision, Recall, F1-score and Accuracy.

TABLE VI. MODEL EVALUATION

Mobiles	Model	Precision	Recall	F1 score	Accuracy
Nord	LR	0.86	1.00	0.92	0.89
	NB	0.86	1.00	0.92	0.89
	SVM	0.86	1.00	0.92	0.89
	NN	0.83	0.83	0.83	0.78
iPhone11	LR	0.90	1.00	0.95	0.90
	NB	0.90	1.00	0.95	0.90
	SVM	0.90	1.00	0.95	0.90
	NN	0.89	0.95	0.92	0.91
Mi 10	LR	0.88	1.00	0.93	0.88
	NB	0.88	1.00	0.93	0.88
	SVM	0.88	1.00	0.93	0.88
	NN	0.88	1.00	0.93	0.88
Galaxy Note9	LR	0.89	1.00	0.94	0.89
	NB	0.89	1.00	0.94	0.89
	SVM	0.89	1.00	0.94	0.89
	NN	0.88	0.88	0.88	0.78

In the above table LR represent Logistic Regression, NB represents Naïve Bayes, SVM represents Support Vector Machine and NN represents Neural Network (Multilayer Perceptron). From Table VI it can be concluded that LR, NB and SVM give similar evaluation value in every metric. NN also gives good values but differ a bit from rest of the three models.

After evaluating all the things suggesting users with their preference list becomes easier. For this percentage table as shown in Table V can be used. Example of a user entered preference list:

List :< Camera, Display, Battery >

The above-mentioned List shows that user prefer camera as the most important feature in the mobile compared to Display and Battery. His second preference is Display and third is Battery. This list is then compared with Table V. On comparing the list, iPhone11 is the best suggested phone for user.

V. CONCLUSION AND FUTURE DIRECTION

In this paper data was collected from different electronic commerce websites. Data was preprocessed and different models were applied to evaluate the results. Some models showed good accuracy values. Although the data in which these models were trained is very less, compared to the machine learning standards. To increase the accuracy of models, increasing the size of dataset is the best option. User preference data analysis is important, as this saves a lot of time of user from going through all those bulk of reviews on the internet. A thorough research needs to be done in this area as the online markets are expanding fast and people are shifting to the digital economy. Better Natural Language Processing tools need to be developed to analyze not only common review but also review with some sarcasm. This area remains as a vast area of research, where scientific communities can come together and build better models.

REFERENCES

[1] Liu, Bing (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1–167. doi:10.2200/s00416ed1v01y201204hlt016

[2] Kao, A., & Poteet, S. R. (Eds.). (2007). Natural Language Processing and Text Mining. doi:10.1007/978-1-84628-754-1

[3] Deng, L. and Liu, Y. eds., 2018. Deep learning in natural language processing. Springer.

[4] J. Ara, M. T. Hasan, A. Al Omar and H. Bhuiyan, "Understanding Customer Sentiment: Lexical Analysis of Restaurant Reviews," 2020 IEEE Region 10 Symposium (TENSYMP), Dhaka, Bangladesh, 2020, pp. 295-299, doi: 10.1109/TENSYMP50017.2020.9230712.

[5] B. B. Alengadan and S. S. Khan, "Modified aspect/feature based opinion mining for a product ranking system," 2018 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC), Bangalore, 2018, pp. 1-5, doi: 10.1109/ICCTAC.2018.8370393.

[6] Sharma, Richa; Nigam, Shweta; Jain, Rekha (2014). Opinion Mining of Movie Reviews At Document Level. International Journal on Information Theory, 3(3), 13–21. doi:10.5121/ijit.2014.3302

[7] Hongting Li, Qinke Peng, Xinyu Guan, "Sentence Level Opinion Mining of Hotel Comm," in IEEE International Conference on Information and Automation, Ningbo, China, 2016.

[8] J. Ramakrishnan, D. Mavaluru, K. Srinivasan, A. Mubarakali, C. Narmatha and G. Malathi, "Opinion Mining using Machine Learning Approaches: A Critical Study," 2020 International Conference on Computing and Information Technology (ICCIT-1441), Tabuk, Saudi Arabia, 2020, pp. 1-4, doi: 10.1109/ICCIT-144147971.2020.9213747.

[9] J. Zhu, H. Wang, M. Zhu, B. K. Tsou and M. Ma, "Aspect-Based Opinion Polling from Customer Reviews," in IEEE Transactions on Affective Computing, vol. 2, no. 1, pp. 37-49, Jan.-June 2011, doi: 10.1109/T-AFFC.2011.2.

[10] Bialecki, Michael; O'Leary, Susan; Smith, David (2016). Judgement devices and the evaluation of singularities: The use of performance ratings and narrative information to guide film viewer choice. Management Accounting Research, (), S1044500516000159–. doi:10.1016/j.mar.2016.01.005

[11] A. Rao and K. Shah, "Model for improving relevant Feature Extraction for Opinion Summarization," 2015 IEEE International Advance Computing Conference (IACC), Banglore, 2015, pp. 1-5, doi: 10.1109/IADCC.2015.7154660.

[12] B. L. Minqing Hu, "Mining Opinion Features in Customer Reviews," American Association for Artificial Intelligence (AAAI), pp. 755–760. 2004.

[13] Carenini, Giuseppe; Ng, Raymond T.; Zwart, Ed (2005). [ACM Press the 3rd international conference - Banff, Alberta, Canada (2005.10.02-2005.10.05)] Proceedings of the 3rd international conference on Knowledge capture - K-CAP '05 - Extracting knowledge from evaluative text. , (), 11–. doi:10.1145/1088622.1088626

[14] L. Yao, Z. Pengzhou and Z. Chi, "Research on News Keyword Extraction Technology Based on TF-IDF and TextRank," 2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS), Beijing, China, 2019, pp. 452-455, doi: 10.1109/ICIS46139.2019.8940293.

[15] P. Sun, L. Wang and Q. Xia, "The Keyword Extraction of Chinese Medical Web Page Based on WF-TF-IDF Algorithm," 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Nanjing, China, 2017, pp. 193-198, doi: 10.1109/CyberC.2017.40.

[16] M. Al-khiza'ay, N. Alallaq, A. Al-Mansoori and A. R. Al-Sudani, "Personalized Reviews Based on Aspect Analysis and Polarity," 2019

8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO), Manama, Bahrain, 2019, pp. 1-6, doi: 10.1109/ICMSAO.2019.8880318.

[17] Chen, N., Lin, J., Hoi, S.C., Xiao, X. and Zhang, B., 2014, May. AR-miner: mining informative reviews for developers from mobile app marketplace. In Proceedings of the 36th international conference on software engineering (pp. 767-778).

[18] S. Hu, A. Kumar, F. Al-Turjman, S. Gupta, S. Seth and Shubham, "Reviewer Credibility and Sentiment Analysis Based User Profile Modelling for Online Product Recommendation," in IEEE Access, vol. 8, pp. 26172-26189, 2020, doi: 10.1109/ACCESS.2020.2971087.

[19] A. Yadav and D. K. Vishwakarma, "A Weighted Text Representation framework for Sentiment Analysis of Medical Drug Reviews," 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM), New Delhi, India, 2020, pp. 326-332, doi: 10.1109/BigMM50055.2020.00057.

[20] K. L. S. Kumar, J. Desai and J. Majumdar, "Opinion mining and sentiment analysis on online customer review," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016, pp. 1-4, doi: 10.1109/ICCIC.2016.7919584.

[21] M. Wongkar and A. Angdresey, "Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter," 2019 Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985884.

[22] S. Naz, A. Sharan and N. Malik, "Sentiment Classification on Twitter Data Using Support Vector Machine," 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Santiago, Chile, 2018, pp. 676-679, doi: 10.1109/WI.2018.00-13.

[23] D. Elangovan and V. Subedha, "An Effective Feature Selection Based Classification model using Firefly with Levy and Multilayer Perceptron based Sentiment Analysis," 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2020, pp. 376-380, doi: 10.1109/ICICT48043.2020.9112425.