

Consumer's Sentiment Analysis of Popular Phone Brands and Operating System Preference

Ashna Bali
BE Student
Department of Computer
Engineering
MES College of Engineering
Pune, India

Payal Agarwal
BE Student
Department of Computer
Engineering
MES College of Engineering
Pune, India

Gayatri Poddar
BE Student
Department of Computer
Engineering
MES College of Engineering
Pune, India

Devyani Harsole
BE Student
Department of Computer Engineering
MES College of Engineering
Pune, India

Nilufar M. Zaman
Assistant Professor
Department of Computer Engineering
MES College of Engineering
Pune, India

ABSTRACT

In today's era, which is symbolic of 'information overload', we are living in a time where knowledge needs to be mined from information which is available in abundance, one such instance being mining sentiments from user textual reviews which plays a decisive role in consumer's decision making processes. Today's customers are intelligent enough and will cross-check before purchasing any product and the best place for that are online reviews which are ubiquitous. In today's competitive market, analyzing user needs and sentiments is of uttermost importance as one bad review can make or break the profits of a brand since Word of Mouth plays a significant role in building a customer base for a brand or product.

An approach for consumer's sentiment analysis has been proposed, which will mine user's sentiments and thus pave a foundation for generating the popularity of the products which will lead to 'personalized' results which are necessary in today's client-centric world. This in turn is helpful for brands wanting to expand the horizon of their user turnover and to devise better retailing strategies to sell their products. The approach is to convert these textual reviews into star ratings that will describe the 'likeability' and popularity of the products.

General Terms

Sentiment Analysis, Data Mining, Opinion Mining.

Keywords

Data Mining, Text Mining, Sentiment Analysis, Information Systems, emotion, socio-affective interaction, word of mouth, online user reviews, feedback systems, sentiment polarity categorization,

1. INTRODUCTION

An exponential increase in review websites in today's age of the web has supplied us with a great possibility to share a stance for various commodities purchased. But, an information surplus problem to extract valuable information from reviews to perceive user's preferences and make an accurate and precise recommendation is essential, is faced.

Sentiment is a general realization, instinct or thought about something. In sentiment analysis, which studies people's sentiments towards certain attributes is studied. User's post

their content through various social platforms, through blogs, forums or social networks. Many social sites release their APIs, invoking data processing for analysis by researchers and developers. [18]

In many cases, numerical values are given more importance rather than decisions. Basically, positive and negative are the two sets of reviews. Selecting an option when all products consider affirmative sentiments or dissenting sentiments is challenging for consumers. For making a decision, customers not only want to understand whether the product is good, but also want to know how good it is. It's also agreed that different people may have different sentimental expression preferences. For instance, some users prefer to use "good" to describe an "excellent" product, while others may prefer to use "good" to describe a "just fine" product. [1]

2. RELATED WORK

In this section, the recent literature related to the proposed approach will be discussed. Firstly, the consequence of online word of mouth on building the brand value of a product and popularizing it will be described [8] [10], then the discussion will be about cloud customer reviews for Software as a Service(SaaS), sentimental dictionary creation for words and then classification of words.

2.1 Impact of Word of Mouth on Sales

For choosing a product, the factors that a consumer reflects upon are: the price of product which plays a decisive role in decision making, online user reviews which a person may refer to before buying a product. Due to easy availability of information about all the products, companies, on the world wide web, consumers have become more aware and make informed choices before buying any product. Maximum amount of word of mouth is generated online because today, every website has a 'comments' section and due to the anonymity offered by the web, people are far more comfortable to state their views online rather than in person. Word of mouth, however, is not only the driving forces of consumer purchase but also the outcome of product sales.

2.2 Insight on Cloud Customer Reviews for SaaS

The exponential growth of cloud services has inflated the count of online consumer reviews and opinions on these

services on different social media platforms. This gives a source of valuable information regarding cloud market position and cloud consumer satisfaction. This explores the cloud consumers' reviews that reflect the user's experience with Software as a Service (SaaS) applications. It examines SaaS consumer's reviews to identify whether they are negative, positive, or neutral in sentiment and then to train an ML classifier to build a classification model to evaluate SaaS reviews. The reviews thus play a significant role in guiding the purchases of potential consumers on cloud services. [3]

2.3 Analysis of User's Response on Smart Phone Brands and Their Operating Systems

The data expressed in the form of blogs, online forums and on Facebook, Twitter, LinkedIn and other social media platforms has given the consumers a plethora of choices to express their opinions about any product or service which very well influences other potential buyers' decision making. The approach is a combination of natural language processing, artificial intelligence and text analytics which is used to evaluate statements found on the various social media websites to determine their polarity, i.e. positive, negative or neutral which targets a particular brand, product or service. Thus, consumers assess other user's reviews for a product or a service before making their own decision. [4] [5]

2.4 Construction of Dictionary of Sentiments

The sentiments are basically determined by the polarity of words, i.e., positive, negative or neutral. The beliefs expressed in the reviews influence how people regard these entities (product, services, institutions, etc.). A product with consistently good reviews is likely to sell well, while a product with numerous bad reviews is likely to sell poorly. A dictionary containing sentimental words is made, which helps in analyzing and reviewing opinionated texts. By using a deduction approach, the resulting sentimental dictionary contains approximately 50% more words than a given sentimental word dictionary. [11]

2.5 Classification of Customer Reviews Followed by Finding Sentiment of the Reviews

The bifurcation of product reviews and service reviews will lead to easier decision making for the users to purchase the product. People often gaze over the products and reviews of the product before buying the product. But the reviews are not necessarily of products but a mixture of product of product review and service review. It builds a design that illustrates the review's sentiment in the form of charts. Thus, Classification of reviews along with sentimental analysis increases the accuracy of the system which in turn provides accurate reviews to the user. [17]

2.6 Categorization of Sentiment Analysis Techniques Without Focusing on Specific Level or Task

People's opinions are often labelled into various groups such as positive and negative from a given piece of text. The main aim of sentiment analysis is to specify whether a given text is subjective, expressing the writer's opinions, or objective. The task level of analysis ranges from document level to the

sentence level and phrase level. Various machine based learning techniques have been used for the categorization which includes: supervised learning, unsupervised learning and semi-supervised learning techniques, techniques based on lexical analysis and techniques which are a subset of these. These methodologies start by fetching the relevant texts, breaking text into parts, analyzing the spelling, and counting the frequency of specific words allowing a more economical passage from (unstructured) textual information to (structured) machine-process able data. [7]

2.7 An Approach to Sentiment Analysis of News Comments

Today sentiment analysis basically focuses on social networks, product reviews and on the financial sector. The focus on analysis of news comments is comparatively less. The audience's reaction and feelings about a particular article is not deduced from the article but from the comments the commentators give regarding the issue addressed in the article. Automatic classification can therefore be more useful than manually reading through the web. Hence, this commences on a hybrid conceptualization to sentiment analysis of news comments which involves using of sentiment lexicon for polarity detection which includes three-way polarity detection (positive, negative, neutral). [13]

2.8 Impact of Supervised Learning on Sentiment Analysis

In today's world, consumers do not make buying decision immediately on shopping website by placing an order. Instead they make use of other sources to access the product related information and then move to final step of ordering. Opinion mining is therefore an important factor of Sentiment Analysis which deals with building a system that explores the user's opinions made in blog spots, comments, reviews, discussions, news, feedback or tweets, about a product, policy, person or topic. So, it becomes easy to give the valuation and finalize the result and hence feature extraction methods is used to enumerate the sentiment polarity of each clustered comments using the supervised learning method for prior distribution. [14]

2.9 Impact of Probabilistic Machine Learning on Sentiment Analysis

The comments on a website are variant but they provide accessible and abundant data for a relatively easy analysis of a variety of applications. The techniques such as Naive Bayes and decision list are used to determine whether a given review is positive or negative. The number of stars a user gives a product is used as training data to perform supervised machine learning. Bag-of-words and bigrams are compared to one another according to how correctly they tag the reviews. Thus, Errors in classification and general difficulties regarding the selection of features are analyzed. [15]

2.10 Analyzing Classification of Sentences Into "Opinion Sentences" and "Fact Sentences" Using SVM

Recently the customer criticism has become an important factor for user accommodations making the vast use of internet sources. This has increasingly grown the user's tendency to determine whether they consult reviews to accommodations and stay. Therefore, the preprocessing of information from reviews is done by classifying the sentences into "Opinion sentences" and "Fact sentences" and we conclude that "Opinion sentences" are more efficient to

estimate values by classification of reviews into Positive and Negative and the Support Vector Machine henceforth classifies it automatically. Hence it shows that if the classification of Opinion sentences and Fact sentences was better, there was no need of classification reviews which divides the sentences into Positive and Negative by using only Opinion sentences. [16]

3. APPLICATIONS

For many years, governments and commercial organizations have been taking efforts to determine the opinions of their demographic groups and audiences. Presently, people deliberately publish their opinions on the Web openly or publicly. This social web allows for an almost instant feedback on stocks, products, policies, etc., and many of the required data, which was almost impossible to find and search earlier, is now openly available. This is in contrast with the conventional surveys and set of questions that often-dubious partakers had to fill without any self-motivation to do so, resulting in non-feasible information. People, many of them are influenced by the opinions found on the web. This is true for product reviews, which were shown to influence buying behavior. Importantly, a user understands another user's emotions and expectations, so people believe on the user's reviews more than the seller. Every person is a potential customer from producers point of view. Hence, knowing their interests and disinterests can be of great help in the manufacturing of new products, as well as management and improvement of existing ones. Moreover, understanding how the information in, for example, product reviews interacts with the details offered by the brands enables them to take leverage of these reviews and increase sales. Sentiments and beliefs on the Web have become an asset to be nurtured by companies, just like the traditional word-of-mouth. In addition, sentiment analysis is also important for other economic areas, for instance trading, finance, etc. [19]

Sentiment analysis also describes other people's perspective about a product. Hence, the first application of opinion analysis is indicating and recommending the choice of products according to the wisdom or thinking of the crowd. When a product is chosen, people are generally attracted to certain aspects or features of the product. A single rating would not be sufficient to give them a birds-eye view of the product. Sentiment analysis can regroup the opinions of the reviewers and predict ratings on certain aspects of the product. Another usage of sentiment analysis is for brands that want to mine and analyze the views of customers on their products which will give them insights on how well are their business strategies faring on the popularity front. As seen nowadays, the companies today have become more 'customer-centric', so their opinion counts the most today. These companies can then improve those areas which they think are unappealing to the consumers by devising new publicity strategies and applying business intelligence ideas. Sentiment analysis can also enumerate which aspects do customers look for before deciding to buy it. [20]

One final approach is to improve information mining in text analysis by omitting the section of the document which is high in subjectivity and automatic recommendation of internet ads for products that reflect the viewer's opinionated choices (and thus personalizing the view). Knowing what people think gives boundless possibilities in the Human-Machine interface domain as it will make it easier for a machine to understand human sentiments which will lead to a future of 'human-centric' machines which in turn will be a big step towards weaving human-machine relationships.

4. THE PROPOSED APPROACH

The purpose of the approach is to determine polarity of consumer's reviews i.e. whether feedback obtained from consumer is positive or negative. In this approach, reviews are given at client side (user side) and then a method is put forward to analyze sentiments at server side. To extract product features from user's reviews, "Bag of Words" approach is applied. The following sections describe more about this approach.

4.1 Giving Product Reviews

In this experiment framework, users can subscribe a product or products, users can write reviews for a product or products or users can view rating for products before subscribing a product. At client side, an android application is created to provide user interface. Then, user reviews are stored in JSON file. The JSON file is then sent to server for further processing. (Refer Fig.1)

4.2 Feature Extraction at Server Side

At server side, server analyses the sentiments received from client, stores analyzed data in JSON file and sends file back to the client. In the proposed system, client and server communicate via web service E.g. JSON.

Fig.1 shows the experiment framework for the proposed approach. In the proposed system, "Bag of Words" approach is used. A bag of words feature vector consists

of all the features of the product as independent features. Words in the review list are compared to the words present in this list and a dictionary is generated that maps each of the features. Table 1 shows an example of Proposed Positive and Negative Word List.

Then a weight for every sentimental word is assigned ranging from -5 to +5 as shown in Table 2.

4.3 Sentiment Calculation

Using proposed approach, the total text in its positive and negative polarity for the feature will be analyzed.

S(P) = Set of Positive Sentiment words

S (N) = Set of Negative Sentiment words

For Nth Feature -

S (P+) = (P1 + P2 + P3 +.....Pn) Set of

Positive Sentiment Words

S (N-) = (N1 + N2 + N3 +.....Wn) Set of

Negative Sentiment Words

First, the positive and negative polarity of every sentiment is calculated by calculating its probability.

After calculating the net mutual opposition of the individual features the mutual information for every feature is calculated for its final product review Analysis.

Table1: Proposed Word List

PROPOSED WORD LIST
POSITIVE: amazing, brilliant, phenomenal, excellent, fantastic, gripping, spectacular, cool, awesome, exciting, great, nice, appealing, superb, best
Negative: suck, terrible, awful, bad, boring, stupid, slow, tacky, bad, worse, poor, hopeless

Table2: Words from A Dictionary with Respective Sentiment Weight

Bad	-2
Like	+2
Poor	-2
Awesome	+2

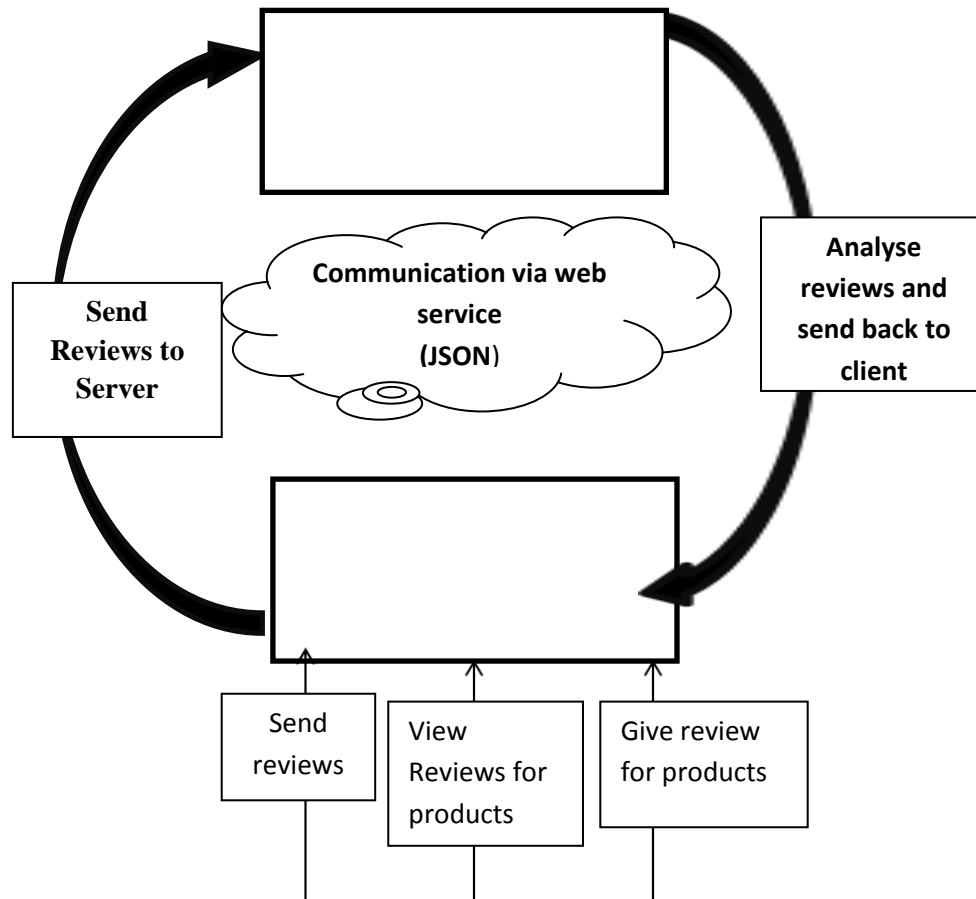


Figure 1: The Experiment Framework

5. CONCLUSION AND FUTURE WORK

In this paper, an approach for mining user sentiments for the various products they buy is proposed by using ‘Bag of Words’ model. It is aimed to provide a useful tool for all the E-commerce companies(brands) for understanding user requirements and needs. This also helps in executing the behavioral analysis of consumers for devising new strategies to make the products more ‘saleable’. The customer of today is not easily duped into becoming a ‘loyal’ customer of some website without researching or taking a feedback of it. Websites/brands today have a big job at hand to deliver according to the customers’ expectations. To attract more and more customers, the approach which has gained popularity in the past few years is ‘Sentiment Analysis’. The approach that has been suggested here is of a system which converts textual reviews of products into star-ratings to give the users a brief and ‘personalized’ idea about the feelings of other consumers who have already used that particular product, also providing the users an option of subscribing products which they want.

As things being ordered online are intangible to the user buying online, it becomes necessary to give the users as much information as possible so that they may have clear idea about that product in their head before actually buying it. These attributes may seem small to some, but these are the factors that play an important role in the consumer’s decision of buying the product or not and thus should not be over-ruled.

In the future work, a richer vocabulary set for the sentiments dictionary is considered which will make the analysis more fine-grained and universal. A machine learning approach can also be implemented where the application can understand the sentiments of the user based on his/her clicks and navigational patterns. This will be a base for Artificial Intelligent systems with machine learning. Another domain for which this approach can be used is Business Intelligence (BI) in decision making, Frequent pattern mining that can help e-commerce companies to understand user buying patterns. Another area of application can be a new product launch with a promotional

campaign to check whether the product is liked by consumers or not. This may also help in brand-building strategies to increase brand popularity.

6. ACKNOWLEDGMENTS

We would like to thank our Department and the faculty for their constant support. We would like to thank our guide and mentor Prof. Nilufar M Zaman for her guidance and help. We would also like to thank our respective families for their valued suggestions and morale boost whenever needed. Lastly, we would like to thank Almighty for giving us the energy to go on.

7. REFERENCES

- [1] Xiaojiang Lei, Xueming Qian, Member, IEEE, and Guoshuai Zhao, "Rating Prediction Based on Social Sentiment From Textual Reviews", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 18, NO. 9, SEPTEMBER 2016.
- [2] Ravendra Ratan Singh Jandail Computing Science and Engineering, Galgotias University, India, "A proposed Novel Approach for Sentiment Analysis and Opinion Mining", International Journal of UbiComp (IJU), Vol.5, No.1/2, April 2014.
- [3] Asma Musabah Alkalbani, Ahmed Mohamed Ghamry, Farookh Khadeer Hussain, Omar Khadeer Hussain, "Sentiment Analysis and classification for Software as a Service Reviews", 2016 IEEE 30th International Conference on Advanced Information Networking and Applications.
- [4] Deepali Arora, Kin Fun Li and Stephen W. Neville, Department of Electrical and Computer Engineering University of Victoria, "Consumers' sentiment analysis of popular phone brands and operating system preference using Twitter data: A feasibility study", 2015 IEEE 29th International Conference on Advanced Information Networking and Applications.
- [5] Wei Yen Chong, Bhawani Selvaretnam, Lay-Ki Soon, "Natural Language Processing for Sentiment Analysis", 2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology.
- [6] Marjan Goodarzi, Maryam Tayefeh Mahmoudi, Ramin Zamani, "A Framework for Sentiment Analysis on Schemabased Research Content Via Lexica Analysis", 2014 7th International Symposium on Telecommunications (IST'2014).
- [7] Zohreh Madhoushi, Abdul Razak Hamdan, Suhaila Zainudin, "Sentiment Analysis Techniques in Recent Works", Science and Information Conference 2015 July 28-30, 2015 — London, UK.
- [8] Duan, Wenjing, Bin Gu, and Andrew B. Whinston. "Do online reviews matter?—An empirical investigation of panel data." *Decision support systems* 45.4 (2008): 1007-1016.
- [9] Liebmann, Michael, Michael Hagenau, and Dirk Neumann. "Information Processing in Electronic Markets: Measuring Subjective Interpretation Using Sentiment Analysis." (2012).
- [10] Chevalier, Judith A., and Dina Mayzlin. "The effect of word of mouth on sales: Online book reviews." *Journal of marketing research* 43.3 (2006): 345-354.
- [11] Dragut, Eduard C., et al. "Construction of a sentimental word dictionary." *Proceedings of the 19th ACM international conference on Information and knowledge management*. ACM, 2010.
- [12] Liu, Xumin, et al. "Extracting, ranking, and evaluating quality features of web services through user review sentiment analysis." *Web Services (ICWS), 2015 IEEE International Conference on*. IEEE, 2015.
- [13] Mukwazvure, Addlight, and K. P. Supreethi. "A hybrid approach to sentiment analysis of news comments." *Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), 2015 4th International Conference on*. IEEE, 2015.
- [14] Nithya, R., and D. Maheswari. "Sentiment Analysis on Unstructured Review." *Intelligent Computing Applications (ICICA), 2014 International Conference on*. IEEE, 2014.
- [15] Rain, Callen. "Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning." *Swarthmore College* (2013).
- [16] Y. Sawakoshi, M. Okada and K. Hashimoto, "An Investigation of Effectiveness of "Opinion" and "Fact" Sentences for Sentiment Analysis of Customer Reviews," *Computer Application Technologies(CCATS), 2015 International Conference on*, Matsue, 2015, pp. 98-102.
- [17] Bhatt, Aashutosh, et al. "Amazon Review Classification and Sentiment Analysis."
- [18] Fang, Xing, and Justin Zhan. "Sentiment analysis using product review data." *Journal of Big Data* 2.1 (2015): 1.
- [19] Schouten, Kim, and Flavius Frasincar. "Survey on aspect-level sentiment analysis." *IEEE Transactions on Knowledge and Data Engineering* 28.3 (2016): 813-830.
- [20] Collomb, Anais, et al. "A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation."