

#### **Open Access**

# Sentiment Analysis of Pharmaceutical Products Evaluation Based on Customer Review Mining

#### Mahboob K\* and Ali F

Department of Software Engineering, Sir Syed University of Engineering and Technology, Karachi, Pakistan

# Abstract

With the emergent of advanced technologies, several methods of extracting huge amount of information have been adopted for many applications and information processing. One of which is mining opinion (or sentiment analysis) that is useful in natural language processing, text mining, and speech recognition, etc., which is widely functional in different information industries. The Online Social Media has created new multiple avenues for sharing of different kind of information among folks for social awareness and sharing views among them. The research based in the context of customer review mining on online various pharmaceutical products evaluation with respect to the frequency of a particular medicine usage, customer personal satisfaction, influential marketing products strategy, online pharmaceutical suggestions and prescriptions, and the impact of medicines on health, etc. The research focus on sentiment analysis of multiple customer reviews on different medicines to evaluate the quality and satisfaction of that product. Further, sentiment approach has been assumed that can help pharmaceutical enterprises in improving and maintaining the quality of product services based on the review of customer. The research aims to use lexicon-based approach for sentiment classification according to positive, negative or neutral type of polarity but to a limited extent. The accuracy and reliability of results is a challenge when parsing complex sentences and words due to insufficient undistinguishable data.

**Keywords:** Drug; Medicine; Pharmacy; Opinion mining; Rating; Customer review; Social media; Sentiment; Polarity

## Introduction

Social media is playing a crucial role in this era nearly in every field like education, health, medical sciences, e-commerce, marketing, finance, travel, demographics, etc. [1]. With the growth of online social networks, everyone can easily interact and engage through different modes and mediums of communication with each other and get the latest updates of information on different fields. One of the major adoptions in a field of information technology is the sentiment analysis which targets to determine the context [2], content and sensitivity of data with respect to any topic or the overall contextual divergence of a text or file. To extract and identify subjective knowledge on huge volumes of digital up-to-date data, Sentiment Analysis (also known as Opinion Mining) is vastly applied in a wide variety of different applications, extending from person to person. This can be well-thought-out as a criterion of judgment and evaluation of a person approach towards subject or topic [3]. Thus, opinion mining or sentiment analysis is used to deal with cost-effective and detailed information [2] related to extracting large volumes of data that helps to determine the product sentiment ratings (also particularly in pharmaceutical industries). The online blogs and reviews are helpful measure for rating the product and its impact day -to-day among people socially.

With the context of health and pharmacological sciences, the patients may also be able to share their reviews and experiences of indications and impact of a medicine so that it can help to rate the medicine according to its usage, cost and chronic effects [4-16]. Several studies carried out to concentrate on the patient's information and their related matters particularly reviewing patient's medication data [12] with varying costs and usage of a drug. In this research, we intend to address customer review mining for pharmacy product evaluation and planned a general pattern (based on lexicon-based approach) using sentiment detection tool for mining the customer reviews [17] to evaluate performance measure of a pharmacy product with grading

The research engaged the individual's users to show online opinions of other users about medicine potential influence. With the explosion of Web 2.0 platforms [3] such as discussion forums, peerto-peer networks, blogs, and various other types of social media [13], consumers are able to share their product experiences and opinions, positive or negative reviews, regarding any product or service [2]. Marketers are always needed to observe information related to their products to gain knowledge of sales, popularity according to usage or demand, etc., which must be necessary to improve the quality of the product and market.

#### **Related work**

**Contextual analysis of health content:** Through online social networking, the communication is vastly improved and different interest of information is available on internet easily at the open pace. Different kind of information needs to share to chronicle highlights of potential benefits and harms and availability of utilities of certain insights, items, people behaviours, products, etc. [12]. One of the important fields is the medical and health sciences to consider social aspects through online discussions, blogs, reviews, and online surveys, etc. [16]. The health-related content shared through various online feedbacks or reviews contains hidden sentiment [18] patterns that are

\*Corresponding author: Mahboob K, Department of Software Engineering, Sir Syed University of Engineering and Technology, Karachi, Pakistan, Tel: +922134988000; E-mail: nedian07@gmail.com

Received March 27, 2018; Accepted April 05, 2018; Published April 30, 2018

Citation: Mahboob K, Ali F (2018) Sentiment Analysis of Pharmaceutical Products Evaluation Based on Customer Review Mining. J Comput Sci Syst Biol 11: 190-194. doi:10.4172/jcsb.1000271

**Copyright:** © 2018 Mahboob K, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

functionalities by analysing different blogs and posts.

need to be identified and extracted through different sources ranging from marketing to service the drugs in a competent way [14]. In this regard, the online mechanism is very popular these days for online shopping, different products through different websites like online purchasing of medicine at door step. After purchases, several websites and blogs offer customers rate their products according to their satisfaction and quality of products and services and also by providing feedback facility by which customer can comment on a particular medicine or on quality of services.

The number of posts available online are analysed to find out responses based on customer reviews available on different online pharmaceutical stores websites in order to detect hidden sentiments on unstructured text data. The information is classified into three forms of orientation as Satisfied, Dissatisfied, and Disinterested based on the result satisfaction obtained by the sentiment classifier from the different datasets [15]. However, the results vary from domain to domain; therefore, a general lexicon is considered using drugs reviews from different sites and blogs related to health, diseases, and drugs [17].

The online reviews area is widespread that provides economical and flexible results by proposing primarily based probabilistic modelling structure [7]. A basic task is to analyse and classify the divergence of a given text at the word, sentence or feature/aspect level related to health content [16]. Several posts are analysed and tested from different online social networking websites like Facebook, Twitter, Google+, etc. [12] showed opinion mining based on customer reviews and ratings is possible with the combination of generalized as well as domain-specific sentiments achieve the desired results.

Mechanism of medicine evaluation: Several measures have been taken to find out the favourable and adverse effects of drugs using a classifier. Numerical ratings are used on a drugs data and the data was evaluated using numeric star rating and word count in a particular or number of reviews associated with each medicine to evaluate the performance of a medicine and experience of a customer with that drug. The mechanism to find out the accuracy, quality, effectiveness, limitations, adversities related to medicine and health was obtained using association rule mining to evaluate correlations among different parameters between drugs and favourable or adverse reactions mentioned in posts online through customer experiences and reviews [1]. A lexicon based manual annotation was developed of user's posts and equitable accuracy using lexical matching was achieved [17-19]. Comparative product review mining was also done by analysing pros and cons from a big volume of customer reviews online through different pharmaceutical stores websites [3]. Different models were testified to mine valuable customer info intelligently, such as perspectives extraction contrastively, high-quality review recommendation [20], review summarization, etc. CRFs model was used to extract different product information, related reviews in order to evaluate the ranking of a medicine currently and in a near future.

# Methods

# Data requirement analysis

As previously discussed there are several online platforms (social networking websites, blogs, forums, etc.) are excessively available where people interact and communicate with each other among the social network and usually involves in different activities from chatting to uploading profile pictures or documents, downloading files, liking, disliking posts (like on Facebook, Twitter) [20], giving reviews on certain topics or products. Particularly, in the domain of

pharmacy, there are several online websites that deal with large number of customers involved in selling and buying different medicines and customers may place reviews about different pharmaceutical products or the purchases they have made. The customer reviews actually is the symbol of representing their satisfaction or dissatisfaction with the product or service [10]. Therefore, the pharmaceutical reviews in the form of text is required to perform sentiment analysis and computation that would help customer as well as business to identify and improving the quality and features probed by the consumer behaviour (in the form of review) [18].

The research intends to evaluate the quality of drugs of different categories that would be helpful in purchasing decisions. To conduct a research, two online websites reviews data on different drugs are analyzed. Sentiment approach is applied on various customer reviews from the two different websites on different medicines and assesses the reviews either holds positive, neutral, or negative words in order to find out the polarity in the review.

## Data collection

The methodology is used in this research is the collection of different online reviews from customers on different medicines to perform sentiment analysis using lexicon-based technique (specifically corpus-based approach that helps in solving different problems of determining opinion words by means of context specific orientations), in order to determine the polarity associated with the sentences or words in the customer reviews [4]. The purpose to identify the polarity in a sentence is helpful to find out customer likes and dislikes towards products. In this way, one can distinguish among good and bad medicines in order to improve the quality and productivity of the medicines based on customer reviews sentiment analysis [16]. Simply, the data containing customer reviews on pharmaceutical products have been gathered from two different online websites. The drugs are grouped into nine different categories of allergy and sinus, children's healthcare, cough, cold and flu, dehydration, digestive health, eye and skin treatments, pain and fever, sexual wellness and women health. In each category, reviews on approximately three drugs have been analysed for sentiment evaluation. For eight categories, the reviews are collected from livewell.pk [11] and for category "Sexual Wellnes" the data is gathered from kaymu.pk [9]. The reviews are then separately analysed for each category. Table 1 contains the list of medicines with nine categories.

## Sampling methods

There are various tools and techniques of sentiment analysis to detect hidden emotions in any phrase or a sentence to find out the associated polarity with them. Sentistrength is one of the tools which are used to perform sentiment analysis specifically selected for this research. It basically estimates the strength in a phrase or in a sentence as a positive, neutral or negative sentiment in the text. It practices a sentiment lexicon for assigning scores to positive, neutral or negative phrases in the text [5]. It essentially organizes human -level accuracy for small social web texts in English, except political texts. SentiStrength principally reports two types of sentiment strengths [6].

-1 (not negative) to -5 (extremely negative) 1 (not positive) to 5 (extremely positive). SentiStrength can also report binary (positive/negative), trinary (positive/negative/neutral) and single scale (-4 to +4) results [6]. SentiStrength was originally developed for English language and optimized for general small social web texts but can be configured for other languages and contexts by changing its input files [6]. The

trinary methodology is basically applied in this research on different medicine reviews. The review on each medicine in each category was analysed by splitting sentences or phrases into parts separately (with respect to full stop). Each sentence is then drastically was given a trinary test to detect whether the text hold the positive (+1), neutral (0) or negative (-1) sentiment. The (+1) indicates the positive sentiment, (0) indicates the neutral sentiment and (-1) indicates the negative sentiment as a result in a customer reviews [3]. The positive, neutral or negative score of each sentence in a review is computed by averaging the scores together. Thus, the sentiment of a given review is measured by averaging the trinary results of all sentences within that review or comment individually medicine wise and category wise separately.

## Results

The medicine wise and category wise average polarity (trinary results) are presented in Table 2. All the nine categories are presented as major evaluation metric for the medicines based on customer reviews. Each term or word in a sentence is analysed by parsing lexicon based trinary sentiment method to classify the sentence as positive, neutral or negative sentiment [5]. The sentiment term hold the sentiment score according to the strength. There are three cases to compute sentiment score in each sentence in a review. In first case, the overall result (+1) is determined by positive>(-3) if possible score. terms are greater than negative score terms, the overall polarity of a sentence is (+1) i.e., positive. In second case, the overall result (0) is determined by positive=1 negative= (-1) if positive score terms and negative score terms are equal means no sentiment term is detected or found, the overall polarity of a sentence is (0) i.e., neutral [6]. Likewise, in third case, the overall result (-1) is determined by positive<negative if positive score terms are lesser than negative score terms, the overall polarity of a sentence is (-1) i.e., negative [6].

The graph in Figure 1 illustrates the category wise result of medicines with comparison to the number of reviews in each category with respect to positive, neutral and negative polarity. The result presented graphically, clearly specifies the categories with totally positive reviews, positive and neutral reviews, negative and neutral reviews and mixed reviews among different medicines in each category according to the polarity distribution.

The precision is defined as a metric for the evaluation of medicines based on customer reviews for individual category as follows [4]:

Precision=Number of Positive Reviews for each Medicine Category/Total Number of Reviews for each Medicine Category

All nine medicine categories have been presented to extract positive reviews from each category [19]. The customers have posted their reviews on medicines belonging to different categories [4]. The positive reviews particularly have been selected in order to determine the most popular category of medicines evaluated with respect to customer satisfaction in terms of product quality, price, usage, and impact. Category 6 and 8 has been positively evaluated by 100% precision value which clearly embodies the most satisfactory and consistent medicine categories.

Category 1 was positively rated 67% precision value which shows the average medicine category. Category 2, 4, 5, 7 and 9 were estimated 33% precision value which indicates the below average or inconsistent medicine categories with the customer reviews perspective [3]. Category 3 was rated 0% precision value which apparently shows the most unsatisfactory medicine category [4]. A graphical depiction of precision value (in %) of nine medicine categories are shown in Figure 2. The total number of positive reviews and their precision (in %) is tabulated and shown in Table 3.

# Conclusion

In this paper, the sentiment analysis using lexicon-based approach is presented on various pharmaceutical products from two different websites based on customer review mining. Nine distinguished categories of medicine were chosen to apply corpus-based sentiment method using SentiStrength tool and trinary results were obtained from the customer reviews on medicines and polarity was found as positive, neutral or negative by breaking sentences in a review and applying sentiment method to detect sentiment related with the sentences in a review. Then, finally averaging the scores of sentences in a customer review about medicines were computed. The high percentage of precision implies the most satisfactory and consistent medicine category in terms of usage, quality, price and frequency of purchases. However, in some reviews, the sentences are not fully analyzed for sentiment detection.

As presented in this research through experiment, the medicine categories: eye and skin treatments and sexual wellness gives the highest and satisfactory precision as compared to other categories based on the evaluation of online customer reviews. A future challenge is to fully analyse the sentences and phrases in reviews in order to develop proper accurate sentiment rating mechanism and the research can extend to more pharmaceutical categories from several resources.







J Comput Sci Syst Biol, an open access journal ISSN: 0974-7231

Category 'Allergy & Sinus'	Telfast BD Fexofenadine HCI Tablets			
	Kestine Ebastine Tablets			
	Senergy-OD Tablets			
Category Calpol 6 Plus Suspension Syrup	Calpol 6 Plus Suspension Syrup			
	Disprol Paediatric Suspension Syrup			
	Tempol 6+ Suspension Syrup			
Category 'Cough, Cold and Flu'	Strepsils Original Lozenges			
	Actifed-P Cold Tablets			
	Fexet Fexofenadine HCI Tablets			
Category 'Dehydration'	Orazaf Powder			
	OEM Orange Powder			
	Peditral Lemon Powder			
Category 'Digestive Health'	Eno Lemon Powder			
	Mucaine Suspension Syrup			
	Dicaine Suspension Syrup			
Category 'Eye & Skin Treatments'	Xyzal Levocetirizine Dihydrochloride Tablets			
	Polyfax Skin Ointment For External Use Only Antiseptic			
Category 'Pain & Fever'	Paracetamol Suspension Syrup			
	Panadol Extra (Paracetamol with Caffeine) Tablets			
	Voltral Emulgel Gel			
Category 'Sexual Wellness'	Pfizer Viagra Pills			
	Fertilex (For Men)			
Category 'Women Health'	Fertile Fast Conception Capsules			
	SlimEasy Tablets			
	Nutrifactor Evening Primrose Oil			

	Medicine wise	Category wise					
Cotogonico with Madising Listing	Average	Average					
Categories with medicine Listing	Polarity	Polarity					
	(Trinary Results)	(Trinary Results)					
Category 'Allergy and Sinus'							
T elfast BD Fexofenadine HCI Tablets	(0) Neutral						
Kestine Ebastine Tablets	(1) Positive	(1) Positive					
Senergy-OD Tablets	(1) Positive						
Category 'Children's Healthcare'							
Calpol 6 Plus Suspension Syrup	(0) Neutral						
Disprol Paediatric Suspension Syrup	(-1) Negative	(0) Neutral					
T empol 6+ Suspension Syrup	(1) Positive						
Category 'Cough, Cold and Flu'							
Strepsils Original Lozenges	(0) Neutral						
Actifed-P Cold Tablets	(-1) Negative	(-1) Negative					
Fexet Fexofenadine HCI Tablets	(0) Neutral						
Category 'Dehydration'							
Orazaf Powder	(0) Neutral						
OEM Orange Powder	(1) Positive	(1) Positive					
Peditral Lemon Powder (0) Neutral							
Category 'Digestive Health'							
Eno Lemon Powder	(1) Positive						
Mucaine Suspension Syrup	(-1) Negative	(-1) Negative					
Dicaine Suspension Syrup	Dicaine Suspension Syrup (-1) Negative						
Category 'Eye and Skin Treatments'							
Xyzal Levocetirizine Dihydrochloride Tablets	(1) Positive	(1) Positive					
Polyfax Skin Ointment For External Use Only Antiseptic	(1) Positive						
Category 'Pain and Fever'							

#### Citation: Mahboob K, Ali F (2018) Sentiment Analysis of Pharmaceutical Products Evaluation Based on Customer Review Mining. J Comput Sci Syst Biol 11: 190-194. doi:10.4172/jcsb.1000271

Paracetamol Suspension Syrup	(1) Positive					
Panadol Extra (Paracetamol with Caffeine) Tablets	(-1) Negative	(-1) Negative				
Voltral Emulgel Gel	(-1) Negative	(-1) Negative				
Category 'Sexual Wellness'						
Pfizer Viagra Pills	(1) Positive	(1) Desitive				
Fertilex (For Men)	(1) Positive (1) Positive					
Category 'Women Health'						
Fertile Fast Conception Capsules	(0) Neutral					
SlimEasy Tablets	(1) Positive	(0) Neutral				
Nutrifactor Evening Primrose Oil	(-1) positive					

Table 2: Average Polarity Results.

	Allergy and Sinus	Children's Healthcare	Cough, Cold and Flu	Dehydration	Digestive Health	Eye and Skin Treatments	Pain and Fever	Sexual Wellness	Women Health
Number of Positive Reviews	2	1	0	1	1	2	1	2	1
% Precision	67	33	0	33	33	100	33	100	33

Table 3: Number of Positive Reviews with precision (in %).

## **Conflicts of Interests**

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

#### References

- Khader N, Lashier A, Won YS (2016) Pharmacy Robotic Dispensing and Planogram Analysis Using Association Rule Mining with Prescription Data. Journal of Expert Systems with Applications 57: 296-310.
- 2. Pang B, Lee L (2008) Opinion Mining and Sentiment Analysis. Foundations and Trends® in Information Retrieval, pp: 1-135.
- 3. https://en.wikipedia.org/wiki/Sentiment analysis [Accessed on: July 2017].
- Denecke K, Nejdl W (2009) How valuable is medical social media data? Content analysis of the medical web. J Inform Sci 179: 1870-1880.
- Wiley MT, Jin C, Hristidis V, Esterling KM (2014) Pharmaceutical drugs chatter on Online Social Networks. Journal of Biomedical Informatics 49: 245-254.
- Goeuriot L, Na JC, Kyaing WYM, Khoo C, Chang YK, et al. (2012) Sentiment lexicons for health -related opinion mining. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, pp: 219-226.
- Salehan M, Kim DJ (2015) Predicting the Performance of Online Consumer Reviews: A Sentiment Mining Approach to Big Data Analytics. Decision Support Systems 2: 163-167.
- Balahur A, Montoyo A (2008) Determining the semantic orientation of opinions of products - a comparative analysis. Procesamiento del Lenguaje Natural 41: 201-208.
- Anto MP, Johny N, Antony M, James V, Muhsina KM, et al. (2016) Product Rating using Sentiment Analysis. International Conference on Electrical, Electronics and Optimization Techniques.

- Qi J, Zhang Z, Jeon S, Zhou Y (2016) Mining Customer Requirements from Online Reviews: A Product Improvement Perspective. Information and Management 53: 951-963.
- 11. Swan M (2012) Scaling crowdsourced health studies: the emergence of a new form of contract research organization. J Pers Med 9: 223-234.
- Popescu AM, Etzioni O (2005) Extracting product features and opinions from reviews. Proceedings of the conference on human language technology and empirical methods in natural language processing. Association for Computational Linguistics, pp: 339-346.
- Lorenzo-Romero C, Constantinides E, Brunink LA (2014) Co-creation: customer integration in social media-based product and service development. Procedia Soc Behav Sci 148: 383-396.
- Mei Q, Ling X, Wondra M, Su H, Zhai C (2007) Topic sentiment mixture: modeling facets and opinions in weblogs. Predictive Modeling of Web Users, pp: 171-180.
- Sohail SS, Siddiqui J, Ali R (2016) Feature Extraction and Analysis of Online Reviews for the Recommendation of Books using Opinion Mining Technique. Perspectives in Science 8: 754-756.
- 16. https://www.livewell.pk/ [Accessed on: August 2017].
- 17. http://www.kaymu.pk/medicine-drugs/ [Accessed on: August 2017].
- Aronson AR (2001) Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. In: AMIA annual symposium proceedings. Washington DC: American Medical Informatics Association, pp: 17-21.
- 19. http://sentistrength.wlv.ac.uk/ [Accessed on: August 2017].
- Van De Belt TH, Engelen LJ, Berben SA, Schoonhoven L (2010) Definition of Health 2.0 and Medicine 2.0: a systematic review. J Med Internet 12: e18.