

# Survey paper on Sentiment Analysis: In General Terms

Abdullah Dar\*, Anurag Jain  
Computer Science & RGPV  
India

## Abstract—

**T**his paper deals with the study of sentiment analysis here the term sentiment analysis is used for identifying sentiments and opinion of the users. In this paper we are going to understand about the different methodology used for sentiment analysis and opinion mining. In the few passed years the scope gets extended to analyze pattern of sentiment analysis. This paper gives the detail of recent study and also spread light on the terms opinion mining and sentiment analysis. The term ontology is used in our paper it deals with arrangement of words we are using ontology based supervised (predefined) learning technique. Helpful to calculate sentiments about (game review , vehicle review, movie review ,product review etc ).Some sentiment analyser are language dependent here, language can be English, German, French,, Chinese or some could be language independent e.g. social networking site sticker store, twitter sentiment analyser etc Made machine learning simpler.

**Keywords—** Ontology, Sentiment, Analyser, Polarity, Sway.

## I. INTRODUCTION

Analysers used for polarity identification. Analysers are of two types manual (domain oriented) and automatic (generalized oriented) we used domain oriented in are methodology. In manual analyser predefined data set exit which similar/ related term have to feed and result occurs and other hand automatic analyser consist huge data set and also capable to handle multiple language at a time. Sentiment analysis is used to classify polarity and the sentiment analyser is used to define polarity opinion expressed is (+) tive, (-) tive or (=) neutral[1].

Tagged reviews are also the important approach for sentiment analysis. After the polarity of the sentiment is identified, another not much difficult but important task is presentation of results.

In our work the sway analysis is used to find out the inclination. Here the word sway is used to define tree and node variations helpful for product ranking and feature extraction.

Large number of users ports their contribution on the World Wide Web [2]. World Wide Web inform of user – reviews for different items ranging from mobile phones, holiday trips, hotel services, movie reviews etc [2]. Intelligent transportation system (IIS) used for traffic management and analysis [3].

With the remarkable advancement of web 2.0 in the last decade, communication platforms, such as blogs, wikis, online forums [3]. Extracting Opinion from text is called opinion mining (OM) [4]. Online community is highly co-related to the fluctuation of the stock index [5]. The online brand community is help firms ( e.g., Hardly-Davidson) to become a global leading brand [5].

The social network has become an increasingly important application in recent years [6]. The ranks of their pages are the measure of their success in social network marketing [6].

## II. FROM WHERE DATA IS OBTAINED[SOURCE]

The major source for obtaining data is websites, software, fact, figures etc. Huge data can be found about company, consumers, producers, retailers, legal documents, data warehouse etc. Opinion mining is used to analyze this data. Details are given below.

### A. Web world

In the world of web there are huge amount of data is present this data is the backbone for researchers and users both in the world of web blogs, comment box, form feeding etc techniques is used for data fetching and to transfer data emails, face book, twitter is used.

### B. Visit related sites

In the current scenario many recognized groups are doing the work of analyzing data as well as maintaining the sites. Industries are keeping people to performing the respective work. Finally on the bases of obtained data such as price, quantity, ranking the product result occurs.

### C. Web based Interface

Interface is a medium between user and the web this consist text messaging, digital audio/video, e-mail, links etc. Interface plays an essential role between user and web because without interface it is not possible to connect or communicate.

**D. Related Terms:**

This part used to describe terms and work to spread light on different approaches for discovering and identifying sentiments, opinion, expression or appearance.

Multiple approaches are present for sentiment detection and analysis.

Few approaches are as described below:

**III. OVERVIEW OF MACHINE LEARNING APPROACH**

In machine learning algorithms are studied as well as constructed and the source is data by using them prediction (event anticipation) and decision (opinion) can be making on the other hand machine is based on instruction.



Machine learning and pattern reorganization are classified as two faces – Approaches are as follows-

**A. Association rule learning:**

In association rule learning the association is finding between different variable. Variables are having different value its changes never be constant. This type of relationship can be finding in huge database where organized collection of data is present.

**B. Decision tree learning:**

It is a tree based approach collection of child and root node which focus on the target value.

**C. Artificial neural network:**

This approach deal with biological neural network based on neurons in this approach pattern matching and statistical data modelling take place.

**D. Naïve bayes method:**

Naïve bayes is a simple type of classifier which based on probability having group of random variable and it is conditional independent. We can find out the probability between the disease and symptom.

**E. Support vector machine (SVM):**

It is used to find out relationship and classification of variables. Here the process of stamina-ting relationship among variable is called Regression.

It consist marked value belonging to weather first category or the second category.

**F. Classification of semantic –orientation Approach (Based on lexicon/Corpus based):**

The corpus deal with structured text and speech all the illegal process and suspect detected by it. A legal term abbreviates “body of the crime”.

Huge corpus is present to calculate outcome in the form of statistical information. It is not efficient as compare to directory based approach.

**G. Directory based approach:**

This approach deal with the synonyms and antonyms but in the recent study sentiwordnet is used in which words are positive, negative, neutral. Sentiwordnet used as language representation lexicon.

TABLE I. DESCRIPTION OF CONCEPT USED IN PAST

S No	Publicati on Year	Authors	Concept
01	2k14	Jianping Cao, Ke Zeng, Hui Wang	Intelligent transport system (Rule based approach) e.g. Online data collection, architectural design etc.
02	2k14	Hanjun Lee, Suveon Jeony, Yongmoo Suh	Reduce and calculate negative emotion through POS tagging.
03	2k14	Phan Trong Ngoc Myungsik	Lexicon – based sentiment analysis use clawer get data from destination. Clawer is used to get data from destination. Clawer consist normal score calculated by lexicon and power score face book data based on number of likes.
04	2k13	Ms Kranti Ghand and Dr. Ketan Shah	Lexicon based analyser (huge data set). Lexicon represents language. The language inventory represented by lexemes here lexemes is the unit it may be the number of words lexicon means of or for words.
05	2k13	V.k. Singh, R.Riryani, A.Uddin and P.Waila, Mansha	Naïve Bayes and SVM. It is a supervised based probabilistic approach.
06	2k13	Neelam Mukhtar, Sara Shahzad, Mohammad Abid Khan, Shah Nazir	Feature based selection of web development.

07	2k12	Farek Lazhar* and Tlili Guaiassa Yamina	Describe the interrelationship between the entities it represents domain oriented concept consist set of values. The general term used in it is lexicon.
08	2k12	Wei Wei and Jon Atle Gulla	The logic behind this paper is HHCP approach. Examine during this study.
09	2k12	Sivaraman Sriram, Xiaobu Yuan	Data set is the collection of elements is to provide general, easy, powerful and control model is used.
10	2k12	Pilar Rodriguez, Alvaro Ortigosa and Rosa M. Carro	In this paper student analysed the concept of ontology.
11	2k11	Hana Jeong, Dongwook Shin, and Joongmin Choi	Distinct characteristics (feature extraction) taken place in this approach. The extraction of right feature by the use of both grammar properties as well as semantic properties.
12	2k11	Usman Tariq, Student Member, IEEE, Kai-Hsiang Lin, Zhen Li, Xi Zhou, Zhaowen Wang, Vuong Le	This paper deal with detection and identification such as identifying the face and detecting the size of face this property is known as visual property. Automatic face detection is its advance approach.
13	2k11	Peter Koncz and Jan Paralic	This approach is good and efficient but accuracy of this approach vary consist random error.
14	2k11	Alexander Hogenboom, Paul van Iterson, Bas Heerschop, Flavius Frasinca, and Uzay Kaymak	It is helpful to determining the negation keyword scope used in sentiment analysis. Example here determine x and not x.
15	2k11	Xiaojun Li, Shanshan He, Hanxiao Shi	This approach give improve accuracy rate used for polarity finding.
16	2k11	Richard Colbaugh and Kristin Glass	This approach deals with high security by analysing public opinion in different languages such as German, French, Arabic etc.
17	2k10	Many authors participated in workshop	Language is the medium of communication which analysed by speech, touch, speaking etc. This approach finds out the emotion present in text. We can understand the concept of natural language.
18	2k09	FAN Na , CAI Wan-dong, ZHAO Yu	It is based on generation model used to refine and analysis text topic identification.
19	2k09	Miho Itoh	Relationship is the linking between the words. Word is the collection of alphabet and through words sentence formed. Which get analysed by this approach.

#### IV. CONCLUSIONS

In general term sentiment analysis popularly also known as opinion mining. Sentiment analysis is helpful in different field for calculating, identifying and expressing sentiment.

There are lots of research work and more work remaining in this field to elaborate and there are work also present which is not get solved yet such as use of negation with maximum accuracy, give detail about items, use of multiple languages at a time.

More and more work done in future which make sentiment analysis nearer to accuracy.

#### ACKNOWLEDGMENT

I am very thankful to my guide and the whole faculty team to make me aware and guide me time to time. I hope in future the treat us in the same way they make us understand how to work and think.

#### REFERENCES

- [1] Jingyi Zhang, Yanhui Lv.” An Approach of Refining the Merged Ontology”. 2014 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2014).
- [2] Hanjun Lee, Suyeon Jeong and Yongmoo Suh, The Influence of Negative Emotions in an Online Brand Community on Customer Innovation Activities, 2014 47th Hawaii International Conference on System Science, 978-1-4799-2504-9/14 \$31.00 © 2014 IEEE DOI 10.1109/HICSS.2014.234
- [3] Phan Trong Ngoc and Myungsik Yoo, The Lexicon-based Sentiment Analysis for Fan Page Ranking in Facebook, 978-1-4799-3689-2/14/\$31.00 ©2014 IEEE
- [4] Ms Kranti Ghag and Dr. Ketan Shah, Comparative Analysis of the Techniques for Sentiment Analysis, ICATE 2013 Paper Identification Number-124
- [5] V.K. Singh, R. Piryani, A. Uddin P. Waila, Marisha Sentiment Analysis of Textual Reviews, 2013 5th International Conference on Knowledge and Smart Technology (KST),978-1-4673-4853-9/13/\$31.00 ©2013 IEEE
- [6] Neelam Mukhtar, Sara Shahzad, Mohammad Abid Khan, Shah Nazir, Ontology for Feature Based Selection of Web Development Tools, 978-1-4799-0615-4/13/\$31.00 ©2013 IEEE
- [7] Farek Lazhar, Tlili Guaiassa Yamina “Identification of Opinions in Arabic Texts Using Ontologies”, J Inform Tech Soft Engg Volume 2 • Issue 2 • 1000108 ISSN: 2165-7866 JITSE, an open access journal,2012.
- [8] Wei Wei and Jon Atle Gulla,” Sentiment Analysis In a Hybrid Hierarchical Classification Process”, IEEE 978-1-4673-2430-4112/\$31.00, 2012.

- [9] Sivaraman Sriram, Xiaobu Yuan,” An Enhanced Approach for Classifying Emotions using Customized Decision Tree Algorithm”, 978-1-4673-1375-9/12/IEEE 2012.
- [10] Pilar Rodriguez, Alvaro Ortigosa, Rosa M. Carro,” Extracting Emotions from Texts in E-learning Environments”, 978-07695-4687-2/12 DOI 10.1109/CISIS.2012.192.
- [11] Hana Jeong, Dongwook Shin, and Joongmin Choi” Feature Extraction and Refinement for Opinion Mining”, ETRI Journal, Volume 33, Number 5, October 2011.
- [12] Usman Tariq,” Recognizing Emotions from an Ensemble of Features”, VOL. 42, NO. 4, AUGUST 2012 1083-4419/ 2012 IEEE.
- [13] Peter Koncz and Jan Paralic,” An approach to feature selection for sentiment Analysis”, 978-1-4244-8956-5/11/JUNE 2011 IEEE.
- [14] Alexander Hogenboom, Paul van Iterson, Bas Heerschop, Flavius Frasincar, and Uzay Kaymak,” Determining Negation Scope and Strength in Sentiment Analysis”, 978-1-4577-0653-0/11/2011 IEEE.
- [15] Xiaojun Li, Shanshan He, Hanxiao Shi,” Construction and Quantization for a Basic Sentiment Lexicon”, 978978-1-61284-181-6/11/2011 IEEE.
- [16] Richard Colbaugh, Kristin Glass,” Agile Sentiment Analysis of Social Media Content for Security Informatics Applications”, 978-0-7695-4406-9/2011 IEEE DOI 10.1109/EISIC.2011.65
- [17] Diana Inkpen, Carlo Strapparava,” Workshop on Computational Approaches to Analysis and Generation of Emotion in Text”, NAACL HLT 2010.
- [18] FAN Na, CAI Wan-dong, ZHAO Yu,” A Method based on Generation Models for Analysing Sentiment-Topic in Texts”, 978-1-4244-4507-3/09/2009 IEEE.
- [19] Miho Itoh,” Contextual Analysis Processing Methods able to interpret Sentiments Evaluation Representations”, 978-0-7695-3800-6/ 2009 IEEE DOI 10.1109/ICSC.2009.98.