



Design of Feedback Analysis System for Deciding Trustworthy E-Commerce Organization

Sonali J.Bagul, Rakhi D.Wajgi

M. Tech Student, Dept. of Computer Science & Engg. Y.C.C.E. Nagpur, India

Asst. Professor, Dept. of Computer Technology, Y.C.C.E., Nagpur, India

ABSTRACT: Online shopping becomes more important in our today's life because of varieties of products are available online. More number of people are using it purchasing online products. As number of online products, their sales and comments are increasing day by day, it is not possible for potential consumer to review all comments and take decision based on them. Buyers' comments are playing important role in taking decision regarding purchasing of products. Therefore in this paper a feedback analysis system is developed for analyze the feedback comments and tell buyers which organization is best for purchasing. For this opinion mining combining with natural language processing techniques are used.

KEYWORDS: Sentiment Analysis, Data mining, opinion mining, information filtering

I. INTRODUCTION

Sentiment analysis is most important part of product review. Sentiment analysis is study in which it analyzes the emotions, beliefs of people towards product, organization. In this lots of things are covered like view mining, emotion mining, influence mining, feeling analysis, product analysis. Opinion is key factor of people behaviors. If there is need to take some decision people want to know the opinion of others. If others opinion are positive towards it then it influence to other people. So many E-Commerce organization providing facility to buyer that comments on purchasing product. So it will help other buyer collect the opinion or aspects towards the product. Even voter can need other people opinion about particular political parties before giving the vote to it. Collecting people opinion is more important for many business, communities, advertising purpose.

Product reviews are a more important contain of information when it comes to making online purchases decision. Some reviews are more important or helpful than others. The goal of this project is to predict the usefulness of a product review. Trust is the most important thing when we want to take any Purchase decision for a particular product. In today's days there are many people which can openly comment on product. They share their experience and thoughts about the product with E-Commerce site.

Trustworthy of organization is important for accomplishment of e-commerce system. People can tell us which points they like about the product as well as dislike of product. So it helps other people to make better decision about which product to purchase. Buyer can leave the comments to seller and rate entire transaction negative, positive, neutral. Normally there exist hundreds and thousands of comments about a product and this number is constantly increasing rapidly as much as people contributing online.

For that problem is how people can read all this comment and getting information about organization. In order to solve this problem we are design feedback comment analysis system for deciding trustworthy organization. Provide correct information that will allow buyers to classify between trustworthy and non-trustworthy Organization. Encourage sellers to be trustworthy.

There are 3 aspects in comment analysis:

1. Positive aspect
2. Negative aspect



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

3. Neutral aspect

II. RELATED WORK

Trust and reputation systems represent a significant trend in decision support for Internet Mediated service provision. The basic idea in [1] is to let parties rate each other, for example after the transaction completion they use the average ratings about a given party to derive a trust or reputation score, which can assist other parties in deciding whether to transact or not with that party in the future.

Though web crawler reviews are easily accessible, it is difficult for a potential consumer to read all the reviews for a particular kind of product. Therefore in paper [2] they have used Weka classifier.

In paper [3] author observed that negative comments affect new product sales more than positive comments. Different from the proposition of the diffusion model, e- word of mouth (WOM) has a great effect on new product sales early on and such effect decreases over time.

In this paper [4] claim e-commerce domain can provide all the right ingredients for successful data mining. They describe an integrated architecture for supporting integration. The architecture can dramatically reduce processing of data, cleaning, and data understanding effort documented to take 80% of the time in knowledge discovery projects. They define the data transformation connection required from the customer event streams and transaction processing systems to the data warehouse.

In paper [5] author presents a technique of ontology-based sentiment classification for classify and analyze online product reviews of consumers. They approach based on a lexical variable ontology. Once the testing is done, it could be demonstrated that the proposed method can provide more efficiency for sentiment classification based on text content.

III. PROPOSED METHODOLOGY

The main objective of proposed work is to develop system which help buyer to decide which product is best and which organization is trustworthy for this we are required to analysis the feedback comment of the organization. This comments are lager in size so buyer cannot read all comments which are place on E-Commerce organization. So it is difficult to take decision about the product and E-Commerce organization. By using this project buyer can easy find out which organization is suitable for purchasing and getting more knowledge about the E-Commerce organization.

There are following phases:-

1. Data Collection
2. Pre-processing
3. Part Of Speech Tagging
4. Opinion feature word Identification
5. Use of SentiWordNet:
6. Classification of Comments
7. Fuzzy Rules.
8. Trust Score Calculation

1. Data Collection:

First phase of this system is collecting the relevant comment dataset from E-Commerce website Like Amaze, e-Bay, Snap deal. Collected data contain lots of information about the products which contain product id, product name, price of product, comments related to product etc .Collected data is not in proper format so it require pre-processing phase for better understanding.

2. Stop word Removal:

Stop word are frequently used word in any language. Stop words are eliminated in most of application because this words frequently use in any application and eliminated this word will focused on important words. For example in search engine query is like "Who is the Prime minister in India." This type of query search engine searches for the words like "who", "Prime", "minister", "India" then it will retrieve more pages containing the words "India", "minister". So it is not related document which we exactly want because any document contain this stop word. By deleting or removing this stop word we can get the proper data to analyse. So this pre-process data can be used for next

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

processing and enhancement. Stop word removal process can be used for improving and standardizes the processing speed of the data and for better result.

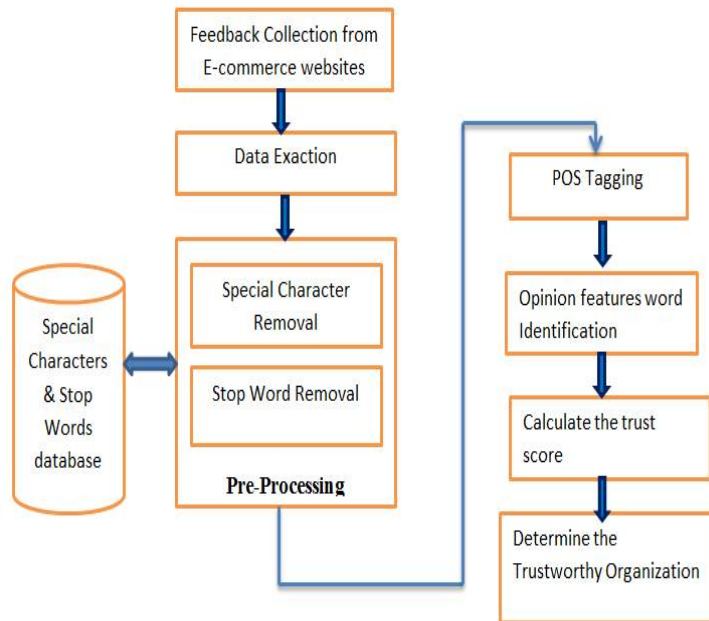


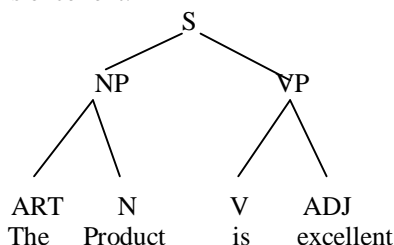
Figure.1 Broad Overview of Proposed System

3. Part of Speech Tagging:

POS Tagging stands for Part-Of-Speech tagging. A Part-Of-Speech Tagging (POS Tagger) annotates the part of speech of the word with particular word according to document in which it is use. For example verb(V), adjective (Adj), noun(N),verb phrase(VP),noun phrase(NP), Preposition(PREP), Conjunction(CONJ) etc. In mostly computational applications use more fine-grained POS tags like 'noun-plural'. Part of speech tagging is necessary for identification of adjective, adverb, noun, conjunction. When POS Tagging is done there may be some disambiguation for some words Like some words can be used as noun as well as adjective so it is difficult to identify the actually sense of that word. POS tagging is mainly use in text mining, opinion mining.

For example:

The product is excellent.



4. Opinion feature word Identification:

In opinion feature word identification phase it will identifying the features from collected comments Like “vacuum cleaner which I purchase yesterday is easy to use and feature is also good and compatible” extract the opinion feature “Quality” .It also specify the opinion is negative, positive, neutral. For example in given comment the word good determine the positive aspect towards the product quality. Other comment like “Laptop keyboard functional is very good but price is very high” express the feature towards “Price”. The feature identification is main done on adjective,



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

adverb. If the adjective is finding in near to the product feature in comment then it is use as opinion word for that noun. Opinion feature identification is very important in product analysis.

5. Use of SentiWordNet:

After opinion word identification it will tokenize the each comment into number of tokens. Then it will assign the hash value to each word by sequentially then it will check each word with hash value in SentiWordnet for calculation of positive score and negative score.

6. Classification of Comments:

In this phase comments can classify into three aspect positive, negative, neutral comments. By using previous opinion feature word identification step. It will extract the opinion word then like adjective, adverb which present in comment. Then check for available words of positive, negative for classification. If retrieved words match with the positive list of words then that comment is classify as positive comment. If retrieved word match with negative word then that comment is classify as negative comment otherwise neutral comment. This classification is necessary for this project.

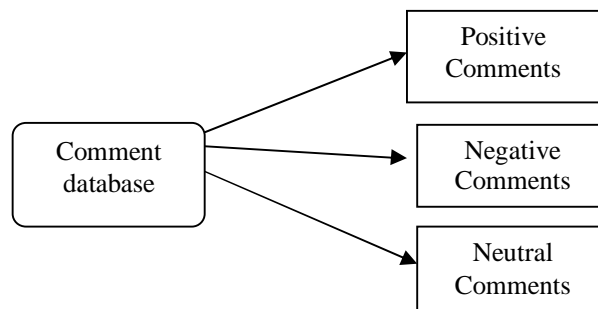


Figure.2 Classification of Comments

7. Fuzzy rules:

In some exceptional cases where it is difficult to predict that comment is positive or negative. Words like not, never, no indicate negative aspect if it comes along with adjective like “not good” then it will change the meaning of entire sentence. Semantic meaning of opinion specified “good” is different from “very good” and it is completely different from “not good”. In second and third cases, there exist modifies term “very” and “not”.

In this case fuzzy rules are used to calculated the fuzzy score .

Case 1:

If (value (ADJ)) \geq 0.5
Then F.S = value (ADJ) $^{1/2}$
Else
F.S = value (ADJ) 2

In this if the value of adjective in SentiWordNet is greater than 0.5 then it will calculate the fuzzy score (F.S) according to equation.

Case 2:

Again not, never, etc. will change the orientation of the opinion like {not good will be bad} {never accepted will be always rejected} etc.

In Case 2, we consider the weight (W) of the opinion like (not/never) (Adj /Verb)

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Fuzzy Score = 1 – Value Of (Adj / Verb)

8. Trust Score Calculation :

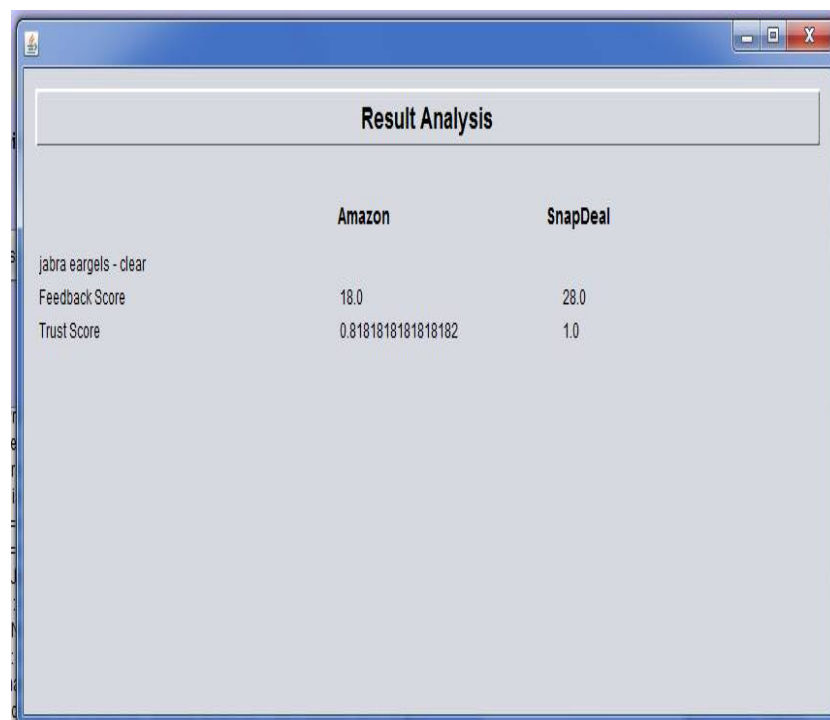
Last step of this analysis is trust score calculation of each E-Commerce organization according to positive score and negative score.

Trust score = TPS/TPS+TNS

TPS stand for total positive score and TNS stands for total negative score.

IV. EXPERIMENTAL RESULT

In this section result analysis is done. Firstly collecting the comment database from E-Commerce website. Then applying the preprocessing technique of collected comment database. Pre-processing technique includes special character removal, stop word removal technique. After that Part of Speech Tagging is done on pre-process data. After that using SentiWordNet and fuzzy rules fuzzy score is calculated. For each E-Commerce organization trust score is calculated according to positive feedback score and negative feedback score .



	Amazon	SnapDeal
jabra eargels - clear		
Feedback Score	18.0	28.0
Trust Score	0.8181818181818182	1.0

Figure.3 Result Analysis

In this figure 3.shows the result analysis for each E-Commerce organization.In that one product is taken i.e Jabra Eargels-Clear which can sale by both the E-Commerce Organization. Calculation of feedback score and Trust score is displayed. Feedback Score is calculated according to the positive comments and negative comments of both E-Commerece organization.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

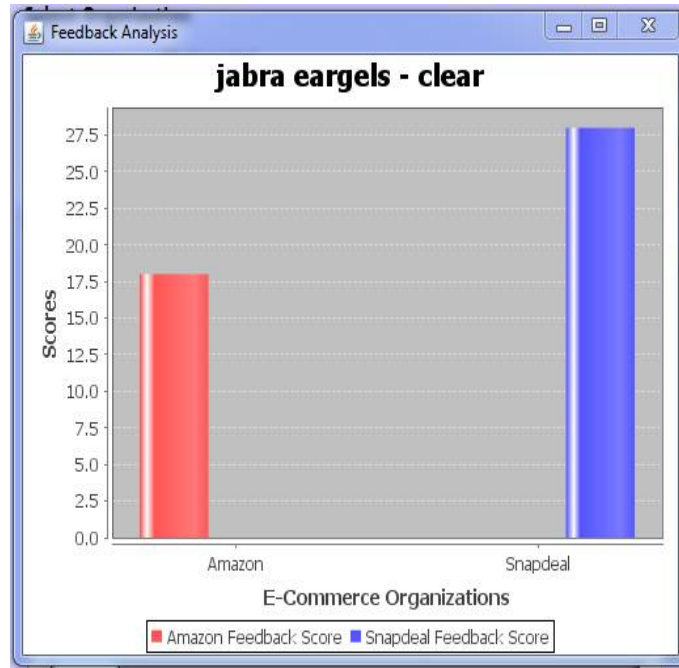


Figure.4 Feedback Score of E-Commerce Organization

In this Figure 4.shows Feedback score for each E-Commerce organization.Feedback score is nothing but total positive feedback score of each organization.It will useful for the further trust score calculation purpose.

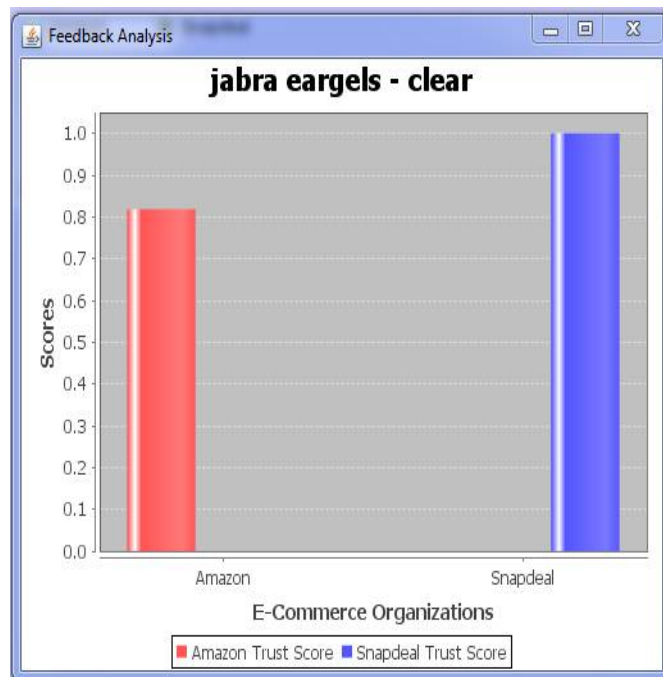


Figure.5 Trust Score of E-Commerce Organization



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

In this Figure 4.shows Trust score for each E-Commerce organization.

Trust score is calculated using total positive score/total positive score + total negative score formula.

V. CONCLUSION AND FUTURE SCOPE

Our approach demonstrates the novel application of combining natural language processing with opinion mining for trust evaluation for e-commerce applications. It helps us to know which organization is more trustworthy than other for buying the product. Result analysis of this analysis may change according to database used. It can be used in many application Like movie review, hotel review, tourism review, product review. In future other natural language technique can be used for more enhancements in result. In future we can apply intra classification on product database for more accurate result.

REFERENCES

- [1] Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *DSS*, vol. 43, 2007
- [2] Arun Monicka Raja M., Godfrey Winster S., Swamynathan S., "Review Analyzer: Analyzing Consumer Product Reviews from Review Collections," *IEEE International Conference on Recent Advances in Computing and Software Systems*, 2012.
- [3] Geng Cui, Hun-Kwong Lui, Xiaoning Guo, "Online Reviews as a Driver of New Product Sales," *IEEE International Conference on Management of e-Commerce and e-Government*, 2010.
- [4] Suhail Ansari, Ron Kohavi, Liew Mason, "Integrating E-commerce and Data Mining," *IEEE International Conference on Data Mining*, 2011
- [5] Jantima Pulpiniji, Aditya K. Ghose, "An Ontology Based Sentiment Classification Methodology for Online Consumer Reviews," *IEEE International Conference on Web Intelligence and Intelligent Agent Technology*, 2008.
- [6] H. Wang, Y. Lu, and C. Zhai, "Latent aspect rating analysis on review text data: A rating regression approach," in *Proc. 16th ACM SIGKDD Int. Conf. KDD*, New York, NY, USA, 2010.
- [7] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Comput. Linguist.*, vol. 37, 2011.
- [8] H. Wang, Y. Lu, and C. Zhai, "Latent aspect rating analysis without aspect keyword supervision," in *Proc. 17th ACM SIGKDD Int. Conf. KDD*, San Diego, CA, USA, 2011, pp. 618–626.
- [9] B. Pang and L. Lee "Opinion mining and sentiment analysis," *Found. Trends Inf. Ret.*, vol. 2, no. 1–2, pp. 1–135, Jan. 2008.
- [10] Y. Hijikata, H. Ohno, Y. Kusumura, and S. Nishida, "Social summarization of text feedback for online auctions and interactive presentation of the summary," *Knowl. Based Syst.*, vol. 20, no. 6, pp. 527–541, 2007
- [11] J. O'Donovan, B. Smyth, V. Evrim, and D. McLeod, "Extracting and visualizing trust relationships from online auction feedback comments," in *Proc. IJCAI*, San Francisco, CA, USA, 2007
- [12] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. 4th Int. Conf. KDD*, Washington, DC, USA, 2004, pp. 168–177.
- [13] X. Wang, L. Liu, and J. Su, "RLM: A general model for trust representation and aggregation," *IEEE Trans. Serv. Comput.*, vol. 5, no. 1, pp. 131–143, Jan.–Mar. 2012.
- [14] C. Lin and Y. He, "Joint sentiment topic model for sentiment analysis" in *Proc. 18th ACM CIKM*, Hong Kong, China, 2009, pp. 375–384.
- [15] I. Titov and R. McDonald, "Modelling online reviews with multigrain topic models," in *Proc. 17th Int. Conf. WWW*, Beijing, China, 2008, pp. 111–120.
- [16] L. Zhuang, F. Jing, X. Zhu, and L. Zhang, "Movie review mining and summarization," in *Proc. 15th ACM CIKM*, Arlington, VA, USA, 2006
- [17] M. De Marneffe and C. Manning, "The Stanford typed dependencies representation," in *Proc. Cross Parser*, Stroudsburg, PA, USA, 2008.
- [18] C. Lin and Y. He, "Joint sentiment/topic model for sentiment analysis," in *Proc. 18th ACM CIKM*, Hong Kong, China, 2009, pp. 375–384
- [19] I. Titov and R. McDonald, "Modeling online reviews with multigrain topic models," in *Proc. 17th Int. Conf. WWW*, Beijing, China, 2008, pp. 111–120.
- [20] S. Brody and N. Elhadad, "An unsupervised aspect-sentiment model for online reviews," in *Proc. HLT*, Los Angeles, CA, USA, 2010, pp. 804–812.
- [21] P. D. Turney, "Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews," in *Proc. 40th ACL*, Philadelphia, PA, USA, 2002, pp. 417–424.
- [22] A. Mukherjee and B. Liu, "Aspect extraction through semisupervised modeling," in *Proc. 50th ACL*, vol. 1. Stroudsburg, PA, USA, 2012, pp. 339–348.