



A Study on Aspect Level Sentiment Analysis of Named Entities using Annotation Scheme

S.Devidhanshrii¹, Dr. T.Karthikeyan²

Research Scholar, Department of Computer Science, PSG College of Arts & Science, Coimbatore, India¹
Associate Professor, Department of Computer Science, PSG College of Arts & Science, Coimbatore, India²

ABSTRACT: Sentiment Analysis is a prolongation of data mining that extracts and inlays data automatically. Mainly sentiment analysis is used for the study of public opinion analysis and text subjectivity analysis and it is a process of identifying and categorizing opinions expressed in a piece of text. People shared their positive and negative thoughts according to their sentiments. Sentiment analysis can be considered as a classification process at different levels. Primarily this calculation is computed based on the famous annotation schema. It can be generated after pre-processing.

Keywords: Sentiment analysis, Classification, Sentiment analysis process, Annotation schema

I. INTRODUCTION

Sentiment analysis is the study about Analyzing people's opinions, evaluations, sentiments, and appraisals, emotions including temperament over against subsistence acting as services, products, organizations, issues, topics, events, individuals and their attributes have a divergence that can be expressed with the opinion of the people in specific a time about any objects and its features. Sentiments also vary in strengths, which can be positive, negative, neutral or mixed sentiments is to find out other people thinking with the growing availability resources. The origin and what's more, quick development of opinion investigation concurs with



the huge volume of digital data available on social media on Web such as online personal blogs, forum discussions, Twitter, micro blogs and review sites Opinions are vital to all human exercises and are key components of our practices. According to Liu B., Zhang L. (2012) [1], Sentiment analysis are technically challenging but practically it is very useful. Organizations to develop their business they wanted to know the opinion about the quality of their product and services from the consumers and public. As per WalaaMedhat, AhmedHassan, HodaKorashy [2], "Supposition Analysis (SA) or Opinion Mining (OM) is the computational investigation of peoples sentiments, states of mind and feelings toward a substance."

Pang and Lee [8] and Liu [9] presented detailed surveys on the functions and challenges in sentimental analysis. They also mentioned the methods and techniques used to clear up problems in sentimental analysis. Cambria and Schuller et al. [9], Feldman [10] and Montoyo and Martínez-Barco [11] have presented surveys illustrating the current trends in sentimental analysis. Tsytarau and Palpanas [12] have made a survey on main topics of sentimental analysis and presented a detailed description about it. For each and every topic they have represented its definition, problems and development categorized the articles with the aid of tables and graphs.

II. SENTIMENT ANALYSIS PROCESS

Sentiment analysis identifies the primary views of a text. Sentiment analysis is the computational study about opinion of the people, attitudes, emotions and appraisals towards individuals, entities, issues, topics, events and their attributes. Here entity represents an event, task or individuals. In this the authors investigated and presented a brief survey on various sentimental analysis application and proposed algorithms and also illustrated the recent trend in opinion mining and sentimental analysis.

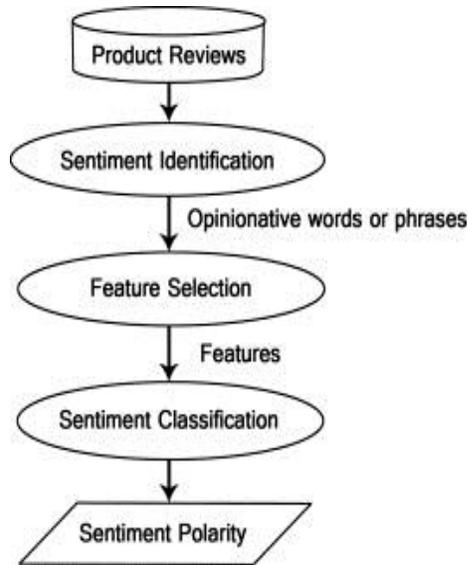


Fig. 1 Sentiment analysis Process

The entire process is then tested using the Figure Eight Simulator to show the inferences drawn from experimental studies. The simulation uses the customer feedback on services offered by Indian Airlines and the twitter data of the customers are taken as the input data set and using the tweets the polarities are identified to derive sentiments. Along with the polarity computation a confidence or trust score is generated to estimate the confidence level of extracted aspects about a comment and arrive at conclusions based on the higher confidence score.

III. PROPOSED WORK

This section demonstrates the method used to mine aspects and opinion from customer reviews and named entities using annotation schema. Initially, the dataset is passed through the pre-processing phases. Then Natural Language Processing (NLP) tools are used for sentence splitting and Part Of Speech (POS) tagging. Next, the annotation scheme is and aspects in the forms of nouns and phrases are extracted along with the corresponding opinion. Finally, the method is tested and the results are generated and analyzed.



ANNOTATION SCHEME

Annotations are a kind of metadata that makes references to the data itself but is not an actual part of it. Annotations can include notes on metadata, such as comments on the data. When performing Sentiment Analysis of stored data, defining an annotation scheme works fine to find out the fine granular points of the customer review expressions and the feedback and expectations of the customer to improve on the entity. Annotation scheme can be specified for events that highlight positive and negative polarity and also represents the attitude of the person towards their agents and objects. This annotation scheme is helpful to find out implicit expressions of opinions as opinion and sentiment tends to focus on explicit expressions of opinions.

On the final word there are 4 types of components that are represented in the annotation scheme they are:

pp-positive polarity

np-negative polarity

agent-initiator of action

influencer-either retain or reverse the polarity

object-represents the actor

IV. EXPERIMENTAL RESULTS

The problem taken into consideration for this research work is to mine the huge set of named entities or product review datasets and to identify the polarity and subjectivity text to identify the customer perception and preference using sentiment analysis. This study uses aspect or entity level sentiment analysis to arrive at customer satisfaction, preference, opinion and sentiments.

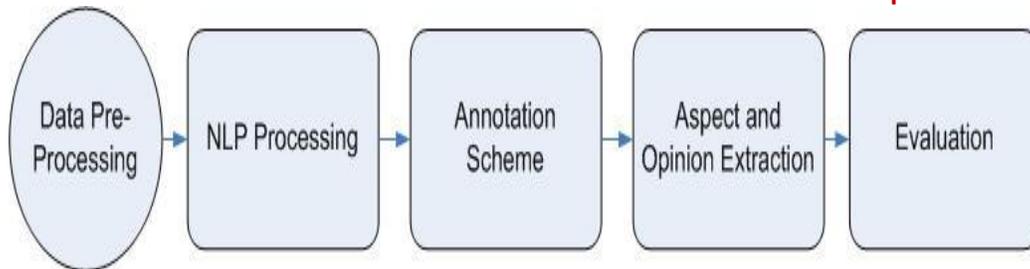


Fig. 2 Workflow of Aspect Level Sentiment Analysis using Annotation Scheme

V. IMPLEMENTATION

1. Data Collection

To perform the experimental analysis the customer feedback on the performance of Indian Airlines is collected from the tweet database of customer feedback. Given below is the simulation of the tweets regarding Indian Airlines from which inferences are drawn regarding the service of the airline company. The simulator used for finding the aspect level sentiment analysis of customer reviews is simulated using the figure eight simulator.

2. Simulation of Sentiments

The first step in the simulation process is to give the customer review or feedback from the customer review database as input. This file can be in the .csv, .tsv or .xls format. The data from this customer review database is arranged into rows and columns and when dragged and dropped the database is uploaded for further processing. Figure 4.3 below shows the database updation process in which the database of customer reviews towards performance of Indian airlines are uploaded for further processing.

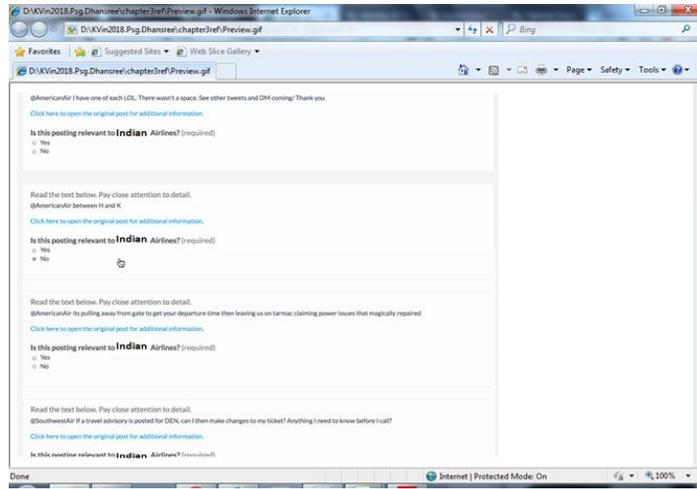


Fig. 3 Preview of the Survey Questionnaire

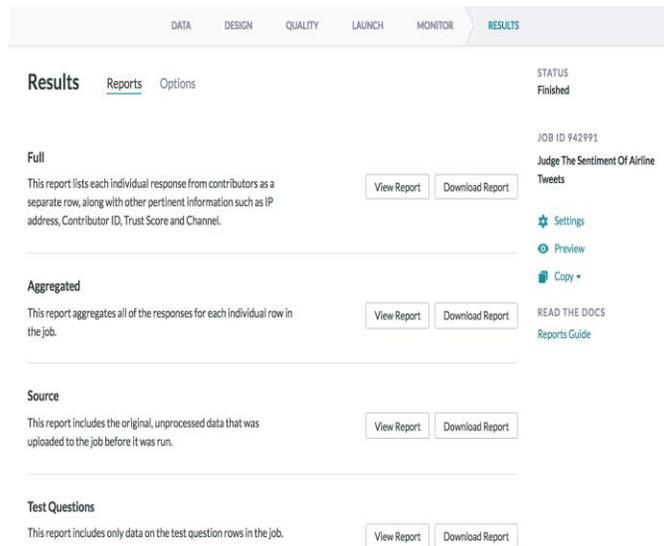


Fig. 4 Result View Page

VI. CONCLUSION

This research work on expected value emotionalism investigation introduces a new model for aspect-level sentiment analysis and a annotation schema to identify sentiments in customer reviews of named entities. This model is largely based on the annotation schema in



which the customer review database is taken as input and then after pre-processing undergoes the process of annotating the statement given by the customer about a specific item or administration and finally inferences in structure of sentiments are drawn from the customer feedback statements.

Instead of modeling overall sentiment explicitly, approximation is done using aspect sentiments and topic distributions. When an annotation is prepared careful attention is paid to identify the agent, the object and the influencer to recognize the assumptions polarity of the sentiment and then based on the polarity to make an analysis to identify the sentiments included in the statement.

To study the execution of the proposed framework is an experimental study was carried out in detail to perform an analysis of sentiment hidden in the customer review statement. Figure Eight is the tool used for identifying the sentiment at the aspect level. The experiment was carried out with the dataset downloaded from TripAdvisor. The TripAdvisor data set contains reviews on airline services. Based on the experimental study it is discovered that the urged work has achieved a high level of performance using the annotation scheme.

VII. FUTURE ENHANCEMENT

The major work of the thesis is to recognize slants hidden in customer reviews and feedbacks and the system could achieve that and the same is highlighted using the experimental study. In future this system can be extended further to include the concept of granularity to further deepen the analysis process. A better option to achieve this functionality combined with the proposed model is to use Multi-Grain LDA (MG-LDA).

This approach performs topic modelling at two levels: high level topics such as the Laptop, MP3 Player, Space Heater topics in the Amazon data set, and low-level aspects such as battery life, image quality, ease of use, etc. With topics and aspects structured in such a way the model would be able to detect the rateable aspects for each topic more successfully and automatically. Another improvement would be to incorporate the overall sentiment into the model itself to gear up the overall sentiment detection up to par.



REFERENCES

- [1] A. Collomb, C. Costea, D. Joyeux, O. Hasan, and L. Brunie. A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation. Technical Report RR-LIRIS-2014-002, LIRIS UMR 5205 CNRS/INSA de Lyon/Universite Claude Bernard Lyon 1/Universite Lumiere Lyon 2/Ecole Centrale de Lyon, Mar 2014.
- [2] M. Daiyan, S. K. Tiwari, M. Kumar, and M. Aftab Alam. A Literature Review on Opinion Mining and Sentiment Analysis. *International Journal of Emerging Technology and Advanced Engineering*, 5(4):262-280, 2015.
- [3] A. N. Jebaseeli and E. Kirubakaran. A Survey on Sentiment Analysis of (Product) Reviews. *International Journal of Computer Applications*, 47(11):36-39, Jun 2012.
- [4] B. Liu. *Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies*. Morgan & Claypool Publishers, 2012.
- [5] W. Medhat, A. Hassan, and H. Korashy. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4):1093-1113, 2014.
- [6] Z. Nanli, Z. Ping, L. Weiguo, and C. Meng. Sentiment analysis: A literature review. In *Proceedings of the 2012 IEEE International International Symposium on Management of Technology (ISMOT'2012)*, pages 572-576, held in Hangzhou, Zhejiang, China, Nov 2012.
- [7] B. Pang and L. Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1-135, 2008.
- [8] R Feldman. Sentiment Analysis Tutorial. In *TF4 Sentiment Mining from User Generated Content*, 23rd International Joint Conference on Artificial Intelligence, held in Beijing, China, Aug 2013.
- [9] K. Schouten and F. Frasincar, "Survey on Aspect-Level Sentiment Analysis," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 813-830, 1 March 2016. doi: 10.1109/TKDE.2015.2485209



- [10] Y. Dang, Y. Zhang and H. Chen, "A Lexicon-Enhanced Method for Sentiment Classification: An Experiment on Online Product Reviews," in *IEEE Intelligent Systems*, vol. 25, no. 4, pp. 46-53, July-Aug. 2010.
- [11] Jonathon Read, "Using Emoticons to reduce Dependency in Machine Learning Techniques for Sentiment Classification", Published in Proceedings ACLstudent '05 Proceedings of the ACL Student Research Workshop Pages 43-48, June 27, 2005.
- [12] Cane W. K. Leung and Stephen C. F. Chan, The Hong Kong Polytechnic University, Hong Kong SAR
- [13] Pang B, Lee L. Opinion mining and sentiment analysis. *Found Trends Inform Retrieval* 2008;2:1–135.
- [14] Cambria E, Schuller B, Xia Y, Havasi C. New avenues in opinion mining and sentiment analysis. *IEEE Intell Syst* 2013;28:15–21.
- [15] Feldman R. Techniques and applications for sentiment analysis. *Commun ACM* 2013;56:82–9.
- [16] Montoyo Andre´ s, Martı´nez-Barco Patricio, Balahur Alexandra. Subjectivity and sentiment analysis: an overview of the current state of the area and envisaged developments. *Decis Support Syst* 2012;53:675–9.
- [17] Tsytsarau Mikalai, Palpanas Themis. Survey on mining subjective data on the web. *Data Min Knowl Discov* 2012;24:478–514.
- [18] Bo Pang and Lillian Lee, Shivakumar Vaithyanathan, "Thumbs up? Sentiment classification using Machine Learning Techniques", Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Philadelphia, July 2002, pp. 79-86. Association for Computational Linguistics.
- [19] Kiplagat Wilfred Kiprono, Elisha Odira Abade. "Comparative Twitter Sentiment Analysis Based on Linear and Probabilistic Models". *International Journal on Data Science and Technology*. Vol. 2, No. 4, 2016, pp. 41-45. doi: 10.11648/j.ijdst.20160204.11



S.Devidhanshri *et al*, International Journal of Computer Science and Mobile Applications,
Vol.7 Issue. 1, January- 2019, pg. 16-25

ISSN: 2321-8363

Impact Factor: 5.515

- [20] Dong Chul Park, “Probabilistic Classifier with Supervised Learning”, International Journal of Computer Science and Electronics Engineering (IJCSEE) Volume 3, Issue 4 (2015) ISSN 2320–4028 (Online)
- [21] Cooper, G. F. & Herskovits, E. (1992), ‘A Bayesian method for the induction of probabilistic networks from data’, Machine learning 9 (4), 309–347
- [22] Sylvester Olubolu Orimaye, “Sentiment Augmented Bayesian Network”, School of Information Technology, MONASH University Malaysia, Proceedings of the 11-th Australasian Data Mining Conference (AusDM'13), Canberra, Australia, sylvester.orimaye@monash.edu
- [23] Murphy K. (1998): A Brief Introduction to Graphical Models and Bayesian Networks
- [24] Nipun Mehra, Shashikant Khandelwal and Priyank Patel, Department of Computer Science, “Sentiment Identification Using Maximum Entropy Analysis of Movie Reviews”
- [25] Nguyen Viet Cuong, Nguyen Thi Thuy Linh, Ha Quang Thuy and Phan Xuan Hieu, “A Maximum Entropy Model for Text Classification“, College of Technology, Vietnam National University, Hanoi, Vietnam, Graduate School of Information Sciences, Tohoku University, Sendai, Japan {mrcuongnv, linhhtt, [@vnu.edu.vn](mailto:thuyhq), hieuxuan@ecei.tohoku.ac.jp}