

A SYSTEM BASED ON OPINION MINING OF USER DATA ON SOCIAL MEDIA

¹Prof. Pragati Goel,
Asso. Prof. M.C.A. Dept.
SIMS, Navi Mumbai,
goelpragati78@gmail.com

²Prof. Sagar Thakare
Asst. Prof. M.C.A. Dept.
SIMS, Navi Mumbai
saghakare@gmail.com

³Prof. Rahul Wantmure
Asst. Prof. M.C.A. Dept.
SIMS, Navi Mumbai
rahul-wan2003@yahoo.co.in

ABSTRACT

Over social media there are plenty of emotional symbols (emojis) which are used in comparison to textual content which is an unstructured sort of text which gets considered and a daily increase in such symbols is moving closer to the new records prediction determination approach. Due to the speedy development of net, big numbers of web pages assigned by readers' emotions are generated via new portals.

Comparing to the previous studies which focused on author's perspective, our research focuses on readers' emotions invoked by news articles. Our studies affords significant assistance in social media software such as sentiment retrieval, opinion summarization and election prediction. In this paper, we are expecting the reader's emotion state-of-information primarily based on the social opinion community. Mainly, we assemble the opinion community based totally on the semantic distance. The communities within the news network suggest specific occasions which might be related to the feelings. Consequently, the opinion community serves as the lexicon among occasions and corresponding emotions. We leverage neighbor courting in network to are expecting readers' feelings. As a result, our techniques achieve better result than the other techniques. Moreover, we developed a developing approach to prune the network for sensible application. The experiment verifies the rationality of the application.

In this paper, we propose a system based on social opinion prediction by generating a real-time social opinion network. In more details, first, we train word vectors according to the most recent Wikipedia word corpus. Second, we calculate semantic distance between news via word vectors.

Keywords: *Opinion mining, SOM, emojis, pattern matching*

1. INTRODUCTION

Social emotion prediction is of value to market analysis and to political decision. With the free and convenient communication environment of internet, people show increasing enthusiasm of online communication. Meanwhile, the internet users prefer to produce and convey online information through expressing personal opinions than just obtain online information. In this way, numerous news articles and comments have been published and shared rapidly via social media services. As a result, abundant underlying positive or negative emotion information spreads and reflects the social sentiment tendency. Most intuitively, emotional label has been widely used in social web services. Fig. 1 indicates the result of voting for a news article using emotion labels from a popular news portal. Large numbers of people concerned about a hot news online. Therefore, valuable and available emotional information is continuously provided for scientific research work [4]. Furthermore, comparing to the traditional methods, which need to do numbers of surveys offline, data processing technology has been developed more feasible in the field of emotional extraction, analysis and prediction with its benefits of lower cost, higher efficiency and more accuracy. Under this circumstance, readers' emotions prediction shows a highly research potential.

Compared with the typical tasks of sentiment analysis, opinion mining or affect recognition which based on subjective text, social opinion prediction focuses on objective text, for example news articles, which may not contain any opinion, but can evoke readers' certain emotion. Due to the particularity of the task, social opinion prediction has potential applications which are different from those of writer-sentiment analysis [5]. Considering the effect of social media on the public sentiment, social emotion analysis engenders large benefits to social and economic problem, such as political issues and brand perception.

In this paper, we implement social opinion prediction by generating real-time social opinion network. In more details, first, we train word vectors according to the most recent Wikipedia word corpus. Second, we calculate semantic distance between news via word vectors. As a metric between opinions, semantic distance allows us to construct the opinions growing network to describe the dynamical social opinions. Last, we predict follow-up news' social emotion based on the network.

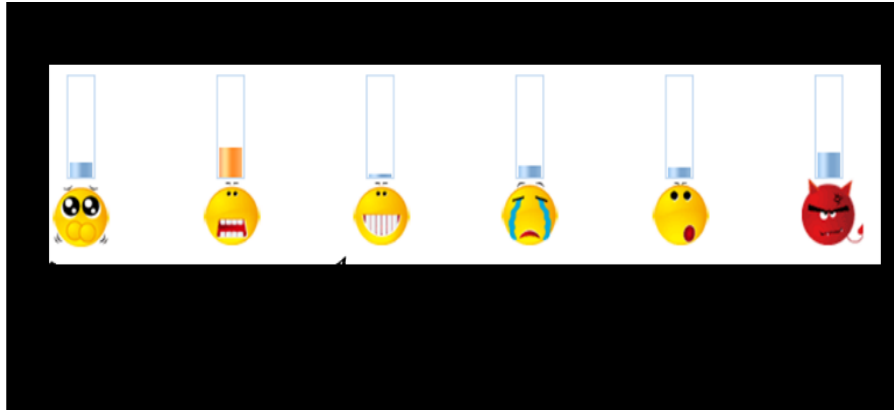


Fig.1: An example of emotion labels and user ratings

2. RELATED WORK

In present paper it is proposed that the system can do the prediction of emotions of the customers which is taken as the reference of the news article which helps us to recognize about the person's feelings regarding to such a piece of writing. In this experiment is proposed on datasets. Social opinion prediction is a difficult research endeavor. As the initial research work on social opinion prediction, "affective text" in SemEval-2007 Tasks [11], [13] intended to annotate news headlines for the evoked emotion of readers. Another research focuses on reader's emotion evoked by news sentences [15]. Current techniques of social opinion prediction may be divided into 3 classes: understanding-primarily based techniques, statistical strategies and hybrid procedures. Due of the deficiency of data text [13], [16] it is unmanageable to annotate the emotions consistently. The popular emotional lexicon includes Affective Lexicon, linguistic annotation scheme [18], Word Net-Affect [19], SentiWordNet [20], and Septic Net [21]. The drawback of knowledge-based techniques is the dependency on the labels present in the emotional lexicon. These techniques can't process terms that do not seem within the emotional lexicon. Statistical methods predict social opinion by training a statistical model based on a large number of well-labeled corpuses. There are two principal categories of statistical methods: word-level [11], [14] and topic-level [22] methods. Word-level methods focus on exploiting the sentiment of individual words on the idea that words are the foundation of consumer sentiments. With a purpose to model the word-emotion association, a variant of Naïve Bayes version named Emotion-term (ET) is created. The phrases extracted from the information articles are considered

as independent functions which indicate the emotion. But, word-stage capabilities in social opinion prediction are always interfered through the background noise words. In particular, the methods treat each word individually also many emotional words are usually mixed with background noise words.

3. THE PROPOSED SYSTEM

After studying the various techniques given in existing researches we are proposing a business intelligence analytic module which is based on emotion detection of user data done through opinion mining. User data is given in form of product remarks, reviews and comments. Our system is based on mining of remarks, reviews and comments given by users. This will help us to give the user instant and fast response which is very essential for proper business development. In this proposed system we can implement the opinion network and emotion opinion model on the datasets which are retrieved from the business data i.e. through E-commerce database. Opinion prediction system will help to predict the trend and decision making in business. In Fig.2. we are showing the flow graph of our proposed system. Here we are also considering the opinion generation done on textual data by using pattern matching to avoid a situation where no emojis are used in the given reviews, feedback and complaints. Following is the algorithm for the proposed system.

Algorithm: Algorithm for proposed architecture is as follows:

- Step 1: Dataset is taken from the e-commerce database or from the company's portal where users have posted their remarks, reviews and comments.
- Step 2: Giving all the emotional symbols used an emotional rating from the unstructured data.
- Step 3: Applying a separator so that textual information and emotional symbols are separated.
- Step 4: Apply a self-organizing map (SOM) which is a type of artificial neural network (ANN) on samples i.e. emotional symbols and get opinion R1. Similarly apply pattern matching on textual data and get opinion R2.

Step 5: Apply a unifier on R1 and R2 to get the best opinion or result.

Decision making can be done now based on the result which can help businesses to grow and take a proper action on a specific product.

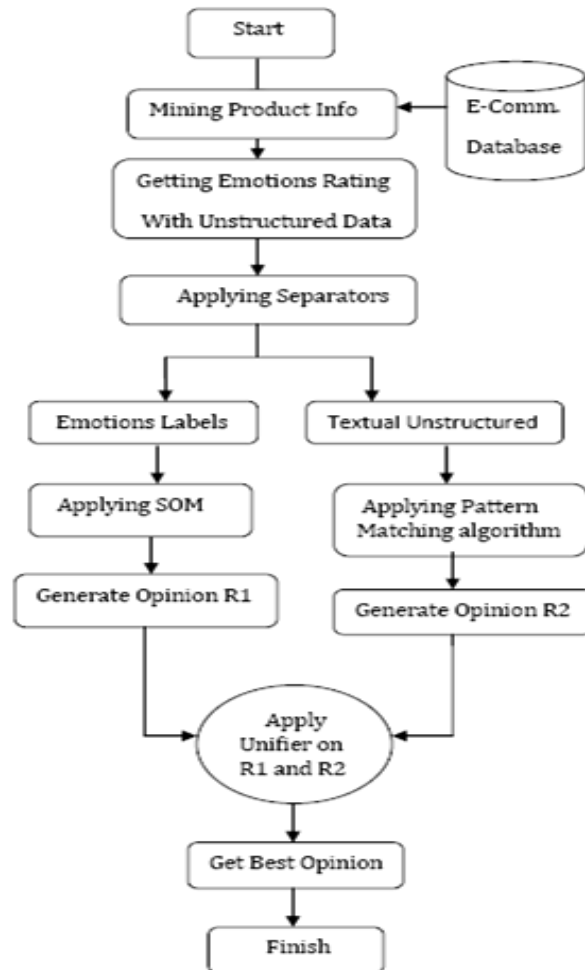


Fig -2: Flow chart of proposed architecture

4. CONCLUSIONS

The proposed system can do the prediction of products advertising and marketing for the e-commerce based businesses. This is an important manner of performing mining on the remarks, reviews and comment given by users and using them to improve the commercial enterprise productivity. The proposed work allows the e-trade platform to analyze the user data in terms of business driven methods.

REFERENCES

- [1] www.researchgate.net/publication/316259881_Predicting_Social_Emotions_from_Readers-27_Perspective, IEEE paper
- [2] E. Cambria, B. Schuller, Y. Xia, and B. White, "New avenues in knowledge bases for Natural language processing," *Knowledge- Based Syst.*, vol. 108, pp. 1–4, Sep. 2016.
- [3] E. Cambria, "Affective Computing and Sentiment Analysis," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, Mar. 2016
- [4] B. Zhang, X. Guan, M. J. Khan, and Y. Zhou, "A time-varying propagation model of hot Topic on BBS sites and Blog networks," *Inf. Sci. (Nay)*, vol. 187, pp. 15–32, 2012.
- [5] Q. Wang, O. Wu, W. Hu, J. Yang, and W. Li, "Ranking social emotions by learning Listwise preference," in 1st Asian Conference on Pattern Recognition, ACPR avenues in knowledge bases for Natural language processing," *Knowledge- Based Syst.*, vol. 108, pp. 1–4, Sep. 2016.
- [6] B. Zhang, X. Guan, M. J. Khan, and Y. Zhou, "A time-varying propagation model of hot Topic on BBS sites and Blog networks," *Inf. Sci. (Nay)*, vol. 187, pp. 15–32, 2012.
- [7] Q. Wang, O. Wu, W. Hu, J. Yang, and W. Li, "Ranking social emotions by learning Listwise preference," in 1st Asian Conference on Pattern Recognition, ACPR 2011, pp.164–168, 2011.
- [8] K. H.-Y. Lin, C. Yang, and H.-H. Chen, "What emotions do news articles trigger in their readers?," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '07*, pp. 733-734, 2007.
- [9] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of Communities in large networks," *J. Stat. Mech. Theory Exp.*, vol. 0008, no. 10, pp. 155-168, 2008.
- [10] https://en.wikipedia.org/wiki/Data_mining
- [11] P. Katz, M. Singleton, and R. Wicentowski, "SWAT-MP: The SemEval-2007 Systems for Task 5 and Task 14," in *Proceedings of the 4th International Workshop on Semantic Evaluations*, pp. 308- -313, 2007
- [12] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *EMNLP*, vol. 21, no. 9, pp. 1746–1751, 2014.
- [13] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv1301.3781*, pp. 1–12, 2013.
- [14] C. Strapparava and R. Mihalcea, "Semeval-2007 task 14: Affective text," *Proc. of SemEval-2007*, no. June, pp. 70–74, 2007.
- [15] P. K. Bhowmick, "Reader Perspective Emotion Analysis in Text through Ensemble based Multi-Label Classification Framework," in *computational intelligence and security*, vol. 2, no. 4, pp. 64-74, 2009
- [16] C. Quan and F. Ren, "An Exploration of Features for Recognizing Word Emotion," in *international conference on computational linguistics*, pp. 922-930, 2010.
- [17] A. Ortony, G. L. Clore, and A. Collins, "*The Cognitive Structure of Emotions*," *The Quarterly Review of Biology*, vol. 18, no. 6. 1988.
- [18] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language Resources and Evaluation*, vol. 39, no. 2–3. pp. 165–210, 2005.
- [19] C. Strapparava and A. Valitutti, "WordNet-Affect: an affective extension of WordNet," *Proc. 4th Int. Conf. Lang. Resour. Eval.*, pp. 1083–1086, 2004.

- [20] A. Esuli and F. Sebastiani, "SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining," *Proc. 5th Conf. Lang. Resour. Eval*, pp. 417–422, 2006.
- [21] E. Cambria, D. Olsher, and D. Rajagopal, "SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis," *Twenty-eighth AAAI Conf.*, pp. 1515–1521, 2014.
- [22] Y. Rao, "Contextual Sentiment Topic Model for Adaptive Social Emotion Classification," *IEEE Intell. Syst.*, vol. 31, no. 1, pp. 41–47, Jan. 2016.