Bibliotheque de Humanisme et Renaissance | ISSN : 0006-1999
Volume 84, Issue 1, 2024

PUBLIC SENTIMENT INTERPRETATION ON TWITTER - A SURVEY

R.Urega[1], Dr.M.Devapriya[2]
[1]M.Phil Scholar, [2]Assistant professor,
PG & Research Department of Computer Science, Government Arts College,Coimbatore-18.

ABSTRACT

Opinion mining and sentiment analysis is the significant keyfield of Natural Language Processing. Number of users shared their thoughts and opinions on micro blog services. Social networks serve as the source of valuable platform for tracking and analyzing public sentiment of different people about different domains/products.Twitter is one of the most popular social sites where the millions of users share their opinions. Public sentiment analysis is a process to identify positive and negative opinions, emotions and evaluations in text. This paper reviewed and analysed number of techniques for public sentiments analysis and its classification.

I.INTRODUCTION

Web mining is the use of data mining techniques to discover automatically and extract information from Web documents and services. Sentiment analysis and opinion mining are sub fields of machine learning. They are very important in the current development because, lots of user opinionated texts are obtainable in the web now. This is a hard problem to be solved because natural language is highly unstructured in nature. The interpretation of the meaning of a particular sentence by a machine is exasperating. But the usefulness of the sentiment analysis is increasing day by day. Machines must be made reliable and efficient in its ability to interpret and understand human emotions and feelings.With the enormousgrowth of user generated messages, Twitter has become a social site where themillions of users can interacttheir opinions. Sentiment analysis on Twitter data has provided an effectiveway to expresspublic opinion timely which is critical for decision making in various domains. For example, when the public sentiment changes on a particular product, forinstance acompany may need to know the feedback of their product.In social media, any politician can adjust their position with respect to the sentiment change of the public. For every important decision making, it is necessary to mine public opinions and to find reasons behind variation of sentiments is valuable.It is usually difficult to find the accurate causes of sentiment variations since they may involve complicated internal and externalfactors.The emergingtopics discussed in the variation period could be highly related to the true reasons behind the variations. This paper includes with following sections. In section 2 the different sentiment analysismethods were discussed. Theevent detection and tracking aretracking the events, and detecting sentiment analysis about events were described in section 3. DataVisualizationand it importance were analysed in section 4. In section 5 the correlation between tweets and events are described.Thispaper endedwithconclusioninsection6.

II RELATED WORKS

2.1 SENTIMENT ANALYSIS

Sentiment analysis isthe process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the attitude of writer towards a particular topic, product, etc. is positive, negative, or neutral. Sentiment classifications have some emotional states such as 'angry','sad', and 'happy'. Sentiment analysis has become popular in judging the opinion of consumers towards various brands [5]. The way in which consumers express their opinion on social networking websites and tweets helps to judge this opinion [13]. O'Connor et al.[12] studied on analysing sentiments of public share on Twitter. This is the main work in micro blogging services to interpret the variations in sentiment.

Pang et al. [1] focused on the existing methods on analysis of sentiments in details i.e. supervised machine learning methods. Machine learning methods minimize the structural risks. To predict sentiment of documents,supervised machine learning approach is used.

M.Hu and B. Liu [2] also works on mining and summarizing customer reviews based on data mining and natural language processing methods. It mines features of products on which customers have been mentioned. Identifies opinions and decides whether it is positive or negative in each review. Itis also summarizing the results of opinions.

W.zhang et al. [3] conducted exhaustive study of opinions recovery from blogs. In this paper, they have presented a three-component opinion retrieval algorithm. The first component is an information retrieval module. The second one classifies documents into opinionative and no opinionative documents, and keeps the former. The third component ensures that the opinions are related to the query, and ranks the documents in certain order.

Meng et al.[3]collected opinions in Twitter for entities by mining hash-tags to conclude the presence of entities and sentiments from each tweets.

Li.Jiang et.al [4] evaluated the target-dependent Twitter sentiment Classification. The state-of-the-art technique solves the problem of target dependent classification which assumes the target-independent strategy. It always assigns irrelevant sentiments to give target.For a given query, it classifies the sentiments of the tweets according to the categories whether they contain positive, or negative or any neutral sentiments about the query. Here the query considers as the target of the sentiments.

III. EVENT DETECTION AND TRACKING

Events are the good explanations behind the distinctions of sentiments related to target. By tracking the events, detecting sentiment analysis about events, tracking variations in sentiments and finding reasons for changes the sentiments completes this task.
Leskovec et al. [5] focused specific works on tracking,for example quoted phrases and sentences. This work offers some of analysis of the global news cycle and the dynamics of information propagation between conventional and social media. It finds a short distinctive phrase that travels rather intact through online text as it develops over time.
B-S Lee and J.Weng [6] focus on detection of events through analysis of the contents which are published in Twitter. This paper proposedEDCoW which refers Event Detection with Clustering of Wavelet based on Signals for detecting events.

IV.DATA VISUALIZATION

D.Tao et.al [7] deeply studied subspace learning algorithms and ranking. Retrieving of images from large databases is very active research field today. For retrieving images, content based image retrieval (CBIR) technique is used. It is related by semantically to query of user from collected database of images.SVM classifier always unstable for a smaller size training set. SVMRF becomes poor if there are number of samples of positive feedback are small.SVM has also suffered from problem of over fitting.

V. CORRELATION BETWEEN TWEETS AND EVENTS

T.Sakaki et al.[8] established novel models to map tweets in a public segmentation. They detect real-time events in Twitter such as earthquakes. They also suggested an algorithm for monitoring tweets detect event. Each Twitter user is considering as a sensor.Kalman filtering and particle filtering are used for estimation of location.

Y.Hu et.al [9] proposed a mutual statistical model ET-LDA that characterizes topical effects among an event and its related Twitter feeds. This model supports the topic modelling of the event and the segmentation of the each event or tweet.

Chakrabarti and Punera [10] have described a variant of Hidden Markov Models in event summarization from tweets. It gets an intermediate illustration for a sequence of tweets relevant for an event. In this paper, author used the sophisticated techniques to summarize the relevant tweets are used for some highly structured and recurring events. Hidden Markov Models gives the hidden events.

Shulong Tan et.al [11] proposes LDA model for interpreting public sentiment variations on Twitter. Where Latent Dirichlet Allocation (LDA) propose two models 1.Foreground and Background LDA (FB-LDA) and 2.Reason Candidate and Background LDA (RCB-LDA) in which FB-LDA model can remove background topics and then extract foreground topics to show possible reasons and Reason Candidate and Background LDA(RCB-LDA) model provides ranking them with respect to their popularity within the variation period. RCB-LDA model also finding the correlation between tweets and their events.

TABLE.1: **Summary of the topics considered in each approach.**

S.NO	TITLE	AUTHOR	YEAR	METHOD	ADVANTAGES	DISADVANTAGES
1.	Opinion mining and sentiment analysis	B.Pang And L.Lee	2008	Supervised machine learning methods.	Machine Learning methods minimizethe structural risks.	It cannot analyse possible reasons behind public sentiments.
2.	Mining and summarizing customer reviews	M.Hu and B. Liu	2004	Natural language processing methods.	It predicts movie sales and elections so it's easy to make decisions.	It cannot determine the opinions strength.
3	Opinion retrieval from blogs	W. Zhang, C. Yu, and W. Meng	2007	Retrieval algorithm.	It mines hash tags for opinions.	It cannot handle more general writings and crossing domains.
4.	Target-dependent twitter sentiment classification	L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao	2011	state-of-the-art	This technique gives the high performance for target-dependent twitter sentiment classification.	It is not search out the relations between a target and its extended targets.
5	Meme-tracking and the dynamics of the news cycle	J.Leskove, L.Backstrom,and J.Kleinberg	2009	Tracking mems	It provides temporal relationships such as the possible.	Fine grained events can be detected very hardly.
6	Event detection in twitter	B-S.Lee and J.Weng	2011	EDCoW	EDCoW (Event Detection with Clustering of Wavelet-based Signals) signal independently treats each word.	It cannot contribute relationship among users to event detection.

7	Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval	D.Tao et.al X. Tang, X. Li, and X. Wu	2006	SVM	SVM improves the performance of relevance feedback	SVM cannot use tested tuning method and not select the parameters of kernel based algorithms.
8	Earthquake shakes twitter users: Real-time event detection by social sensors.	T.Sakaki M.Okazaki and Y. Matsuo	2010	Sensor. Kalman filtering and particle filtering	Earthquakes are detected by this system and sends email to users who are registered users.	It cannot detect multiple event occurrences.
9	Et-lda: Joint topic modeling for aligning events andtheir twitter feedback	Y. Hu, A. John, F. Wang, and D. D. Seligmann	2012	ET-LDA	It extracts a variety of dimensions such as sentiment and polarity.	They model each tweet as a multinomial mixture of all events, which is obviously unreasonable due to short lengths of tweets.
10	Event summarization using tweets	Chakrabarti and Punera	2011	Hidden Markov	It provides benefits for existing query matching technologies.	SIT does not use the continuous time stamps present in tweets.

5. CONCLUSION

The paper reviews and summarizesthe methodology for analysing public sentiments.This survey presents various approaches to Opinion Mining and Sentiment analysis. It provides a detailed view of different applications and potential challenges of Sentiment Classification .The emerging topics are related to the actual or genuine reasons behind the variations are very important. So it is necessary to interpret sentiment variation and finding the reasons behind them to overcome above mentioned limitations on different models.The Latent Dirichlet Allocation (LDA) techniques which proposed two models such as Foreground and Background LDA (FB-LDA) and Reason Candidate and Background LDA (RCB-LDA) interpret sentiment variations in an efficient manner.It concludes that these models can mine number of reasons behind variations of publicsentiments.

REFERENCES:

[1] B. Pang and L. Lee, ―"Opinion mining and sentiment analysis",‖Found. Trends Inform. Retrieval, vol. 2, no.(1–2), pp. 1–135, 2008.
[2] M. Hu and B. Liu, ―"Mining and summarizing customer reviews",‖inProc. 10th ACM SIGKDD, Washington, DC, USA, 2004.

[3] W. Zhang, C. Yu, and W. Meng, ―"Opinion retrieval from blogs",‖inProc. 16th ACM CIKM, Lisbon, Portugal, 2007.
[4] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, ―"Target-dependent twitter sentiment classification",‖ in Proc. 49th HLT, Portland, OR,USA, 2011.
[5] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in Proc. 15th ACM SIGKDD, Paris, France, 2009
[6] J. Weng and B.-S. Lee, ―Event detection in twitter,‖ in Proc. 5thInt. AAAI Conf. Weblogs Social Media, Barcelona, Spain, 2011
[7] D. Tao, X. Tang, X. Li, and X. Wu, ―Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval,‖IEEE Trans. Patt. Anal. Mach. Intell.,vol. 28, no. 7, pp. 1088–1099, Jul. 2006 .
[8] T. Sakaki, M. Okazaki, and Y. Matsuo, ―Earthquake shakes twitter users: Real-time event detection by social sensors,‖ in Proc. 19thInt. Conf. WWW, Raleigh, NC, USA, 2010.
9]Y. Hu, A. John, F. Wang, and D. D. Seligmann, ―Et-lda: Joint topic modeling for aligning events andtheir twitter feedback,‖ in Proc.26th AAAI Conf. Artif. Intell., Vancouver, BC, Canada, 2012.
10] D. Chakrabarti and K. Punera, ―Event summarization using tweets,‖ in Proc. 5th Int. AAAI Conf. Weblogs Social Media, Barcelona, Spain, 2011.
[11] Shulong Tan, Yang Li, Huan Sun, Ziyu Guan, Xifeng Yan, ―Interpreting the Public Sentiment Variations on Twitter,‖ IEEE Transactions on Knowledge and Data Engineering, VOL. 26, NO.5, MAY 2014.
12)B. O'Connor, R. Balasubramanyan, B. R. Routledge, and N. A.Smith, "From tweets to polls: Linking text sentiment to public opinion time series," in Proc. 4th Int. AAAI Conf. Weblogs Social Media, Washington, DC, USA, 2010.
[13] J. Bollen, H. Mao, and A. Pepe, "Modeling public mood and emotion:Twitter sentiment and socio-economic phenomena," in Proc.5th Int. AAAI Conf. Weblogs Social Media, Barcelona, Spain, 2011.
[14] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment in twitter events," J. Amer. Soc. Inform. Sci. Technol., vol. 62, no. 2,pp. 406–418, 2011.