

Sentiment Analysis as Prediction Tool in Multiple Domain Applications

Baldev Singh

Lyallpur Khalsa College, Jalandhar (India)

ABSTRACT

Sentiment Analysis is considered as a tool for prediction. The use of sentiment analysis is in multiple domain applications. The list of domain specific areas is exhausted where sentiment analysis is applied. Some of the real life domain specific areas of sentiment analysis are as services sector, healthcare, sports, fashion, politics, business sector, financial services, movies reviews, stock market reviews, book reviews, and social relations. The sentiments analysis and opining mining is made by various researchers from time to time. The techniques suggested by various authors are mainly classification, clustering, machine learning, natural language processing techniques, supervised pattern discovery method, quantification, fuzzy neural network, correlation, probabilistic model, regression, probabilistic model, supervised learning. Although various techniques are presented by various authors still there is vast scope of research for deploying best sentiment analysis mechanisms specifically on the basis of its explicit domain of applications.

Keywords: *sensitive analysis, opinion mining, social network, tweets, blog.*

I. INTRODUCTION

Sentiments refers to the human behavior of an individual. The outcome generated from the analysis or mining of the sentiments of a group or cluster can be used for decision making and may be helpful for an individual or an enterprise to take effective decisions. Lots of depend upon the predictability power of the analysis. Sentiment concept is also related to attitude, opinion [7], evaluation or emotions [2, 18]. In the contemporary world, social media and Internet is a widely used source to collect data and information related to opinions and sentiments of the humans of a required group or groups. Social networks, blogs, tweets, sites, discussion forums etc. are the generators of huge data related to the sentiments and is available in digital form. From early 2000, natural language processing is in use for sentiment analysis by way of data mining [15,17], text mining and web mining. Sentiment analysis [1] is a thriving subject now and is not in demand by individuals but by other govt. and not govt. commercial and other enterprises. Analysis of sentiments is classified as qualitative and quantitative in nature. Sentiment analysis includes the entities like issues, events, products, individuals, organizations, services, capital and many more. Sentiment analysis is coincide with opinion mining or sentiment mining and sentiment analysis [3,14,18] is more commonly used term in industrial domain rather than academia as opinion mining is more preferred term in use by academia. Although natural language processing has long standing in research area of different domains but now it is widely used for sentiment analysis. One of the reason is huge volume of opinion and sentiments related data available in digital form online.

II. DOMAINS AND LEVELS OF SENTIMENT ANALYSIS

Sentiment analysis or opinion mining leads to decision making regarding some purchases, marketing, campaign or services and the like. Individuals both consumers and producers as well as organizations need information regarding sentiments of public or consumers in the real world. There is explosive growth in Internet of Everything and social media that generates enormous content online and such abundance information is further analyzed for the multiple domain of applications as a prediction tool. Although large information related to sentiments and opinions is available but it is not so easy to decipher from large blogs or postings etc. Secondly it is also many times very difficult to identify the most relevant sites to extract and summarize the sentiments for analysis. No doubt there exists specific posts over social media for opinion mining. Real life domain specific areas [18] of sentiment analysis is very wide like

services, healthcare, sports, fashion, politics, business sector, financial services, movies reviews, stock market reviews, book reviews, social relations, B2B, B2C, C2G, C2B, C2C, B2G (B-Business, C-Consumer, G-Govt.) are some of them.

Sentiment analysis is performed at different levels of sources of information. These levels are categorized as document level, sentence level and entity as well as feature level. Sentiment analysis at document level mainly consists of sentiments or opinions as reviews which are either positive or reviews/sentiments. The sentiment is in the form of whole document about a particular aspect, product, service etc. The sentence level sentiment analysis further consists of negative, positive or neutral opinion or sentiments [11,14]. This analysis is basically subjective in nature and can have clauses in it. The third level of sentiment analysis categorize as entity and aspect level analysis which is known as finer-grained analysis [16,18]. Aspect level works at opinion level and the construct used for input is immaterial here whether it is at sentences level, document level, paragraph level or clause level.

Sentiment is either positive or negative and the opinion generated from the sentiment analysis is known as target. The sentence “Although laptop I have (of XY company) is heavy but it is having worth”. Here opinion generated as target and is emphasized on entity and its aspects. Opinion [7] generated using sentiment analysis can be categorize as regular opinion or comparative opinion. For example, College ABC is reputed college and academically at top”. There is regular opinion about positive sentiment. The other sentence “The college ABC is good in academic and sports as compare to college XYZ” There is comparative opinion.

III. SENTIMENT-PREDICTION TECHNIQUES

Sentiment Analysis is characterized as a tool for predictive mining used in multiple domain applications. Sentiments analysis and opining mining is made by various researchers from time to time. Some of techniques proposed and presented by various authors are depicted in table-1 below. The techniques suggested basically includes, classification [4,6,11,12], clustering [2,3,5], machine learning, natural language processing techniques [1], supervised pattern discovery method [2,13], quantification, fuzzy neural network [8], correlation [9], probabilistic model [10], regression, probabilistic model, supervised learning. The usage of different techniques of sentiment analysis is also categorically different on the basis of its domain of application areas.

Table-1: Sentiment-Prediction Techniques

Sr.No	Authors	Research Paper Title	Publication in Journal/Book/Conference	Techniques
1	B. Pang, L. Lee, and S. Vaithyanathan	Thumbs Up? Sentiment Classification Using Machine Learning Techniques	Proc. ACL-02 Conf. Empirical Methods in Natural Language Processing (EMNLP), 2002.	NLP, Classification Using Machine Learning
2	Bing Liu, Minqing Hu, Junsheng Cheng	Opinion Observer: Analyzing and Comparing Opinions on the Web	WWW 2005, May 10-14, 2005	Supervised pattern discovery method
3	M. Farhadloo and E. Rolland	Multi-Class Sentiment Analysis with Clustering and Score Representation	13th International Conference on Data Mining Workshops, 2013.	Clustering
4	P.D. Turney	Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised	Proc. 40th Ann. Meeting on Assoc. for Computational Linguistics (ACL), pp. 417-424, 2001	Classification

		Classification of Reviews,		
5	S. Siamala Devi, Dr. A. Shanmugam	Hybridization of K-means and Harmony Search Method for Text Clustering Using Concept Factorization	International Journal of Advanced Research in Computer Engineering & Technology, Vol. 3 Issue 8, 2014.	Clustering
6	Xia Hu, Lei Tang, Jiliang Tang, Huan Liu	Exploiting Social Relations for Sentiment Analysis in Microblogging	WSDM '13, February 4–8, 2013.	Classification
7	Wei Gao, Fabrizio Sebastiani	Tweet Sentiment: From Classification to Quantification	2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2015	Quantification
8	Johan Bollen, Huina Mao, Xiao-Jun Zeng	Twitter mood predicts the stock market.	Journal of Computational Science 2(1),October,2010.	Fuzzy Neural Network
9	Matt Mcvicar, Tim Freeman	Mining the Correlation between Lyrical and Audio Features and the Emergence of Mood	Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR, 2011.	Correlation
10	Yang Liu, Xiangji Huang, Aijun An, Xiaohui Yu	ARSA: A Sentiment-Aware Model for Predicting Sales Performance Using Blogs	SIGIR'07, July 23–27, 2007.	Probabilistic Model, Regression
11	Casey Whitelaw, Navendu Garg, Shlomo Argamon	Using Appraisal Groups for Sentiment Analysis	CIKM '05 Proceedings of the 14th ACM international conference on Information and knowledge management.	Classification
12	Yun Wan, Qigang Gao	An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis	2015 IEEE International Conference on Data Mining Workshop (ICDMW)	Classification
13	L B Shyamasundar, P Jhansi Rani	Twitter sentiment analysis with different feature extractors and dimensionality reduction using supervised learning algorithms	India Conference (INDICON) 2016 IEEE Annual	Supervised learning
14	M.Govindarajan , Romina M	A Survey of Classification Methods and Applications for Sentiment Analysis	The International Journal Of Engineering And Science, Vol. 2, Issue-12, 2013	Classification

15	M. Karanasou, A. Ampla, C. Doulkeridis and M. Halkid	Scalable and Real-Time Sentiment Analysis of Twitter Data	016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), Barcelona, 2016,	Supervised learning
16	Y. Haimovitch, K. Crammer, S. Mannor	More is better: Large scale partially-supervised sentiment classification	Proc. of ACML, 2012.	Classification
17	Singh and Vivek Kumar	A clustering and opinion mining approach to socio-political analysis of the blogosphere	Computational Intelligence and Computing Research (ICCIC), 2010 IEEE	Clustering and opinion mining

IV. CONCLUSION

Sentiment is a valuable metric on the basis of which some decision can be made by the individuals or an organization. Decision or strategy can only be successful if accurate and appropriate analysis is made. Along with any metric of sentiment analysis to be used, context also plays a vital role. The sensitive analysis also be selected on the basis of context that is domain of application where prediction or mining operations is required. Multi model of sensitive analysis can be advantageous in some area of application as well as single model or technique can give best outcomes in some other area. Because of huge subjective information produced by various social media tools like Facebook and Twitter, as well as different variety of opinions, views, sentiments, feelings etc. effective sensitive analysis is must. The effectiveness of the technique is vital irrespective of any single or multi model techniques so that proper prediction can be realized. The ability related to velocity, subjectivity, influence, trajectory, accuracy, relevance, cross-reference intensity of every “mention” at social media sensitive analysis count a lot for prediction in multiple domain applications.

REFERENCES

- [1] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs Up? Sentiment Classification Using Machine Learning Techniques”, Proc. ACL-02 Conf. Empirical Methods in Natural Language Processing (EMNLP), 2002.
- [2] Bing Liu, Minqing Hu, Junsheng Cheng, “Opinion Observer: Analyzing and Comparing Opinions on the Web”, WWW 2005, May 10-14, 2005.
- [3] M. Farhadloo and E. Rolland, “Multi-Class Sentiment Analysis with Clustering and Score Representation”, 13th International Conference on Data Mining Workshops, 2013.
- [4] P.D. Turney, “Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews”, Proc. 40th Ann. Meeting on Assoc. for Computational Linguistics (ACL), pp. 417-424, 2001.
- [5] S. Siamala Devi, Dr. A. Shanmugam, “Hybridization of K-means and Harmony Search Method for Text Clustering Using Concept Factorization”, International Journal of Advanced Research in Computer Engineering & Technology, Vol. 3 Issue 8, 2014.
- [6] Xia Hu, Lei Tang, Jiliang Tang, Huan Liu, “Exploiting Social Relations for Sentiment Analysis in Microblogging”, WSDM '13, February 4–8, 2013.
- [7] Wei Gao, Fabrizio Sebastiani, “Tweet Sentiment: From Classification to Quantification”, 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.
- [8] Johan Bollen, Huina Mao, Xiao-Jun Zeng, “Twitter mood predicts the stock market”, Journal of Computational Science 2(1), October, 2010.

- [9] Matt Mcvicar, Tim Freeman, "Mining the Correlation between Lyrical and Audio Features and the Emergence of Mood, Proceedings of the 12th International Society for Music Information Retrieval Conference, ISMIR 2011.
- [10] Yang Liu, Xiangji Huang, Aijun An, Xiaohui Yu, "ARSA: A Sentiment-Aware Model for Predicting Sales Performance Using Blogs", SIGIR'07, July 23–27, 2007.
- [11] Casey Whitelaw, Navendu Garg, Shlomo Argamon, "Using Appraisal Groups for Sentiment Analysis", CIKM '05 Proceedings of the 14th ACM international conference on Information and knowledge management.
- [12] Yun Wan, Qigang Gao. "An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis", 2015 IEEE International Conference on Data Mining Workshop (ICDMW).
- [13] L B Shyamasundar, P Jhansi Rani, "Twitter sentiment analysis with different feature extractors and dimensionality reduction using supervised learning algorithms", India Conference (INDICON) 2016 IEEE Annual.
- [14] Eka Surya Usop, R. Rizal Isnanto, Retno Kusumaningrum, "Part of speech features for sentiment classification based on Latent Dirichlet Allocation Information Technology Computer and Electrical Engineering (ICITACEE) 2017 4th International Conference.
- [15] M. Karanasou, A. Ampla, C. Doulkeridis and M. Halkidi, "Scalable and Real-Time Sentiment Analysis of Twitter Data," 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), Barcelona, 2016.
- [16] Y. Haimovitch, K. Crammer, S. Mannor, "More is better: Large scale partially-supervised sentiment classification", Proc. of ACML, 2012.
- [17] Singh and Vivek Kumar, "A clustering and opinion mining approach to socio-political analysis of the blogosphere", Computational Intelligence and Computing Research (ICCIC), 2010 IEEE.
- [18] Bing Liu, "Sentiment Analysis and Opinion Mining", Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.