Opinion Mining on Twitter Data using Unsupervised Learning Technique

Abstract

Social media is one of the biggest forums to express opinions. Sentiment analysis is the procedure by which information is extracted from the opinions, appraisal and emotions of people in regards to entities, events and their attributes. Sentiment analysis is also known as opinion mining. Opinion mining is to analyze and cluster the user generated data like reviews, blogs, comments, articles etc. These data find its way on social networking sites like twitter, facebook etc. Twitter has provided a very gigantic space for prediction of consumer brands, movie reviews, democratic electoral events, stock market, and popularity of celebrities.

The main objective of opinion mining is to cluster the tweets into positive and negative clusters. An earlier work is based on supervised machine learning (Naïve bayes, maximum entropy classification and support vector machines). The proposed work is able to collect information from social networking sites like Twitter and the same is used for sentiment analysis. The processed meaningful tweets are cluster into two different clusters positive and negative using unsupervised machine learning technique such as spectral clustering. Manual analysis of such
large number of tweets is impossible. So the automated approach of unsupervised learning as spectral clustering is used. The results are also visualized using scatter plot graph and hierarchical graph.

References

1. Influence factor based opinion mining of Twitter data using supervised learning Malhar Anjaria; Ram Mohana Reddy Guddeti 2014 Sixth International Conference on Communication Systems and Networks (COMSNETS) Year: 2014
5. Twitter as a Corpus for Sentiment Analysis and Opinion Mining, Alexander Pak, Patrick Paroubek. (2014)
6. Parikh and Movassate , Sentiment Analysis of User-Generated Twitter Updates using Various Classification Techniques , Stanford University, 2009
9. Sajib Dasgupta and Vincent Ng “Mine the Easy, Classify the Hard: A Semi-Supervised Approach to Automatic Sentiment Classification”, Human Language Technology Research Institute, University of Texas at Dallas.
10. M Ashraf et. al. "Multinomial Naïve Bayes for Text Categorization Revisited", University of Waikato
17. Agglomerative Hierarchical Clustering Algorithm- A Review K.Sasirekha, P.Baby
Department of CS, Dr.SNS.Rajalakshmi College of Arts & Science, International Journal of
Scientific and Research Publications, Volume 3, Issue 3, March 2013 1 ISSN 2250-3153
(2011) Performance Analysis of Hierarchical Clustering Algorithm K.Ranjini Department of
Computer Science and Engineering, Einstein College of Engineering, Tirunelveli, Ind
19. A Survey on Supervised Learning for Word Sense Disambiguation Abhishek Fulmari1 ,
Manoj B. Chandak2 International Journal of Advanced Research in Computer and
20. Ahamed Shafeeq BM and Hareesha K S , “Dynamic Clustering of Data with Modified
K-Means Algorithm,” proceeding of the 2012, International Conference on Information and
Computer Networks (ICICN 2012).
21. Sentiment Analysis and Opinion Mining: A Survey, Volume 2, Issue 6, June 2012 ISSN:
2277 128X International Journal of Advanced Research in Computer Science and Software
Engineering.
22. A Survey Paper on Twitter Opinion Mining Geetanjali S. Potdar1 , Prof R. N. Phursule2
International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index
23. Eigenvalues and Eigenvectors: Formal, Symbolic and Embodied Thinking Michael O. J.
Thomas The University of Auckland.
24. Spectral Clustering: Advanced Clustering Techniques 1 S. V. Suryanarayana (Ph.D),
2Guttula Rama Krishna (M.Tech), 3Dr. G. Venkateswara Rao (Ph.D) International Journal of
Advanced Research in Computer Science and Software Engg. 4(11), November - 2014, pp.
625-62.
25. Feature selection and classification approach for sentiment analysis gautami tripathi1
and naganna s.2 machine learning and applications: an international journal (mlaij) vol.2, no.2,
June 2015
2, 2013 Comparison of Supervised and Unsupervised Learning Algorithms for Pattern
Classification R. Sathya Professor.

Index Terms

Computer Science
Information Sciences

Keywords
Opinion mining; feature extraction; feature vector; spectral clustering; k-means clustering; hierarchical clustering.