IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-5, Issue-12 E-ISSN: 2347-2693

# Future Opportunities and Challenges in Sentiment Analysis: An Overview

**Pranay Kumar B.V[1*] and M.Sadanandam[2]**

[1*]Dept of CSE, Christu Jyothi Institute of Technology and Science, JNTUH, Hyderabad, India
[2] Dept of CSE, UCE Kakatiya University, Warangal, India

*[*]Corresponding Author: pranaybv4u@gmail.com, Tel.: +91-99853-31180*

**Available online at: www.ijcseonline.org**

Received: 19/Nov/2017, Revised: 30/Nov/2017, Accepted: 20/Dec/2017, Published: 31/Dec/2017

***Abstract*—** Today the evolution of technology and fair access to the internet in countries such as India, public opinions over social media, expression of sentiment on products and services are fast and furious in present days. These opinions have value for companies to materialize profits and understand the market for their future strategic decisions. Present technology adoption energized by the healthy growth in big data framework, caused applications based on Sentiment Analysis (SA) in big data to become common for businesses. But there is wide gap and scope for SA application in big data. This paper discusses various Sentiment analysis approaches and algorithms, including sentiment polarity detection, SA features (explicit and implicit), sentiment classification techniques, applications of SA. Future opportunities and challenges of sentiment analysis is explored. Scalable, automated, accurate, sophisticated sentiment analysis is a much sought-after technology that almost no one has truly nailed yet.

***Keywords*—** Sentiment Analysis Approaches, tools, techniques, machine Learning, opportunities, and challenges, supervised learning and unsupervised learning

## I. Introduction

Sentiment Analysis is typically used to analyze people's sentiments, opinions, appraisals, attitudes, evaluations and emotions towards such entities as organizations, products, services, individuals, topics, issues, events and their attributes, as presented online via text, video and other means of communication. These communications can fall into three broad categories, namely positive, neutral and negative. These categories involve many names and slightly different tasks, such as opinion mining, opinion extraction, sentiment mining, subjectivity analysis, customer complaint, effect analysis, emotion analysis, review mining and review analysis. These techniques are categorized into the following: Application-oriented, which ranges from stock price predictions to public voice analysis, crowd surveillance and SA-based customer care; fundamental approaches, including word-level sentiment disambiguation, sentence-level SA, aspect-level SA, concept-level SA, multilingual SA and linguistic features analysis; and social intelligence, which exploits the public's online content generation to analyze such inputs as pandemic spreading, emotion and responses towards local events

This review paper is organised with different approaches of sentiment analysis with discussion on features of SA in section II. Various classification techniques that aid in SA is discussed in section III. Various applications of SA and tools used in market are discussed in section IV and section V respectively. Section VI deals with the various challenges and opportunities in SA that may explored for future research.

## II. General Approaches of SA

Sentiment Polarity Detection SA, also known as opinion mining, is the extraction of positive or negative opinions from (unstructured) text (Pang et al., 2002). Sentiment classification has several important characteristics, including various tasks, features, and techniques. Three important sentiment polarity tasks are as follows:
• Identifying whether text is objective/subjective or whether subjective text has a positive/negative orientation
• Determining the level of the classification (document/sentence level)
• Identifying the source/target of the sentiment

Polarity classification of sentiment is classified into document-level, sentence-level and phrase level classification. Document-level classification classifies the document as positive, negative, or neutral. Sentence-level classification considers and classifies only a sentence determining whether a sentence is subjective or objective. To capture multiple sentiments that might exist within a single sentence, phrase-level classification is performed (Wilson et al., 2005). Furthermore, to categorize levels and sentiment classes, different assumptions have also been made about sentiment sources and targets.

**2.1 SA Features:** Four types of explicit features have been used, namely syntactic, semantic, link-based and stylistic features. Most common set of features for SA, Syntactic attributes contain word n-grams tags and punctuation. Moreover, these attributes contain phrase patterns, which make use of POS tag n-gram patterns. They illustrated that phrase patterns like 'n+aj' (noun followed by a positive adjective) usually denote positive sentiment orientation, whereas 'n+dj' (noun followed by a negative adjective) often expresses a negative sentiment. In 2004, Wiebe applied collections, where certain parts of fixed n-grams were exchanged with general word tags. Whitelaw et al. (2005) applied a set of modifier features (e.g., very, mostly and not). The presence of these features transformed appraisal attributes for lexicon items. Link/citation analysis is applied in link-based features to detect sentiment from the web and documents. Efron et al. (2004) demonstrated that opinion web pages are linked to one another. Link-based features have been used in limited studies. Thus, the effectiveness of such features for SA remains unclear. Stylistic features contain structural and lexical attributes, which are used in many previous stylometric/authorship works.

**2.2 Lexical and structural style markers:** This has been used in limited sentiment analysis studies. Bernabé-Moreno et al. applied hapax legomena (unique/once occurring words) for subjectivity and opinion perception. They found that the presence of unique words in the subjective text is higher than in an objective document. Lexical features, such as length of sentence, for the classification of feedback surveys, is used. Lexical style markers (words per message and words per sentence) were used in Cambria et al. to analyze web blogs. Previous studies have shown style markers to be highly common in web discourse.

**2.3 Implicit Features**: Studies on implicit features in SA have focused on semantic and linguistic rules to identify the embedded message, which is not typically expressed using predefined keywords. Instead, the meaning is delivered using similar conceptual-based expressions. Semantic features try to identify polarity or provide intensity-related scores to words and phrases. In Semantic Orientation (SO) method mutual information was calculated to compute the SO score of each word/phrase automatically. Manual or semi-automatically produced sentiment lexicons commonly use a primary set of automatically generated terms that are manually filtered and coded with polarity and intensity information. User-defined tags are used to indicate whether certain phrases have positive or negative sentiment. Semi-automatic lexicon generation tools were used by to construct a set of strong subjectivity, weak subjectivity and objective nouns. They also used other features, such as bag-of-words, to classify English documents as either subjective or objective.

**2.4 Appraisal group**: Initial term lists are created using WordNet. These lists are then filtered manually to construct the lexicon. In Appraisal Theory, each expression is manually classified into several appraisal classes, such as attitude, the polarity of phrases, orientation and graduation. Manually generated lexicons have also been used for affect analysis. Affect lexicons are used with fuzzy semantic typing to analyze movie reviews and news articles. Hate and violence in extremist web forums were analyzed using manually constructed affect lexicons

**2.5 Semantic attributes:** These contain contextual features that represent the semantic orientation of surrounding text. Semantic attributes have been useful for sentence-level sentiment classification.

**2.6 WordNet:** It is a large electronic lexical database for English and it continues to be developed and maintained. WordNet consists of synsets from major syntactic categories, such as nouns, verbs, adjectives and adverbs. The current version of WordNet (3.0) contains over 117,000 synsets, comprising over 81,000 noun synsets, 3,600 verb synsets, 19,000 adjective synsets and 3,600 adverb synsets (Poli et al., 2010). WordNet has been used for synonym collection, whereas SentiWordNet has been used to identify the semantic orientation of each sentence or extracted feature.

**2.7 SentiWordNet:** It is a lexical resource for opinion mining. It is a lexicon base that is similar to WordNet, but it is extended with the lexical information about the sentiment of each synset contained in WordNet. Three different polarities, namely positivity, negativity and objectivity, are assigned to each synset in WordNet. The two most common versions of SentiWordNet used in many studies are SentiWordNet 1.0 and SentiWordNet 3.0. Apart from being used in monolingual studies, SentiWordNet can also be used in multilingual SA

**2.7 SenticNet:** This is built by using semantic computing. It is the latest semantic resource specifically developed for concept-level SA. It exploits both Artificial Intelligence (AI) and semantic web technique to recognize, interpret and process natural language opinions better over the web. SenticNet is a knowledge base that can be applied in the development of many fields, such as big social data analysis, human-computer interaction, electronic health etc.

**2.8 Linguistic Rules:** Most of the rule-based linguistics approaches are applied to clause-level or concept-level sentiment classification. The algorithm adopts a pure linguistic approach and considers the grammatical dependency structure of the clause by using SA rules. Linguistic rules are useful for dealing with the semantic orientation of context-dependent words and they are very helpful for extracting implicit features. These features are those that are not clearly mentioned but are rather implied in a sentence. All existing works on implicit aspect extraction were based on the use of Implicit Aspect Clue (IAC) and rule-based method to extract implicit aspects. They mapped the implicit aspect to the corresponding explicit aspect.

## III. SENTIMENT CLASSIFICATION TECHNIQUES

**3.1 Sentiment Classification through Machine Learning**: The Machine Learning (ML) approach applies the ML algorithm and uses linguistic features with the aim of optimizing the performance of the system using example data. Typically, two sets of documents are required in an ML-based classification. These documents are the training and testing sets. A training set is used by the classifier to learn the document characteristics, whereas a testing set is used to validate classifier performance. The text classification methods using the ML approach can be divided into supervised and unsupervised learning methods. The supervised methods use a large number of labelled training documents. The unsupervised methods are used when these labelled training documents are difficult to find. The supervised methods achieve reasonable effectiveness but are usually domain specific and language dependent and they require labelled data, which is often labor intensive. Meanwhile, the unsupervised methods have high demand because publicly available data are often unlabelled and thus require robust solutions. Therefore, semi-supervised learning has been introduced and has attracted considerable attention in sentiment classification. In unsupervised learning, it uses a large amount of unlabelled data along with labeled data to build better learning models.

The most famous ML systems that have made extraordinary progress in text classification are Support Vector Machine, Naive Bayes and Maximum Entropy. The other surely understood ML strategies in characteristic dialect preparing are K-Nearest neighbor, ID3, C5, centroid classifier, winnow classifier and the N-gram model.

**3.1 Decision Tree Classifier**

In Decision Tree classifier, the interior nodes were marked with features and edges that are leaving the node were named as a trial on the data set weight. Leaves in the tree are good, by categorization. This categories whole document by starting at the root of the tree and moving successfully down through its branches till a leaf node is reached. Learning indecision tree adopts a decision tree classifier as an anticipated model in which it maps information of an item to conclusions of that item's expected value. In a decision tree, the large amount of input can figure out by using authoritative computing assets in the finite time. The main advantages of decision tree classifier are, it is easy to understand and to interpret. This classifier requires small data preparation. But these concepts can create complicated trees that do not generalized easily.

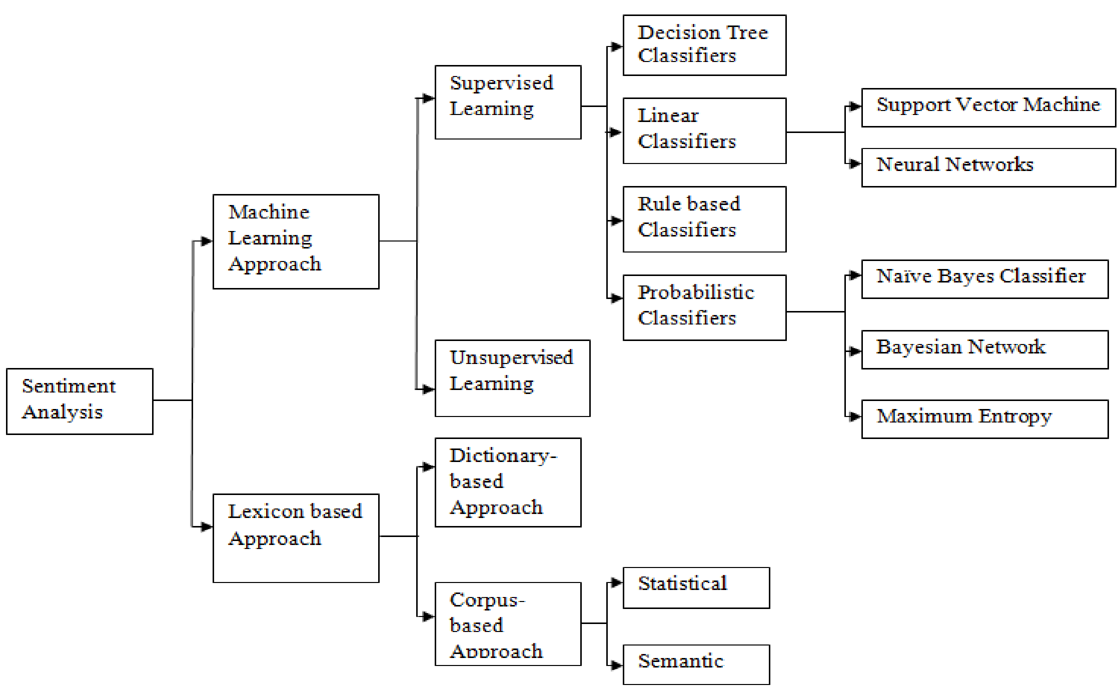


**Fig 2:** Sentiment Analysis Techniques

 

### 3.2 Linear Classifier

In linear classifier, for classifying input vectors to classes they use linear decision margins. There are many types of linear classifiers. Support vector machine is one of them. This classifier provides a good , scatter between various classes.

#### 3.2.1 Neural Network

Neural network includes numerous neurons in which this neuron is its elemental unit. Multilayer neural network was used with non-linear margins. The results of the neuron in the previous layer will be given as input for the next layer. In this type of classifier training of data set is more complicated, because the faults must be back-propagated for various layers.

#### 3.2.2 Support Vector Machine

Most accurate results in speech classification problems are achieved through Support vector machine. Hyper plane with maximal Euclidean distance for the nearest trained examples achieves solutions to the problem of classification. Support Vector Machine hyperplane is completely resolved by a comparatively minute subset of the trained data sets which are treated as support vectors. The remaining training data sets have no access to the qualified classifier. So scattered text classification, the classifier SVMs have been applied successfully and also used in different sequence processing application. SVMs are used in hypertext and text classification since they do not require labeled training dataset.

### 3.3 Rule-Based Classifier

As the name indicates in rule-based classifiers, data set is designed along with a group of rules. In rules left-hand side indicates the condition of aspect set and right hand indicates the class label.

### 3.4 Probabilistic Classifier

These classifiers utilize different types of arrangement. This assortment of structures takes every single class as a feature of that blend. Every different component are the gainful model in which it gives the likelihood of investigating an unmistakable word for that component. These classifiers are otherwise called generative classifiers. A portion of the probabilistic classifiers are Naïve Bayes, Bayesian Network and Maximum Entropy.

#### 3.4.1 Naïve Bayes Classifier

A Naive Bayesian classifier is one of the familiar supervised learning techniques which are frequently used for classification purpose. Their classifier is named as naïve since it considers the contingency that is actually linked are not depending on the further. Calculation of whole document feasibility would be the substance in an aggregation of all the feasibility report of a single word in the file. These Naïve Bayesian classifiers were frequently applied in sentiment categorization since they are having lower computing power when comparing to the other approach but independence assumptions will provide inaccurate results.

#### 3.4.2 Bayesian Network

The main disadvantage of Naïve Bayes classifier is its independent assumption of aspects in data sets. This assumption is the reason for the start of using Bayesian Network. This Bayesian network is directed non-cyclic graph where nodes correspond to variables and those edges are correspond to conditional independence. In text classification, Bayesian Network is not usually used since it is expensive in computation.

#### 3.4.3 Maximum Entropy

Maximum Entropy classifier is parameterized by a weight set that is used to associate with the joint-future, accomplished by a trained data set by encoding it. This Maximum Entropy classifier appears with the group of classifiers such as log-linear and exponential classifier, as its job is done by deriving some data sets against the input binding them directly and the result will be treated as its exponent.

### 3.5 K- Nearest Neighbour Classifier

K-Nearest Neighbour is an unsupervised learning algorithm for text classification. In this algorithm, the entity is classified with various trained data set along with their nearest distance against each entity. The advantage of this algorithm is its simplicity in text categorization. It also works well with multi-class text classification. The main drawback of KNN is it needs a large amount of time for categorizing entities where a huge data set is inclined.

**3.6 Strength/Sentiment Scoring:** Sentiment strength is calculated by manipulating the frequency of matched lexicons according to polarity. Extended studies in this challenge include prior polarity, dependency rules, negation identification and summarization. These approaches, however, are still far from being able to infer the cognitive and affective information associated with natural language, given that they mainly rely on knowledge bases that are still too limited to process text efficiently at the sentence level. Moreover, such text analysis granularity might still be insufficient, given that a single sentence may contain different opinions about different facets of the same product or service. To this end, concept-level SA aims to go beyond a mere word-level analysis of text to provide novel approaches to opinion mining and SA that enable more efficient passage from unstructured textual information to structured machine-processable data in any domain.

## IV. APPLICATIONS OF SENTIMENT ANALYSIS

Corporate brands want to hear from customers as social media access has become very easy with the advent of mobile and e-commerce. Analysing large number of reviews, complaints and compliments posted and shared just seconds after a new product is released, helps companies to accommodate this growing trend in order to achieve some business values like increasing the number of customers; enhancing customer loyalty, customer satisfaction and company reputation; and achieving higher sales and total revenue.

On the other hand, SA helps in extracting the strengths and weaknesses of the distinguishable features of each product, as well as finding the satisfaction levels of other users of those products. Besides the benefits of entrepreneurship, other benefits include political polarity, people's view on programmes. This information can be obtained by analyzing the sentiment orientation of comments, the number of likes, shares or comments on posted topics. Applications of SA range from public voice analysis, crowd surveillance, customer care and social intelligence-based SA to exploit the publics' online content generation for analyzing inputs such as pandemic spreading, emotion and responses towards local events. SA that focuses on microblogging is very typical because this is the main source that taps the public's voice. SA on microblogging data is more challenging compared to conventional texts such as documents review, due to the length, repeated use of some unofficial and atypical words and the rapid progress of language variation usage. The social network is also exploited to identify the most influential opinionators as a communication strategy which is useful during elections and disasters.

Affective computing through SA facilitates answers to questions such as 'What are the important themes that repeatedly feature in user comments?', 'What is the sentiment orientation of a specific gender about a specific post?' and 'What are the trends of happiness and sadness of the user over time?' Emotions in a text may be expressed explicitly (for example, emoticons and lexicon) as well as implicitly. Affective computing enables companies to care more about their customers and is useful for market prediction, assists in diagnosing patients' suicidal levels and allows the related parties to gauge public perception towards events). The advancements in affective computing allow applications to sense and deliver services tailored to customer needs, but issues such as privacy need to be observed. SA has also been tested in multilingual perspectives where the focus was to resolve the limitations of language dependent sentiment lexicons. Several approaches exist in this study, such as translating text into a reference language in which a sentiment lexicon is available before subsequently analyzing the text and mapping sentiment scores from a semantically enabled reference lexicon to a target lexicon by traversing relations between language-specific lexicons. These principles have encouraged many languages such as Dutch, Czech and Arabic to explore the potential of SA.

## V. Sentiment Analysis Tools

Shortlist of 10 practical tools to track user sentiment:

1. Meltwater: Assess the tone of the commentary as a proxy for reputation of brand and uncover new inputs that help you understand who the target audience are ?
2. Google Alerts: A simple and very useful way to monitor your search queries. It is used to track "content marketing" and get regular email updates on the latest relevant Google results. This is a good starting point for tracking influencers, trends and competitors.
3. People Browser: Find all the mentions of your brand, industry and competitors and analyze sentiment. This tool allows you to compare the volume of mentions before, during and after your marketing campaigns.
4. Google Analytics: A powerful tool for discovering which channels influenced your subscribers and buyers. Create custom reports, annotations to keep uninterrupted records of your marketing and web design actions, as well as advanced segments to breakdown visitor data and gain valuable insights on their online experiences.
5. HootSuite: A great freemium tool that allows you to manage and measure your social networks
6. Tweetstats: This is a fun, free tool that allows you to graph your Twitter stats.
7. Facebook Insights: If you have more than 30 Likes on your Facebook Page you can start measuring its performance with Insights. See total page Likes, a number of fans, daily active users, new Likes/Unlikes, Like sources, demographics, page views and unique page views, tab views, external referrers, media consumption and more!
8. Pagelever: This is another tool for measuring Facebook activity. Pagelever gives you the ability to precisely measure each stage of how content is consumed and shared on the Facebook platform.
9. Social Mention: The social media equivalent to Google Alerts, this is a useful tool that allows you to track mentions for identified keywords in a video, blogs, microblogs, events, bookmarks, comments, news, Q&A, hashtags and even audio media. It also indicates if mentions are positive, negative, or neutral.

 

10. Marketing Grader: Hubspot's Marketing Grader is a tool for grading your entire marketing funnel. It uses over 35 metrics to calculate your grade by looking at if you are regularly blog posting, Tweeting, updating on Facebook, converting visitors into leads, and more.

## VI. FUTURE OPPORTUNITIES AND CHALLENGES

SA has focused on the techniques, applications and web services but none of the available studies have focused on the SA approaches' adaptability for big data

Data generated for analysis is continuous and ever-changing, hence, there is increasing the possibility of new linguistic features being created, such as new acronyms, emoticons, idioms and terminologies, which require an update of the SA model.

Messages in social media are short with no proper grammar and wide use of acronyms and internet slag with minimizes the accuracy of text classification. Hence there is a need for more advanced SA techniques to adapt to new to new linguistic features and change in human communication text.
An existing approach based on fuzzy logic has been introduced for opinion mining on large-scale twitter data which was an attempt at mining the meaning of the texts according to the sentiment of the attributes in the text. Map Reduce framework was used to increase the speed for scanning the texts before the multi-attribute mining. Hierarchical Dirichlet Process-Latent Dirichlet Allocation (HDP-LDA) was applied for unsupervised aspect identification in the SA. These methods achieved good results. However, they have been tested on a prepared dataset mainly used for research. In fact, real data generated on social media contains vast amounts of noise. This indicates the need for a capability to sense and identify useful messages from the online media to be used as input for any strategic marketing.

SA techniques are commonly based on textual sources. In fact, many other multimedia sources should also be processed, some of which are important sources for examples exhibiting expressions of mocking, sabotaging and sarcasm, which is sensitive content for companies' reputations and for competitiveness planning. Therefore, multi-modal SA techniques are probably going to be in high demand in the near future.

Trustworthiness of the data in SA analysis is a challenge. Some SA techniques have focused on detecting deceptive reviews and cyber bullying messages Determining trustworthiness of the data demands more norms and logical reasoning which should be considered using many factors and not limited to only the current message being processed but also other messages being posted by the same message sender, for his profile to be considered. SA techniques should also be updated to be able to reason and determine the levels of uncertainty, validity, messiness and trustworthiness of the data.

Is Unsupervised learning capable of delivering similar accuracy to the one provided by Supervised Learning techniques in the determination of subjectivity in Sentiment Analysis?

Can we represent with more accuracy sentiments expressed in natural languages by using as a bedrock concepts of emotions that originate in psychology? Can we get closer to the heart of the matter by using this foundation and looking into the cognitive model of emotions or is doing so not worth?

Is our model flexible enough to attempt to accommodate the recognition and understanding of metaphors? Can this be achieved without the use of supervised or semi-supervised machine learning approaches?

## VII. CONCLUSION

The future of sentiment analysis vests not only in improving the accuracy and speed of various algorithms discussed but also in the area of whether we can correlate sentiment with behavior. To associate sentiment with behavior needs to be explored in the area of predictive analytics and Lexi analytics. The new demand would not be to say if its overall positive or negative but users want which part of the discussion or feedback is positive and negative. And also sentences of comparison in forums don't result in any polarity which needs semantic analysis. So I suppose there will be a trend towards greater use of NLP techniques (such as syntactic parsing, coreference resolution, etc), in addition to machine learning methods. Based on the applications and context of analysis various approaches are used , sometimes using more than one techniques.SA needs to be explored to other local languages such as Telugu and Hindi that would help organisation to target customers and products specifically to linguistic regions. To handle analysis that needs accuracy and efficiency an integrated system ( multi modal) that uses multiple techniques of sentimental analysis is the need of the day.
Trustworthiness of data in SA analysis and information for analysis in multimedia sources is a challenge. Unsupervised learning for classification of prediction is new and yet to be nailed still.

 

## REFERENCES


[1]. Harnani Mat Zin and Samaneh Nadali, Nurfadhlina Mohd Sharef. "Overview and Future Opportunities of Sentiment Analysis Approaches for Big Data." thescipub.com. http://thescipub.com/PDF/jcssp.2016.153.168.pdf

[2]. .Agnihotri, R., R. Dingus, M.Y. Hu and M.T. Krush, 2015. Social media: Influencing customer satisfaction in B2B sales. Industrial Market. Manage. DOI: 10.1016/j.indmarman.2015.09.003

[3]. **A**sur, S. and B.A. Huberman, 2010. Predicting the future with social media. Proceedings of the International Conference on Web Intelligence and Intelligent Agent Technology, Aug. 31 2010-Sept. 3, IEEE Xplore Press, Toronto, ON, pp: 492-499. DOI: 10.1109/WI-IAT.2010.63

[4]. Bagheri, A., M. Saraee and F. de Jong, 2013. Care more about customers: Unsupervised domain independent aspect detection for sentiment analysis of customer reviews. Knowledge-Based Syst., 52: 201-213. DOI: 10.1016/j.knosys.2013.08.011

[5]. Balahur, A., J.M. Hermida and A. Montoya, 2012. Detecting implicit expressions of emotion in a text: A comparative analysis. Decision Support Syst., 53: 742-753. DOI: 10.1016/j.dss.2012.05.024

[6]. Balahur, A., R. Mihalcea and A. Montoya, 2014. Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications. Comput. Speech Lang., 28: 1-6. DOI: 10.1016/j.csl.2013.09.003

[7]. Banea, C., R. Mihalcea and J. Wiebe, 2014. Sense-level subjectivity in a multilingual setting. Comput. Speech Lang., 28: 7-19. DOI: 10.1016/j.csl.2013.03.002

[8]. Batrinca, B. and P.C. Treleaven, 2014. Social media analytics: A survey of techniques, tools and platforms. Ai Society, 30: 89-116. DOI: 10.1007/s00146-014-0549-4

[9]. Bernabé-Moreno, J., A. Tejeda-Lorente, C. Porcel, H. Fujita and E. Herrera-Viedma, 2015. CARESOME: A system to enrich marketing customers acquisition and retention campaigns using social media information. Knowledge-Based Syst., 80: 163-179. DOI: 10.1016/j.knosys.2014.12.033

[10]. Bing, L.I. and K.C.C. Chan, 2014. A fuzzy logic approach for opinion mining on large-scale twitter data. Proceedings of the ACM 7th International Conference on Utility and Cloud Computing, 0Dec. 8-11, IEEE Xplore Press, London, pp: 652-657. DOI: 10.1109/UCC.2014.105

[11]. Bravo-Marquez, F., M. Mendoza and B. Poblete, 2014. Meta-level sentiment models for big social data analysis. Knowledge-Based Syst., 69: 86-99. DOI: 10.1016/j.knosys.2014.05.016

[12]. Cambria, E., M. Grassi, A. Hussain and C. Havasi, 2011. Sentic Computing for social media marketing. Multimedia Tools Applic., 59: 557-577. DOI: 10.1007/s11042-011-0815-0

[13]. Chaumartin, F., L. Talana and U. Paris, 2007. UPAR7: A knowledge-based system for headline sentiment tagging. Proceedings of the 4th International Workshop on Semantic Evaluations, (WSE' 07), Stroudsburg, PA, USA, pp: 422-425. DOI: 10.3115/1621474.1621568

[14]. Cheong, M. and V.C.S. Lee, 2010. A microblogging based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter. Inform. Syst. Frontiers, 13: 45-59. DOI: 10.1007/s10796-010-9273-x

[15]. Conejero, J., P. Burnap, O. Rana and J. Morgan, 2013. Scaling archived social media data analysis using a Hadoop cloud. Proceedings of the IEEE 6th International Conference on Cloud Computing, Jun. 28 Jul. 3, IEEE Xplore Press, Santa Clara, CA, pp: 685-692. DOI: 10.1109/CLOUD.2013.120

[16]. Da Silva, N.F.F., E.R. Hruschka and E.R. Hruschka, 2014. Tweet sentiment analysis with classifier ensembles. Decision Support Syst., 66: 170-179. DOI: 10.1016/j.dss.2014.07.003 De Vel, O., A. Anderson, M. Corney and G. Mohay, 2001. Mining e-mail content for author identification forensics. ACM SIGMOD Record, 30: 55-64. DOI: 10.1145/604264.604272

[17]. Denecke, K., 2008. Using SentiWordNet for multilingual sentiment analysis. Proceedings of the 24th International Conference on Data Engineering Workshop, Apr. 7-12, IEEE Xplore Press, Cancun, pp: 507-512. DOI: 10.1109/ICDEW.2008.4498370

[18]. Derczynski, L. and K. Bontcheva, 2014a. PHEME: Veracity in digital social networks. Proceedings of the Workshop on UMAP Projects Synergy, (UPS' 14), pp: 1-4.

[19]. Ding, X., B. Liu and S.M. Street, 2007. The utility of linguistic rules in opinion mining. Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Jul. 23-27, Amsterdam, pp: 811-812. DOI: 10.1145/1277741.1277921

[20]. Efron, M., M. Gary and Z. Julinang, 2004. Cultural orientation: Classifying subjective documents by Cociation analysis. Proceedings of the AAAI Fall Symposium Series on Style and Meaning in Language, Art, Music and Design, (AMD' 04), pp: 41-48.

[21]. Eirinaki, M., S. Pisal and J. Singh, 2012. Feature-based opinion mining and ranking. J. Comput. Syst. Sci., 78: 1175-1184. DOI: 10.1016/j.jcss.2011.10.007

[22]. Fei, X., H. Wang and J. Zhu, 2010. Sentiment word identification using the maximum entropy model. Proceedings of the 6th International Conference on Natural Language Processing and Knowledge Engineering, Aug. 21-23, IEEE Xplore Press, Beijing, pp: 1-4. DOI: 10.1109/NLPKE.2010.5587811

[23]. Fei, Z., J. Liu and G. Wu, 2004. Sentiment classification using phrase patterns. Proceedings of the 4th IEEE International Conference on Computer Information Technology, Sept. 14-16, IEEE Xplore Press, pp: 1147-1152. DOI: 10.1109/CIT.2004.1357349

[24]. Fulse, S., R. Sugandhi and A. Mahajan, 2014. A survey on multimodal sentiment analysis. Int. J. Eng. Res. Technol., 3: 1233-1238. Gamon, M., 2004.

[25]. Sentiment classification on customer feedback data: Noisy data, large feature vectors and the role of linguistic analysis. Proceedings of the 20th International Conference on Computational Linguistics, (CCL' 04), Stroudsburg, PA, USA, pp: 1-7. DOI: 10.3115/1220355.1220476

[26]. Ghazi, D., D. Inkpen and S. Szpakowicz, 2014. The prior and contextual emotion of words in sentential context. Comput. Speech Lang., 28: 76-92. DOI: 10.1016/j.csl.2013.04.009

[27]. Go, A., R. Bhayani and L. Huang, 2009. Twitter sentiment classification using distant supervision. Grefenstette, G., Y. Qu, J. Shanahan and D. Evans, 2004.

[28]. Guo, Y., Z. Shao and N. Hua, 2010. Automatic text categorization based on content analysis of cognitive situation models. Inform. Sci., 180: 613-630. DOI: 10.1016/j.ins.2009.11.012

[29]. Habernal, I., T. Ptáček and J. Steinberger, 2014. Supervised sentiment analysis in Czech social media. Inform. Process. Manage., 50: 693-707. DOI: 10.1016/j.ipm.2014.05.001

[30]. Hai, Z., K. Chang and J. Kim, 2011. Implicit feature identification via co-occurrence association rule mining. Proceedings of the 12th International Conference on Computational Linguistics and Intelligent Text Processing, Feb. 20-26, Springer Berlin Heidelberg, pp: 393-404. DOI: 10.1007/978-3-642-19400-9_31

[31]. Harrigan, P., G. Soutar, M.M. Choudhury and M. Lowe, 2014. Modelling CRM in a social media age. Aus. Market. J., 23: 27-37. DOI: 10.1016/j.ausmj.2014.11.001

[32]. Hatzivassiloglou, V. and K.R. McKeown, 1997. Predicting the semantic orientation of adjectives. Proceedings of the 8th Conference on European chapter of the Association for Computational Linguistics, (ACL' 97), Stroudsburg, PA, USA, pp: 174-181. DOI: 10.3115/976909.979640

[33]. He, W., H. Wu, G. Yan, V. Akula and J. Shen, 2015. A novel social media competitive analytics framework with sentiment benchmarks. Inform. Manage., 52: 801-812. DOI: 10.1016/j.im.2015.04.006

## Authors Profile

*Mr. Pranay Kumar BV :* pursed Bachelor of Engineering from Bangalore University Benguluru and Master of Technology from JNTU Hyderabada University in year 2009. He is currently working as Assistant Professor in Department of Computer Science and Engineering, Christu Jyothi institute of Technology and Science. He is a Life member of ISTE. He has published more than 8 research papers in reputed international His main research work focuses on Sentiment Analysis, Machine Learning, Big Data Analytics, Cryptography Algorithms, Network Security, Cloud Security and Privacy, Data Mining He has 13 years of teaching experience and 1 year of Research Experience.

*Dr.M Sadanandam :* pursed Bachelor of Technology and Master of Technology from JNT university Hyderabad.He completed his Ph.D in Computer Science and Engineering from JNT University Hyderabad .He is chairman, Board of studies in Department of Computer science and engineering, Kakatiya University Warangal India.He is member of IEEE, ISTE and computer society of India. He has published more than 20 research papers in reputed international journals including Thomson Reuters (SCI & Web of Science) and conferences including IEEE and it's also available online. His main research work focuses on Image processing, Cryptography Algorithms, Network Security, Cloud Security and Privacy, Big Data Analytics, Data Mining, IoT and Computational Intelligence based education. He has 15 years of teaching experience and 9 years of Research Experience.

 