

Opinion Mining Using Classification Techniques: A Review

Pooja Rani¹, Jaswinder Singh²

¹M.Tech Scholar, Department of Computer Science and Engineering

²Assistant Professor, Department of Computer Science and Engineering

Guru Jambheshwar University of Science & Technology, Hisar (Haryana) India

(E-mail: poojabhatia46@gmail.com)

Abstract- The concept of opinion mining or sentiment analysis is considered as the computational analysis of people's appraisals, opinions, behaviour, and feelings towards organizations, individuals, problems, events, themes and their characteristics. Reviews extracted from sentiment analysis are technically interesting and challenging. But it is also helpful in various fields for decision making. For instance, business always tries to find consumer or public opinions about their products. Prospective customers also need to know the views of existing users prior to use a service or buy a product. This paper highlights certain discussions regarding the review-based classification of internet data through the use of methods based on the concept of machine learning.

Keywords- sentiment, review, machine learning

I. INTRODUCTION

The process of sentiment analysis refers to emotions, feelings, attitude or opinion. With the development of technology based on World Wide Web, an individual or a group often explicit their opinions or sentiments over the technology of internet with the help of reviews, blogs, and ratings. With the help of such a rise in textual form of data, organization feels the desire to evaluate and figure out this type of text and further evaluate the business-based insights. Advertising companies and business owners often engage the process of sentiment analysis in order to explore new strategies of business and campaign of advertising. The algorithms of machine learning are mainly used for classifying and predicting the negative or positive kind of sentiments [1] [2]. The most famous categories of machine learning algorithm are the supervised and un-supervised categories. The supervised form of algorithm mainly uses a dataset of labelled form where each of the trained set is marked or labelled with an appropriate form of sentiment. Whereas, in the case of unsupervised learning includes an unlabelled form of dataset [10]. The algorithms of unsupervised learning are more of more complex form and it needs an additional algorithm of clustering in its initial implementation phase. This kind of study usually concerns with the techniques of supervised learning on a labelled form of dataset. The technique of SA can be operated on three distinct levels such as, aspect level, sentence level, and document level [3] [5] [11]. In case of classification of sentiments based on sentence level, there occurs an individual sentenced based polarity of a document whereas in case of classification base on aspect level, firstly it helps in identifying distinct forms of aspects. Document Level mainly aims in classification of the overall document or the topic as negative or positive [27].

1.1 Social Media Evolvement

In 1990s, the clients of the system were capable to make the content on their own in order to show the process by broadband internet representing a very popular/common technology. In early

1997, first website based on social network was appeared famously known as SixDegrees.com. From the year 2002 onwards, most number of the informal type of organization-based locales are derived. For instance, most of them are similar to Friendster whereas the others are similar to the concept of MySpace. In 2000, there was a broad increase in the process of online-based networking and it also increased the client's quantity/number. A large number of elements were involved into the area of online-based networking process [2] [4] [8]. These mainly embodies broadband connection accessibility, new devices of programming, and cell phones and personal computers development. In this 21st century, with enhancement of Internet, platform of social media for instance, the websites helps to allow their user clients to work and connect together in practical group [23]. Openness, rich experience of client, adaptability, insight aggregation, etc. are added form of newly built attributes in the process of online-based networking. The user of the system (clients) usually get supplementary facilities of make post pictures, like, audio upload, video-based content on the environment of online-based networking [6] [7] [24]. The posted data can be easily shared with most of the chosen users or clients or it navigates freely over the technology of web. The concept of social media has quickly expanded its roots all over the world due to merger of social links and innovations for the worth co-making.

1.2 Social Media: Characteristics

These usually represent traditional or conventional unbinding type of services.

- It helps in providing a space/venue for interpersonal or mass and online communicational process.
- It helps in permitting the individuals from all the corners of the globe providing a chance for communication with other.
- The locales-based fullest extent usually offers an ease to use the features that gets updated on itself basis in a constant manner [7] [9].
- These may particularize in a single repeated topic or obscure (covers up) a large number of themes.
- It allows the individuals to perform an activity that helps in advancement of their own while performing a mechanism of online-based finger printing, which helps in identifying of themselves [1] [4].
- It relies over give-and-take ideas and contributes newly formed communicational freeway process.
- Posted data permanent availability on the social media builds up a powerful platform.
- It helps to provide the ability of accessibility, search and replicability.

1.3 Different types of Social Media

The concept of networking is expanded everywhere throughout the technology of Internet. Clients of the system have several regions

of interest in order to utilize the outlets of online-based networking. Distinct media types are discussed as follows:

1. *Blogs*: It presents a focused streamline for journal or personal diary. It helps in providing a huge platform for expressing passions and thoughts across the world. It represents an online-based diary where the pages are shown typically in an inverted (backward) sequential order [2] [22]. Blog is of dynamic nature that can be remodelled and it helps in permitting the user guest to transmit through remarking segments combined to each and every single post. Web based journals are allowed to post freely on several networking websites.

2. *Wikis*: It presents the site-based aggregation allowing all to model a page by using the technology of Web browser. In actual manner, wikis usually present a CGI (common gateway interface) script blend and collection of transparent (plain) content-based document permitting the users to frame the pages of Web. Wikis mainly offer an efficient, effective yet flexible kind of communitarianism appropriate apparatus for designing specific sites of the Web.

3. *Social networking sites*: Social networking sites are popularly known as SNS which explains the expression mainly used to characterize any type of website permitting persons to grab the benefits of networking services which are available online presenting definition-based challenges. Some of the features present normally are: (a) individual to individual communication (b) UGC i.e. user-generated content known as the SNS lifeblood (c) user-based profiles in terms of specific services for several application that are kept by association of SNS, and (d) interpersonal-type of communication supports the online groups/communities progression by consumer profile interfacing mechanism with several packs and/or individuals. Social sites are continuously updated with addition of newly built devices, for instance video, photographs, blogging and sharing [10] [12] [13].

4. *Status-updating services*: Popularly known as administration-based micro-blogging process, for instance, Twitter helps to allow the clients to share or post tweets (short) regarding events or a person, and enables to watch the updates framed by other individuals [1] [8]. The technology of micro blogging represents a medium of broadcast existing in blogging form. Microblog variates from a traditional site as the content is of small size. It authorizes the users for using little content or text in terms of a substance, for instance "singular images, video joins or small tenses". These small kind of posts are known to be the micro posts [23].

5. *Media-shared sites*: It allows the users to share photographs or recordings on Pinterest or YouTube, etc. It further allows the users to share it on global basis or it may just choose selected person (s) or group (s). In addition, several destinations of media permits you to locate the media on distinct sites by the 'embedding' process. The technologies of Web 2.0 have resulted in an easy placement to offer and post media over numerous informal groups or communities and other enhanced phases [9] [26] [27]. The sharing of Media is viable for organizations or people to develop their reach and impact all over the world. Such type of mechanism authorizes the people or clients to share and post photographs or recordings [17] [20].

1.4 Classification Techniques

So many techniques of machine learning are adopted to categorize several reviews.

1. *Decision Tree*: The approach of Decision Tree useful is mainly used for classifying issues. In case of option-based testing, the experts use split percentage as the favored method.

2. *Naive Bayes*: The NB classifier presents basic probabilistic form of classifier on the basis of applying the Bayes based theorem. Naive Bayes evaluates probabilities set or group by consolidation-

based values in a known dataset [5] [14]. Also, this classifier has a quick process of decision-based making.

3. *K-nearest neighbours*: Also known as K-NN representing an algorithm based on lazy learning and it also presents an approach of non-parametric nature for classifying objects on the basis of closest form of training. The algorithm of K-NN indicates a very easy type of algorithm for the process of machine learning [2] [12]. The performance of algorithm based on K-NN classifier mainly depends over various distinct key factors, such as proper measure of distance, K parameter, and a measure of similarity [15].

4. *Support Vector Machine*: in machine learning, SVM presents a learning supervised model that helps in examining data and identifies several patterns. This technology is mainly used for analysis of classification and regression. Recently, a large number of classification-based algorithms have been planned and scheduled, yet SVM is considered as the most popular and widely used classifiers.

5. *Artificial neural networks*: Connectionist systems or ANNs represents the system of computing as encouraged by constituting the brain of animals [21]. Here, neural network does not represent an algorithm on itself basis, instead it presents a design for various distinct algorithms of in order to perform combined analysis and process the complex form of data-based inputs. Such type of systems mainly learns to execute various tasks by using examples that are not programmed in general with any kind of task-particularized rules. For example, in case of identifying, that on physical or manual basis has been as "cat" or "no cat" and in addition by using numerous results, the identification of cats in other type of images can be done [2] [13]. Experts perform the analysis without any kind of prior information about cats. ANN is usually based on an assortment of associated nodes or units known as that loosely designs in biologically built brain. Each of its connection, such as the property of in it, helps in transmitting a signal from a single artificially built neuron to the other. Such type of neuron which receives a signal can perform the operating process and further the signal artificial neurons in addition are associated to it. Commonly, in case of implementing ANNs, signal at a links among several artificial neurons presents a, and each artificial neuron output is mainly evaluated by various non-linearized functions of the input sums. ANN approach main aim was to solve issues. However, the analysis based on attention, over time moved to perform particular tasks, resulting in discrepancy. ANN has been used excellently over large number of tasks involving computer-based vision, recognition of speech, machine-based translation, social networking filtering process, medical and video games.

6. *Maximum Entropy*: The maximum entropy classifiers, does not contain assumptions regarding the relation among features. These kind of classifiers always help in trying to maximize the system-based entropy by estimation of conditional class label distribution.

II. RELATED WORK

Shidaganti, *et al.* [2] investigated over a method that presented a combined process of machine learning and data mining. The work proposed was done over tweeter data for the purpose of analyzing the tweet to collect the opinions of the user in regard to a specific type of topic or issue. The platform of tweeter is mainly used by individuals for expressing their sentiments (views) in a short type of message in response to several types of products, brands, celebrities, along with the political criticisms. In this type of work, clustering and TF-IDF process is mainly conversed about their effective efficiency. Mumtaz, *et al.* [3] viewed and expected a methodology presenting a combined form of lexical-based and machine learning approach. The hybrid approach i.e. proposed

provides high amount of accuracy than the classical-type of lexical method and it helps in providing the enhanced form of redundancy than the approach of machine learning. The approach i.e. proposed is mainly used for opinion/sentiment mining through the natural language processing (NLP) which helps in extracting the opinions /sentiments from the text associated with an entity. Balachandran, *et al.* [6] explained the present status about selecting the right type of institute which is considered as the most demanding task. In case of aspect-based, the process of SA is implemented directly over the reviews providing positive and negative reviews of the specific institution. Numerous techniques have been used for the purpose of aspect-based identification approaches like Machine Learning, Dictionary-based, NLP-based technique, unsupervised technique, Corpus based approaches. The suitable and the best possible result analytics is provided by the ML and the NLP type of classifiers. Hagge Marvin, *et al.* [7] discussed customer opinion by service of micro blogging service for instance, Twitter. The views of the consumer perform the sentiment analysis using aspect-based methodology with the help of parsing dependency and part-of-speech tagging from NLP and in return it help in extracting the tweet based negative, positive, or neutral aspects. In the proposed methodology, tool kit software was modelled such that it firstly extracted the tweets, then performed the process of filtering and after that analysed the polarity of sentiments thereby displaying the result. Here, persons have the ability to rent out their houses to other over a platform based on web technology. The considered aspects were view, place, room, apartment, home, people, day, night, for the purpose of analysis. The results obtained were displayed by using graphs. The future work depends over the Airbnb website reviews. Keumhee Kang, *et al.* [12] discussed a method for identification of clients with harsh moods by investigating their tweets on daily basis. Author usually exploited all the types of tweet-based media, i.e., emoticons and images along with texts. In order to check the proposed method validity, the mechanics of operation relied over the experimental observation of two experiments: 1) the proposed multimodal analysis which was tested with wide number of tweets, and the achievement obtained was measured with strength of sentiments i.e. senti-strength; 2) it was further applied in classification of 45 user-based mental states. Results based on experiment committed that approach proposed indicates the high accuracy than the existing-type of methods and such a model may help in prediction of the moods of an individual in a more efficiently manner. Zahrotun, *et al.* [15] examined about several techniques of clustering used in the process of mining the data. The technique of clustering presents the categorization of essential data which usually belongs to similar class or it may come under similar type of group. Cosine and Jaccard similarity, along with combinational analysis of both the processes is used for getting the flawless value-based similarity. The Cosine-based similarity helps in measuring the similarity provided within two of the non-zero type of vectors on the basis of inner space of product which evaluates the cosine angle between the product spaces. Virmani, *et al.* [17] explained the sentiment analysis collaboration with summarization, sentiment extraction and further maintains document of each student. Firstly, it sets the score for any kind of sentiment or opinion word. In this process, when an opinion word is encountered in sentence, it performs the operation of matching with the set of database and further it sets the score as per the requirement. Then from such type of scores the value based on cumulative opinion was evaluated. The algorithm provides an opinion-based numerical value. If in

case the value of numerical score is large then it provides a positive remark conclusion and if in case the value is low then it shows a negative remark. For instance, if the remarks provided by two teachers are extremely high and correspondingly the remarks provided by one teacher is very low then the process of collaboration occurs which will help in providing an averaging score. The performance on overall basis usually depends on the remarking state. Opinion word being used by faculty does not match with database word, which further affects the score on overall basis. Donglin, *et al.* [19] discussed the work based on visual SA. The researchers have presented several techniques used for analysis of visual sentiments. In such type of methodology, images were used for prediction. This survey describes new platform for researcher over textual content but visual-based sentiment ontology presented an advanced method for performing something new. The concept of deep learning approach must be adopted for visual-based sentiment analysis to work more effective. Basari, *et al.* [22] discussed the concept of opinions-based mining which referred to the application of computational linguistics, natural language processing and mining of the text in order to classify the good or bad movie on the basis of message opinion. SVM mainly presents a supervised method of learning which helps in analyzing the data and recognizes the patterns used for the purpose of classification. Abdul-Mageed, *et al.* [24] presented work on standardized version of Arabic data for the case of sentiment analysis. Here, a set of data was collected and further it performs an automatic classification of step where the process of tokenization was performed over the data. The process of two-stage classification was performed. The results obtained shows that the used approach works in an efficient and effective manner. A. M. Popescu, *et al.* [25] presented a method of unsupervised extraction of information that was used in extracting the review-based opinions. This kind of work was done using the following steps. In the first case, the product-based features were identified and secondly, the opinions associated with the product were identified and in the third case, the opinion's polarity was identified. Finally, the method proposed was rank and the opinions were based on analysis of strength. The semantically built orientation was obtained using an approach named relaxation-labelling approach. The results based on recall and precision method of the proposed approach effectively shows the effectiveness in identification of sentiments of an individual. X. Song, *et al.* [26] analysed the social-media based usage of the platform through the process of micro blogging and it extracts the informational data from them. The experts have analysed that the users of the system mainly use the platform of social media for updating their routine-based tasks and to know what is used by them on daily basis. Such type of analysis was based on distinct type of micro blogs platforms in distinct kind of geographical area based on time. P. D. Turney, *et al.* [27] suggested a supervised algorithm for learning that helps in classifying various reviews like thumbs up or down. The average-based semantic orientation was used to forecast the review-based classification. The negative and positive communication with the help of review presents the review-based orientation. The semantic-based orientation was evaluated using Point wise Mutual Information and Information Retrieval considered as vital step of the study. The algorithm proposed provides accuracy on distinct tweet types such as, automobiles and banks 80% and travel 84 %, and movies 74%. The summary of other related work is shown in Table 1.

Table.1 Existing Work

Author's Name	Year	Methodology Used	Proposed Work
Abdul-Mageed, <i>et al.</i> [24]	2011	Automatic Classification	Presented their work on standardized version of Arabic data for the case of sentiment analysis.
Virmani, <i>et al.</i> [17]	2014	Sentiment Analysis	Explained the sentiment analysis collaboration with summarization, extraction and further maintains document.
N. U. Pannala, <i>et al.</i> [11]	2016	Aspect based Sentiment Analysis	Discussed the opinion mining based existing work not on sentence-based level, but over the word-level.
Donglin, <i>et al.</i> [19]	2014	Leading Edge Methods	Surveyed the approaches based on visual SA. Such kind of survey was mainly presented defining the distinct used techniques for the analysis of visual sentiments.
Mohammad, <i>et al.</i> [9]	2017	Multimodal Sentiment Analysis	Performed the work on the analysis of multimodal kind of sentiment related to images, text video and audio shared by the individuals or the users
M. M. Fouad, <i>et al.</i> [1]	2018	Machine Learning	Proposed a model for tweeter sentiment analysis which describes the tweet is positive or negative by using the concept of machine learning
D. Mumtaz, <i>et al.</i> [3]	2018	Hybrid Approach	Proposed a methodology presenting a combined form of lexical-based and machine learning approach.
J. K. Rout, <i>et al.</i> [4]	2018	Supervised and Unsupervised Approach	Investigated the social-media based unstructured data such as tweeter for sentiment, emotion, and blogs analysis.
T. M. Jones, <i>et al.</i> [8]	2017	PMJDY scheme	Focused on the implication of PMJDY scheme for the development all over the India.
J. Singh, <i>et al.</i> [28]	2014	Jaccard and Cosine Similarity	Discussed about different similarity functions that are used to measure the similarity of document with the query in the field of information retrieval.
R. Rajnish, <i>et al.</i> [13]	2016	Machine Learning and Fuzzy Approach	Proposed a model on fuzzy logic for feature based opinion mining and sentiment analysis.
M. Al-Smadi, <i>et al.</i> [14]	2016	Long-Short Term Memory Based Neural Networks	Inspected the term sentiment analysis on the basis of hotel-based reviews in Arabic language using the methodology of memory-based neural networks on long-term basis.
L. Zahrotun, <i>et al.</i> [15]	2016	Jaccard and Cosine Similarity	Examined about several techniques of clustering used in the process of mining the data.
D. Cao, <i>et al.</i> [16]	2016	TagNet Tag clouds with improved node-link diagrams	Proposed a visualization approach called TagNet which is used for sentiment analysis.
W. Medhat, <i>et al.</i> [20]	2014	Pont-wise Mutual Information, Chi-square, Latent Semantic Indexing	Conferred about various applications of SA, recent modernized advancements in algorithm which were presented and investigated briefly in paper.
N. Agarwal, <i>et al.</i> [21]	2014	Jaccard and Cosine Similarity	Researched about several documents in the form of semi-structured, structured and unstructured data.

III. CONCLUSION

Sentiment analysis play essential role to create conclusion in appraisal domain. Due to its tremendous value, there has been an explosive growth of both research in academia and applications in

the industry. The developing field of sentiment analysis using various supervised learning techniques still requires a proper

platform for the research. In this paper, various supervised learning approaches have been used to classify the reviews so it would be valuable to combine different types of classifiers together to produce more accurate result.

REFERENCES

- [1] M. M. Fouad, T. F. Gharib, and A. S. Mashat, "Efficient Twitter Sentiment Analysis System with Feature Selection and Classifier Ensemble," in *The International Conference on Advanced Machine Learning Technologies and Applications (AMTA2018)*, vol. 723, pp. 516–527, 2018.
- [2] G. Shidaganti, R. G. Hulkund, and S. Prakash, "Analysis and Exploitation of Twitter Data Using Machine Learning Techniques," in *International Proceedings on Advances in Soft Computing, Intelligent Systems and Applications*, vol. 628, pp. 135–146, Springer Singapore, 2018.
- [3] D. Mumtaz and B. Ahuja, "A Lexical and Machine Learning-Based Hybrid System for Sentiment Analysis", in *Innovations in Computational Intelligence*, vol. 713, pp. 165–175, Springer Singapore, 2018.
- [4] J. K. Rout, K. K. R. Choo, A. K. Dash, S. Bakshi, S. K. Jena, and K. L. Williams, "A model for sentiment and emotion analysis of unstructured social media text," *Electron. Commer. Res.*, vol. 18, no. 1, pp. 181–199, 2018.
- [5] C. Bhadane, H. Dalal, and H. Doshi, "Sentiment Analysis: Measuring Opinions," *Procedia Computer Science*, vol. 45, pp. 808–814, 2015.
- [6] L. Balachandran and A. Kirupananda, "Online reviews evaluation system for higher education institution: An aspect based sentiment analysis tool," in *2017 11th Int. Conf. on Software, Knowledge, Inf. Manag. Appl. (SKIMA)*, pp. 1–7, 2017.
- [7] M. Hagge, M. Von Hoffen, J. H. Betzing, and J. Becker, "Design and implementation of a toolkit for the aspect-based sentiment analysis of tweets," *Proc. - 2017 IEEE 19th Conf. Bus. Informatics, CBI 2017*, vol. 1, pp. 379–387, 2017.
- [8] T. M. Jones, S. DivyaSri, and G. Bavani, "A Study on the Implications of Pradhan Mantri Jan Dhan Yojana on the Growth of Indian Economy," *IRA-International Journal of Management & Social Sciences*, vol. 06, no. 03, pp. 461–466, 2017.
- [9] M. A. Ullah, M. M. Islam, N. B. Azman, and Z. M. Zaki, "An overview of Multimodal Sentiment Analysis research: Opportunities and Difficulties," in *2017 IEEE Int. Conf. Imaging, Vis. Pattern Recognition, ICIVPR*, 2017.
- [10] N. Sharma, "Pradhan Mantri Jan Dhan Yojana (PMJDY) - A Conceptual Study," *Int. J. Res. Granthaalayah*, vol. 5, pp. 143–152, 2017.
- [11] N. U. Pannala, C. P. Nawarathna, J. T. K. Jayakody, L. Rupasinghe and K. Krishnadeva, "Supervised Learning Based Approach to Aspect Based Sentiment Analysis," *2016 IEEE International Conference on Computer and Information Technology (CIT)*, pp. 662-666, 2016.
- [12] K. Kang, C. Yoon, and E. Y. Kim, "Identifying depressive users in Twitter using multimodal analysis," *2016 Int. Conf. Big Data Smart Comput. BigComp 2016*, pp. 231–238, 2016.
- [13] R. Rajnish "Fuzzy Aspects in Sentiment Analysis and Opinion Mining" pp. 7750–7755, 2016.
- [14] M. Al-Smadi, M. Al-Ayyoub, H. Al-Sarhan, and Y. Jararwell, "An aspect-based sentiment analysis approach to evaluating Arabic news affect on readers," *J. Univers. Comput. Sci.*, vol. 22, no. 5, pp. 630–649, 2016.
- [15] L. Zahrotun, "Comparison Jaccard similarity, Cosine Similarity and Combined Both of the Data Clustering With Shared Nearest Neighbor Method," *Computer Engineering and Applications Journal*, vol. 5, no. 11, pp. 2252–4274, 2016.
- [16] R. Ji, D. Cao, Y. Zhou, and F. Chen, "Survey of visual sentiment prediction for social media analysis," *Front. Comput. Sci.*, vol. 10, no. 4, pp. 602–611, 2016.
- [17] D. Virmani, V. Malhotra, and R. Tyagi, "Sentiment Analysis Using Collaborated Opinion Mining," vol. 4, p. 4, January, 2014.
- [18] F. H. Khan, S. Bashir, and U. Qamar, "TOM: Twitter opinion mining framework using hybrid classification scheme," *Decision Support System*, vol. 57, pp. 245–257, 2014.
- [19] S. L. Ji, Donglin Cao, Rongrong Ji, Dazhen Lin, "Visual sentiment topic model based microblog image sentiment analysis," *Springer*, vol. 75, no. 15, pp. 8955–8968, 2014.
- [20] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications : A survey," *Ain Shams Eng. Jour.*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [21] N. Agarwal, M. Rawat, and M. Vijay, "Comparative Analysis Of Jaccard Coefficient and Cosine Similarity for Web Document Similarity Measure," *Int. J. Adv. Res. Eng. Technol.*, vol. 2, no. 5, pp. 18–21, 2014.
- [22] A. S. H. Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization," *Procedia Engineering*, vol. 53, pp. 453–462, 2013.
- [23] R. Socher, J. Pennington, E. H. Huang, A. Y. Ng, and C. D. Manning, "Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions," *EMNLP 2011 - Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, no. 2, pp. 151–161, 2011.
- [24] M. Abdul-Mageed, M. T. Diab, and M. Korayem, "Subjectivity and Sentiment Analysis of Modern Standard Arabic," *Proc. 49th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol.*, vol. 27, no. 1, pp. 587–591, 2011.
- [25] A. M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," *Nat. Lang. Process. Text Min.*, October, pp. 9–28, 2007.
- [26] A. Java, X. Song, T. Finin, and B. Tseng, "Why We Twitter: Understanding Microblogging Usage and Communities," in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis - WebKDD/SNA-KDD '07*, San Jose, California, pp. 56–65, 2007.
- [27] P. D. Turney, "Thumbs up or thumbs down? Semantic Orientation applied to Unsupervised Classification of Reviews," *Proc. 40th Annu. Meet. Assoc. Comput. Linguist.*, July, pp. 417–424, 2002.
- [28] Jaswinder Singh, P. Singh, and Y. Chaba, "A Study of Similarity Functions Used in Textual Information Retrieval in Wide Area Networks," in *International Journal of Computer Science and Information Technologies (IJCSIT)*, vol. 5, no. 6, pp. 7880-7884, 2014.