

Addressing Challenges of Opinion Mining in Businesses

Tabassum H Khan¹,

Ph.D Research Scholar at Department of Computer Science & Engineering¹

Dr. Vijay B Gadicha²

Associate Professor at Department of Computer Science & Engineering²

G H Raisonni University Amravati, India, vijay.gadicha@raisonni.net²

Abstract— Customer satisfaction has become the primary target for each successful running business currently. They have the true potential in uplifting a tiny business overnight to completely shutting down a brand. Real time analysis of customer opinion (sentiment) is the backbone of this success. Thus this piece of research tries to focus light on the challenges encountered by businesses in four categories of opinion mining. Natural Language Processing with machine learning algorithms works wonders for opinion mining. What is important is the correct choice of opinion mining for a specific type of business as each one of them demands an idiosyncratic approach. The research indeed shares a quick infusion for businesses to safeguard. Countless advice is beared by businesses to have real time analysis, but this research endeavour by showcasing its ideal infusion.

Keywords—Opinion Mining, Natural Language Processing, Machine Learning, Computational Linguistics, Sentiments, Customers , Businesses.

I. INTRODUCTION:

Opinion mining, or *sentiment analysis*, is a text analysis technique that uses computational linguistics and natural language processing to automatically identify and extract sentiment or opinion from within text (positive, negative, neutral, etc.).It allows you to get inside your customers' heads and find out what they like and dislike, and why, so you can create products and services that meet their needs. When you have the right tools, you can perform opinion mining automatically, on almost any form of unstructured text, with very little human input needed.

Are there certain aspects of your business that receive regular accolades but others that are just lukewarm, or worse? Do you want to perform market research to see how sentiment toward your brand compares to that of your competitors? Sentiment analysis can process thousands of pages, comments, emails, or surveys in just minutes for real-time results. Or you can perform

opinion mining over time to see how sentiment classification rises or falls.

NLP software allows you to train models to the specific terminology and criteria of your business for a consistently accurate and objective analysis of your customers' conversations. Save time and money and leave behind the

wavering subjectivity of manual human processing. Sentiment analysis is not a novel research theme. The use of automation in sentiment analysis has increased significantly in the past couple of years. Natural Language Processing, Machine Learning and Opinion mining are few streams of computer science on which the research theme is dependent. Nowadays, lots and lots of data is available and can be adopted for use. For making any important decisions, devising business strategies it is a necessity to analyse enormous amounts of data available from various sources. Social media is one of the upcoming sources providing indefinite data which is used for sentiment analysis. But the

data obtained from social media is disorganized. Obsolete content analysis methods focused on predicting topics. In the past few years opinions, emotions and sentiments are some qualities which are reflected by the content from social media. The volume of data obtained from social media is gigantic making it more complex to analyze. As a result, decrease in interest of semantic-based application and inclination towards statistics and visualization is observed[1]. Opinion Mining is a process for automatic extraction of knowledge and information from the opinion of others. Sentiment classification and analysis are trending fields of research predominant in Sentiment Analysis. They are the formalization for studying and construing opinions and sentiments. Opinion mining is used interchangeably with Sentiment Analysis, Opinion Extraction, Sentiment Mining, Subjective Analysis, Emotion Analysis, Review Mining to create a knowledgebase of opinions, views, recommendations, preferences in a more structured and explicit form. This knowledge base is very crucial to us when we have to make a decision. This is not only true for individuals but also true for organizations[2].

Text can be written using two writing styles: formal and informal writing style. Formal writing consists of Poetry, novels, plays, government/official documents. And informal text consists of chat room data, short message on social media, SMS. As Literary arts contains a lot of emotions, these literature pieces especially poems can be used for task of Sentiment Classification which is very challenging in computational point of view. And secondly, short messages like tweets, face book status, are also become the useful source for Opinion Mining and their sentiment analysis. Because of length constraints in these kinds of messages, Opinion Mining is very difficult on this dataset.[5]

II. PROCESS OF OPINION MINING:

The Architecture of Opinion mining is shown below. Right from the data collection feedback and surveys from the internet (it can be of way also) is faded to the Information Retrieval section where all the data is been cleaned by eliminating unnecessary data for the feedback and sending it to the Subjective Text section. In this section the data is clustered in a subjective manner like its been written in a specific point of opinion. Their

by the given to Opinion Classification which will compile the input to predict the input data to be positive, neutral or negative. Henceforth it is passed on to Opinion summarization which will now showcase the overall sentiment for the inputted data as positive or negative as a summary result.

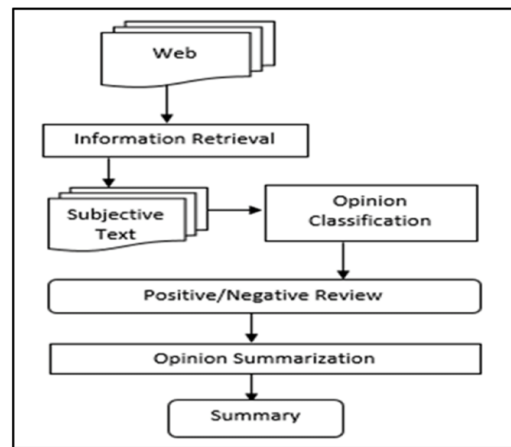


Figure1. Architecture of Opinion Mining[6]

III. CATEGORIES OF OPINION MINING:

Opinion mining and sentiment analysis models can focus on polarity of opinion (positive, neutral, negative), personal feelings (sad, angry, upset, happy, excited, and more), and intentions or objectives (*interested* or *not interested*).

Types of opinion mining:

- Fine-grained sentiment analysis
- Emotion detection
- Aspect-based sentiment analysis
- Multilingual sentiment analysis

Fine-grained sentiment analysis:

The most common use of opinion mining works to categorize comments and statements on a scale of opinion polarity. This can be simply positive, negative, or neutral, or you can go beyond this into fine-grained sentiment analysis with a larger scale of categories that include:

- Very positive
- Positive
- Neutral
- Negative
- Very negative

This is commonly used in opinion polls or surveys, as:

- Very Positive = 5 stars
- Very Negative = 1 star

However, text analysis tools allow you to analyze open-ended survey responses or track sentiment on social media, for example and categorize texts into pre-defined sentiments. It's a method for quantifying qualitative data.

Emotion detection:

This is opinion mining aimed at finding and extracting specific emotions (anger, disappointment, irritation, happiness, etc.) from text. Some emotion detection tools use lexicons, or word lists defined by the emotions they denote. This can be problematic as some words that often convey negative emotions, like *bad* or *kill* could also be used to express happiness or approval: "Your brand is killing it!" Advanced machine learning algorithms, on the other hand, allow text analysis programs to learn directly from sample text, so they can understand the nuances of human language, even to the point of detecting irony and sarcasm.

Aspect-based sentiment analysis:

When opinion mining text about your brand, you'll probably want to organize it into categories. If you're analyzing customer feedback, for example, you'd be able to categorize the text into aspects, like *Usability*, *Features*, *Shipping*, etc., then analyze each statement as positive, negative, or neutral. Aspect-based sentiment analysis would read the comment: "The new UI is amazing," as a positive opinion about the feature, "User Interface."

Multilingual sentiment analysis:

Multilingual sentiment analysis is often very difficult, as it involves a lot of preprocessing and resources. Some resources, like sentiment lexicons, are available online, while some, like translated corpora and noise detection algorithms, have to be built. And you'll need coding experience to put them into practice. Machines with emotions are no longer sci-fi with the headways of technology. It is essential to have a

superior There are various methods used for opinion mining and sentiment analysis among which following are the important ones [3]:

- 1) Naïve Bayes Classifier.
- 2) Support Vector Machine (SVM).
- 3) Multilayer Perceptron.
- 4) Clustering.

IV. CHALLENGES IN OPINION MINING :

Opinion mining has become really important as nowadays

for devising even smaller strategies, corpuses of data has to be analyzed. So solutions for opinion and sentiment mining are evolving at a great pace and reducing the burden of human shoulders. There are many techniques available for sentiment analysis. But still it's very difficult to say which technique works best because each technique has its own issues and certain challenges. Mainly there are two techniques used for opinion mining.

1) Lexicon based and 2) Learning based Lexicon based techniques involves high precision but on the other hand gives low recall. Also, another issue in this technique is that lexicons aren't available in all the languages. Learning based techniques make use of labeled examples to classify text. But it requires learning training as well as training dataset which becomes an issue too. Another technique-syntactic technique does yield good results but then it isn't language independent. Few other challenges that erupt during the process of classifying text in opinion mining have dependency on factors like: Value for n-when n-gram framework is used, choosing a higher value of n will degrade the performance. Occurrence of word-the number of occurrences of a word should be at the most 2-3 times for sentiment analysis.

Features to be used-Deciding what features must be used is also a challenge as the list of features returned by tokenization contains few irrelevant features. Hence it becomes necessary to select relevant features which will determine the precision of a classifier. People are habitual to using slang and casual language on social websites which makes it really difficult to predict people's opinion. So to eradicate such issues methods must be developed and existing ones

should be modified to adapt to the kind of language used on social websites. Now the term “spam messages” or “fake reviews” have become really common and used on a large scale on social websites. This creates a hurdle in the process of sentiment mining. So one of the biggest challenges is to identify such spam messages and fake reviews which can mainly be done through the comparison of qualitative with summary reviews.

Also, identifying duplicates, detecting outliers and knowing the reputation of the reviewer is to be kept in mind while implementing opinion mining. There are limitations in collaborative filtering which is responsible for identifying most famous concepts and suggest some out of the box thinking. Another challenge is the risk of filter bubbles, where a combination of automated content analysis with behavioral analysis proves to be very effective but eventually deviates from the selection of useful opinions making the user unaware of content that is different from what he expects in some manner. Integrating opinion with implicit data and behaviour to validate data and provide analysis beyond the opinion expressed is another common challenge. Asymmetry in available opinion mining software and the need for continuous improvement in the usability and user-friendliness of opinion mining software and other tools is another key challenge in the field of sentiment and opinion mining. Skewness in the dataset that is responsible for impacting recall is another challenge in sentiment mining[study].

There are many research challenges arises in this field, on the basis of our research study we analyze following issues:

- Major challenge arise in handling ambiguity in NLP as user may use correct semantics or may not use correct syntax.
- Some linguistic issue arise in opinion mining as language is not necessarily always English.
- Another challenge is the cost of tools which can only be afforded by large organizations and government funded groups.
- Another challenge is the domain dependency of words. One feature set may give good performance in one domain and poor in another.
- There is an asymmetry in the availability of opinion mining tools.[6]

Text used in social media, in particular micro-tweets, presents numerous challenges when

compared to formally structured text such as that presented in newspapers and scientific journals. As illustrated by Nabil et.al, the challenges related to the sentiment analysis systems to Twitter blogs arise due to several features of tweets. Tweets can contain unstructured language, numerous orthographic mistakes, slang words, ironic sentences, contractions, colloquial expressions, abbreviations or idiomatic expressions. Analysing tweets composed in Arabic is a particularly challenging task due to spelling inconsistencies, the use of connected words and a lack of capitalisation, which would otherwise be used to identify features. In addition, most people write tweets as they speak; for instance, the emotional character of Arabic tweeters and the frequent tendency to repeat letters for exaggeration, examples of which include ‘sorrriiiiiii’ and ‘noooooo’. Moreover, some characters have more than one form, an issue that highlights the need for normalisation, i.e., the unification of Arabic characters[12].

V. INFUSION:

Emotional Intelligence is an entire new horizon for the robot to still understand and conceive, Thinking about current generation expectation from human are fare beyond, can this be achieved from a robot pet or any other similar type machine [23] . Some shows some waved off literature context in the evolution of robotics in Artificial Intelligence. EI is being friendly, affectation, loving, concerning, reacting calmly in tuff situation, feelings, anger control, stress handling, being right at right time and so much .

Suggestion is to design an approach which will be capable of simulating all the EI emotions at a time and at right place. Thus, quantitative information shown in this article may perhaps best be regarded as indicative rather than authoritative. Despite such shortcomings, this article should be helpful in providing a base, for individuals in and out planning for future robotic research and development, applications and utilizations. Robotics and AI are playing an incredible vital responsibility in any country’s economic growth and be prepared likewise for its impact on the workforce structure .

VI.CONCLUSION:

Emotional Intelligence deep understanding is itself quite difficulty in Humans. Making EI work

for Robots will take miles for us. Certainly many stones have been unturned by researchers. The precision, emotions, affection, felling and decision making will still be a challenge in this domain. Emotional Intelligence Robot shall be a bliss for mankind in near future.

VII. REFERENCES

- [1] M.B. Nasreen Taj & G.S. Girisha, Insights of strength and weakness of evolving methodologies of sentiment analysis, *Global Transitions Proceedings 2* (2021) 157-162.
<http://www.keaipublishing.com/en/journals/global-transitions-proceedings/>
- [2] Prajna. P & Anubha. R, Opinion Mining: A Communication Tool For Career Making, *Gedrag & Organisatie Review - Issn:0921-5077, Volume 33 : Issue 02 - 2020.*
- [3] Pravesh Kumar Singh & Mohd Shahid Husain, METHODOLOGICAL STUDY OF OPINION MINING AND SENTIMENT ANALYSIS TECHNIQUES, *International Journal on Soft Computing (IJSC) Vol. 5, No. 1, February 2014.*
- [4] Kumar Ravi and Vadlamani Ravi, A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems 89* (2015) 14–46
<http://dx.doi.org/10.1016/j.knosys.2015.06.015>, June 2015.
- [5] Jasleen Kaur and Dr. Jatinder Kumar R. Saini¹, An Analysis of Opinion Mining Research Works Based on Language, Writing Style and Feature Selection Parameters, *International Journal of Advanced Networking Applications (IJANA) ISSN No. : 0975-0290 November 2014.*
- [6] Harshi Garg and Niranjana Lal, *Data Analysis: Opinion Mining and Sentiment Analysis of Opinionated Unstructured Data*, © Springer Nature Singapore Pte Ltd. 2018. M. Singh et al. (Eds.): ICACDS 2018, CCIS 906, pp. 249–258, 2018. https://doi.org/10.1007/978-981-13-1813-9_25.
- [7] Elham Asani a, Hamed Vahdat-Nejad a & Javad Sadri, Restaurant recommender system based on sentiment analysis, *Machine Learning with Applications 6* (2021) 100114- journal homepage: www.elsevier.com/locate/mlwa
<https://doi.org/10.1016/j.mlwa.2021.100114>,
- [8] Ana Reyes-Menendez, Jose Ramon Saura & Ferrao Filipe, Marketing challenges in the #MeToo era: gaining business insights using an exploratory sentiment analysis, <https://doi.org/10.1016/j.heliyon.2020.e03626>, September 2019; 2405-8440/© 2020 Published by Elsevier Ltd.
- [9] Walaa Medhat, Ahmed Hassan & Hoda Korashy, Sentiment analysis algorithms and applications: A survey, *Ain Shams Engineering Journal* (2014) 5, 1093-1113, www.elsevier.com/locate/asej, www.sciencedirect.com, 2090-4479 2014 Production and hosting by Elsevier B.V. on behalf of Ain Shams University. <http://dx.doi.org/10.1016/j.asej.2014.04.011>
- [10] Ravinder Ahujaa, Aakarsha Chuga, Shruti Kohlia, Shaurya Gupta and Pratyush Ahujaa, The Impact of Features Extraction on the Sentiment Analysis, *Procedia Computer Science 152* (2019) 341–348, www.sciencedirect.com
- [11] Akilandeswari Ja & Jothi Gb, Sentiment Classification of Tweets with Non-Language Features, *Procedia Computer Science 143* (2018) 426–433, www.sciencedirect.com, 8th International Conference on Advances in Computing and Communication (ICACC-2018)
- [12] Ghadah Alwakid, Taha Osman and Thomas Hughes-Roberts, Challenges in Sentiment Analysis for Arabic Social Networks, 3rd International Conference on Arabic Computational Linguistics, ACLing 2017, 5-6 November 2017, Dubai, United Arab Emirates, *Procedia Computer Science 117* (2017) 89–100
- [13] Abinash Tripathya, Ankit Agrawalb & Santanu Kumar Rath, Classification of

Sentimental Reviews Using Machine Learning Techniques, 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015) doi: 10.1016/j.procs.2015.07.523, Procedia Computer Science 57 (2015) 821 – 829, www.sciencedirect.com, 1877-0509 © 2015

[14] Doaa Mohey El-Din Mohamed Hussein, A survey on sentiment analysis challenges, Journal of King Saud University – Engineering Sciences (2018) 30, 330-338 , www.ksu.edu.sa, www.sciencedirect.com

[15] Khairullah Khan, Baharum Baharudin , Aurnagzeb Khan & Ashraf Ullah, Mining opinion components from unstructured reviews: A review, Journal of King Saud University – Computer and Information Sciences (2014) 26, 258-275. www.ksu.edu.sa , www.sciencedirect.com

[16] Rushabh Shah and Bhoomit Patel, Procedure of Opinion Mining and Sentiment Analysis: A Study, International Journal of Current Engineering and Technology, E-ISSN 2277 – 4106, P-ISSN 2347 - 5161 ©2014 INPRESSCO.

[17] Emma Haddia, Xiaohui Liua, Yong Shib (2013) The Role of Text Pre-processing in Sentiment Analysis. Procedia Computer Science 17, 26 – 32.

[18] Mita K. Dalal and Mukesh A. Zaveri (2013) Semi Supervised Learning Based Opinion Summarization and Classification for Online Product Reviews. Hindawi Publishing Corporation, Applied Computational Intelligence and Soft Computing, Article ID 910706.

[19] Walaa Medhat , Ahmed Hassan , Hoda Korashy (2014)

Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal.

[20] Lalita Sharma, Shweta Shukla. Classification of Web Blog Mining for Movie Review. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958.

[20] Meishan Hu, Aixin Sun and Ee-Peng Lim (2007) Comments-oriented blog summarization by sentence-extraction. In Proceedings of the

ACM SIGIR Conference on Information and Knowledge Management (CIKM), 901-904,. ISBN 78-1- 59593-803-9, Post paper.

[21] Charlotta Engstrom (2004) Topic dependence in sentiment classification. Master's thesis, University of Cambridge.

[22] Soo-Min Kim and Eduard Hovy (2004) Determining the sentiment of opinions. In Proceedings of the International Conference on Computational Linguistics (COLING).

[23] Abbasi, "Affect intensity analysis of dark web forums," in Proceedings of Intelligence and Security Informatics (ISI), pp. 282–288, 2007.

[24] K. Dave, S. Lawrence & D. Pennock. \Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews." Proceedings of the 12th International Conference on World Wide Web, pp. 519-528, 2003.

[24] B. Liu. \Web Data Mining: Exploring hyperlinks, contents, and usage data," Opinion Mining. Springer, 2007.

[25] B. Pang & L. Lee, \Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales." Proceedings of the Association for Computational Linguistics (ACL), pp. 15124, 2005.

[26] Nilesh M. Shelke, Shriniwas Deshpande, Vilas Thakre, Survey of Techniques for Opinion Mining, International Journal of Computer Applications (0975 – 8887) Volume 57– No.13, November 2012.

[27] Ion SMEUREANU, Cristian BUCUR, Applying Supervised Opinion Mining Techniques on Online User Reviews, Informatica Economică vol. 16, no. 2/2012.

[28] Bo Pang and Lillian Lee, "Opinion Mining and Sentiment Analysis", Foundations and Trends_R in Information Retrieval Vol. 2, Nos. 1–2 (2008).

[29] Vaghela, V.B., Jadav, B.M.: Analysis of various sentiment classification techniques. Int.J. Comput. Appl. 140, 1–6 (2016)

- [30]Hamzah, A., Widyastuti, N.: Opinion classification using maximum entropy and K-means clustering. In: International Conference on Information, Communication Technology and System, IEEE International Conference, pp. 162–166 (2016)
- [31]Songpan, W.: The analysis and prediction of customer review rating using opinion mining. In: IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA), pp. 71–77 (2017)
- [32]Hailong, Z., Wenyan, G., Bo, J.: Machine learning and lexicon based methods for sentiment classification: a survey. In: IEEE International Conference on Web Information System and Application (WISA), pp. 262–265 (2014)
- [33]Guellil, I., Boukhalfa, K.: Social big data mining: a survey focused on opinion mining and sentiments analysis. In: 12th International Symposium on Programming and Systems (ISPS). IEEE (2015)
- [34]Osimo, D., Mureddu, F.: Research challenge on opinion mining and sentiment analysis. In: The CROSSROAD Roadmap on ICT for Governance and Policy Modeling (2010)
- [35]Vo, A.-D., Nguyen, Q.-P., Ock, C.-Y.: Opinion—aspect relations in cognizing customer feelings via reviews. In: IEEE, pp. 5415–5426 (2018)
- [36]Denecke, K. (2008). Using sentiwordnet for multilingual sentiment analysis. In IEEE 24th international conference on data engineering workshop. Cancun, Mexico.
- [37]Esmaeili, L., Mardani, S., Hashemi Golpayegani, S. A., & Zanganeh Madar, Z. (2020). A novel tourism recommender system in the context of social commerce. *Expert Systems with Applications*, 149(1), A novel tourism recommender system in the context of social commerce.
- [38]García-Sánchez, F., Colomo-Palacios, R & alencia-García, R. (2020). A social-semantic recommender system for advertisements. *Information Processing & Management*, 57(2), Article 102153.
- [39]Ghiassi, M., & Lee, S. (2018). A domain transferable lexicon set for Twitter senti- ment analysis using a supervised machine learning approach. *Expert Systems with Applications*, 106(1), 197–216.
- [40]Hedge, S. B., Satyappanavar, S., & Setty, S. (2018). Sentiment based food classification for restaurant business. In *International conference on advances in computing, communications and informatics* (pp. 1455–1462). Bangalore, India: IEEE.