

# Performing Sentiment Analysis on Product Based Twitter Data Using Fuzzification and Defuzzification Process

Akkireddi Vara Prasad <sup>1</sup>, V.Satya Aruna <sup>2</sup>

<sup>1</sup> Assistant Professor, GITAM CDL, GITAM Deemed to be University, Visakhapatnam, Andhra Pradesh, India.

<sup>2</sup> Associate Professor, Lingaya's Institute of Management and Technology, Vijayawada, Andhra Pradesh, India.

## Abstract

In the current trend of business intelligence, internet based social media network is playing key role on decision making process. Much of this data depends on consumer perceptions and opinions of organizations such as interest to business intelligence analysts in marketing, customer relationship management and customer retention. Minute to minute, the opinion mining and sentiment analysis is changing rapidly due to the availability of vast and recent data in the social media networks. Also, the customers who are tech savvy and intelligent consumers mostly responding on social media network especially they are using twitter as their main platform. In this regard, twitter data for market analysis and business intelligence has been used to develop a model which would facilitate a decision making. The proposed model considers all sorts of data like, text reviews, symbols, emojis, hash tags; due to the fact that, business analyst, who needs actual strategic information should consider all types of data to get pure accurate results finally. The statements which are not precise, vague and with ambiguity needs filtering of unnecessary parts and extract necessary information. Boolean logic used in this research provide clear verdict either positive or negative. Sometimes, it is also mandatory to consider other opinions at extreme ends like completely positive, partially positive, completely negative and partially negative in order to make certain decisions. By applying Nearest Neighborhood concept, the specific information will be transformed in to generic form so that uncertain ranges can be identified. In this paper, by using fundamental Intelligence control algorithm, the fuzzification and defuzzification process estimates accurate results. At the same time all sorts of data with different structures are allowed to investigate in the current process.

**KEYWORDS:** Fuzzification, Defuzzification, Sentiment Analysis, Nearest Neighborhood, Business Intelligence

## 1. Introduction

Emotions are the complex state of feelings that cannot be represented by traditional Boolean logic. Emotion is a subjective, conscious experience characterized mainly by physiological expressions, biological reactions, and mental states [6]. Emotion is associated significantly with mood, nature, personality, disposition and motivation. Emotion is a positive or negative experience that is associated with a particular pattern of physiological activity. Humans carry lot of emotions like happiness, sadness, angry, disgust, surprise, panic, etc.

## 1.1. Opinion Mining

The term Sentiment Analysis[1] first appeared in (Nasukawa& YI 2003) and the term Opinion Mining first appeared in (Dave, Lawrence & Pennock 2003). Opinion Mining also called as sentiment analysis is a field that is used to extract the users opinions about any product, service or any topic. Long time back before emergence of Web 2.0, when an individual want to buy any product or utilize any service, he asks or discuss with the peer/friends about particular product & then take decision accordingly. But now-a-days Social media emerge as a boom with Terabytes of data available [17] about every product or a service. Opinion mining's main objective is to classify the reviews expressed by users into positive or negative & then present it in summarize form that is understood by user. Different Data Sources like Blogs, Social network Sites, E-commerce sites are available where users can post their reviews about particular topic (like Modi's campaign of Swachh Bharat) or any product like mobiles, cameras, etc. In order to promote marketing, large companies and business people are making use of Opinion Mining [1], thus we can say that opinion mining and Sentiment Analysis had tremendous opportunity in every domain. The growing importance of sentiment analysis coincides with the growth of social media such as reviews, forums, discussion groups, chatting, blogs, micro-blogs, twitter and social networks.

## 1.2. Sentiment Analysis

Sentiment analysis is consumers review [11] on products and services which helps both the producers and consumers (stakeholders)[7] to take effective and efficient decision within a shortest period of time. Producers can have better knowledge of their products and services through the sentiment analysis (ex. positive and negative comments or consumers likes and dislikes) which will help them to know their products status (ex. product limitations or market status). Consumers can have better knowledge of their interested products and services through the sentiment analysis (ex. positive and negative comments or consumers likes and dislikes) which will help them to know their deserving products status (ex. product limitations or market status)[8]. For more specification of the sentiment values, fuzzy logic could be introduced. Therefore, sentiment analysis with the help of fuzzy logic (deals with reasoning and gives closer views to the exact sentiment values) will help the producers or consumers for taking the effective decision according to their product or service interest.

Fuzzy Aspects in Sentiment Analysis and Opinion Mining with the growing availability and popularity of opinion-rich resources such as online review sites and personal blogs where new opportunities and challenges arises as people are now actively using information technologies to seek out and understand the opinions of others. Opinion Mining and Sentiment Analysis have thus become more important and significant. Opinions are fuzzy in nature and expressions can be beyond true or false or 0 or 1. Fuzzy sets constitute the oldest and most reported soft computing paradigm. They are well-suited to modeling different forms of uncertainties and ambiguities, often encountered in real life. Integration of fuzzy sets with other soft computing tools has led to the generation of more powerful, intelligent and efficient systems. This paper discusses about different levels of opinion mining, sentiment lexicon, issues and challenges in Opinion Mining & Sentiment Analysis, concept of Fuzzy logic and how well it suits to the problem of Opinion Mining & Sentiment Analysis.

## 2. Methodology, Development and Analysis

### 2.1. Dataset

The dataset consists of tweets which contain different weightages for every tweet. The weightages for the tweets can be done based on the words that are used in the tweet. Based on the weightages given to a tweet, the tweets are categorized into positive, negative, partially positive, partially negative, etc. Also, in order to ascertain, a tweet is positive/negative/partially positive/partially negative, the principles of fuzzy logic have been applied for this purpose.

➤ **Membership function design and assessment to the nearest neighborhood valued:**

The crisp values of the linguistic variables are transformed to the given fuzzy sets by the function known as Membership Function (MF). The inputs are mapped into the degree of membership by suitable membership functions for the partitions of linguistic variables. There are many forms of membership functions. The selection of appropriate membership functions for fuzzy sets is significant in a fuzzy system which perfectly represents the fuzzy modeling [9]. In the proposed research, the classification labels for the intensity of semantic orientation and positive or negative polarity [11][12] of a given sentence are identified as: logic and fuzzy logic both are differed with transparent line. Boolean logic states “true” and “false”, but adding with fuzziness describes partial and pure results also i.e., along with range and degree. It varies and describes completely true, partially true, completely false, and partially false which states the model of uncertainty of natural language processing.

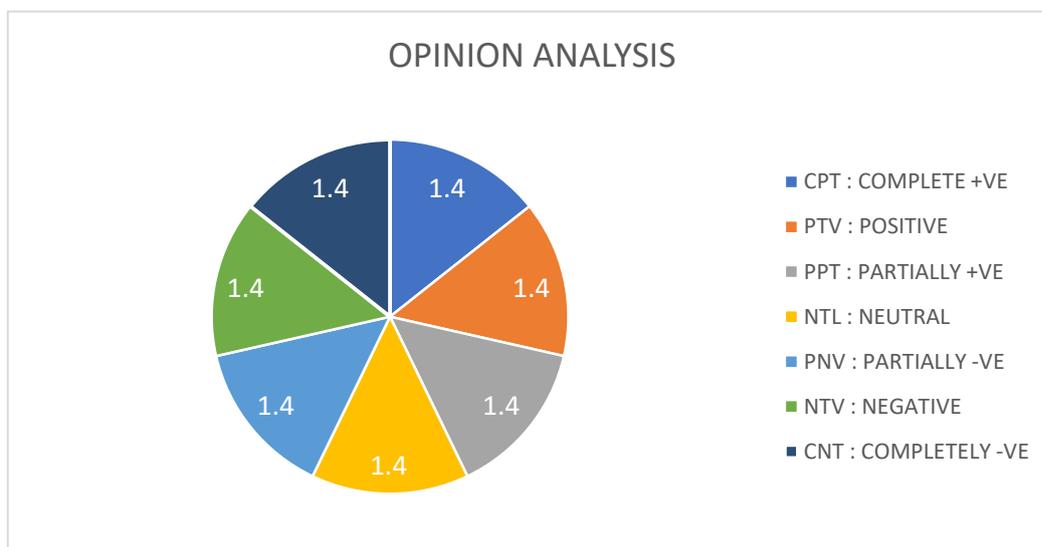
### 2.2. Fuzzification

It is the process of fuzzy theory to generalize crisp [18] or discrete values. That is about the way of transformation specific to generic. Fuzzy set estimation is gradually done [9] in a systematic measure from specific to generic. Membership to non membership will be done for various classes or objects. With the proper observation on human thinking and the nature of reasoning is always not traditional like true, false or neutral [10]. But sometimes the inference rules along with fuzziness will be obtained in a natural manner. In general, the set theory order pair (0,1) ranges clearly from (false, true) as in Boolean logic. So the reference of membership functions as fuzzy subset. Hence the appliances like electronic controlled machines in the industries and medical instruments in the hospitals, all the decision making systems used in the current market trends using fuzzy logic from fuzzy theory. Reasons to apply Fuzzy Logic on decision support systems on large amount of data which is freely and widely available continuously with rapid speed like internet social media network [2].

1. It is flexible.
2. Can be applied on imprecise data.
3. Can recognize non classical words, etc.
4. Human expressions and emotions can be recognized.

5. Sentiment Analysis on Bayesian N/W are used.
6. Can built an intelligent controlled system [13][14].
7. Fuzziness in the sentiment analysis can be straight forward.
8. Existing system like Opinion Mining is a holistic lexicon method having no further classifications other than Positive, Negative and Neutral.
9. Increases accuracy gradually while in classification approach.
10. Lexicon series of phases can be established basing on range of fuzzy set.

**A. Fuzzy System Opinion Analysis – Ranges**



**B. Fuzzy Logic and Truth Table:** The assessment[3] of the performance[8] of any product in the current market trend exclusively taken from twitter data[17] which is recently available is considered with the following measurement. This assessment[3] requires fuzzy sets. Let X be a product and ranges[15] from 0 to 1 i.e.,  $X \rightarrow \{0,1\}$ . It is the multilevel feedback Truth table (Table 1) for fuzzy logic

**Table 1. Fuzzy Controlled System reaching nearest Neighborhood**

$X \rightarrow \{0..1\}$	CPT	PTV	PPT	NTL	PNV	NTV	CNT
CPT	CPT	PTV	PTV	NTL	NTV	NTV	NTV
PTV	PTV	PTV	PPT	NTL	PNV	NTV	NTV
PPT	PTV	PPT	PPT	NTL	PNV	NTV	CNT
NTL	NTL	NTL	NTL	NTL	NTL	NTV	CNT
PNV	NTV	PNV	PNV	NTL	PNV	NTV	NTV
NTV	NTV	NTV	NTV	NTV	NTV	NTV	CNT
CNT	NTV	NTV	CNT	CNT	NTV	CNT	CNT

**C. Basis for Algorithm development:**

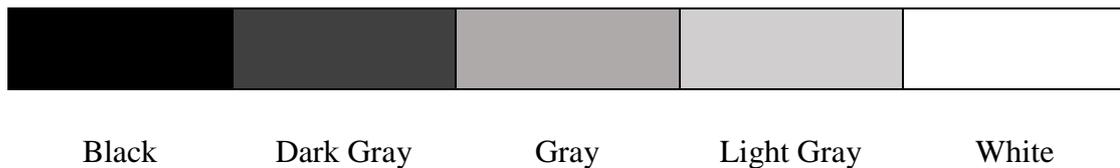
**Rules and ranges for defuzzification:** *X ranges from 0 to 1*

**0.0    0.125    0.25    0.375    0.5    0.625    0.75    0.875    1.0**

X is the product	( $x \geq 0$ and $x \leq 0.125$ )	completely positive.
	( $x > 0.125$ and $x \leq 0.25$ )	positive.
	( $x > 0.25$ and $x \leq 0.375$ )	partially positive.
	( $x > 0.375$ and $x \leq 0.5$ )	Neutral.
	( $x < 0.5$ and $x \leq 0.625$ )	Neutral.
	( $x > 0.625$ and $x \leq 0.75$ )	partially negative.
	( $x > 0.75$ and $x \leq 0.875$ )	negative.
	( $x > 0.875$ and $x \leq 1.0$ )	completely negative.

Fuzzy logic along with fuzzy thinking can be in a fuzzy set fashion[18] with lower and upper bound the above ranges can classify data with appropriate reasoning.

➤ **End to End Product based performance calculation:**



**3. Fuzzy Logic Controlled Intelligence Algorithm**

Step 1: Start process

Step 2: Collect exclusive twitter data.

Step 3: Extract product(X) based reviews.

Extraction based upon

- Text
- Sentences
- Hash tags
- Emojis
- Symbols

Step 4: start processing

Range\_value\_Count( )

i=0

repeat

Get each data item value

If sent\_val = range(i) then

Get sent\_val\_classify\_word.  
i=i+1.  
Until i>=max\_range\_value..

Step 5: Apply nearest neighborhood value of the fuzzy set.

Step 6: Calculate resultant value from defuzzification.

Step 7: Get the result. Add to range value\_count()

Step 8: Repeat the process from step 3 for the next\_sent

Until the sent\_data = "NULL".

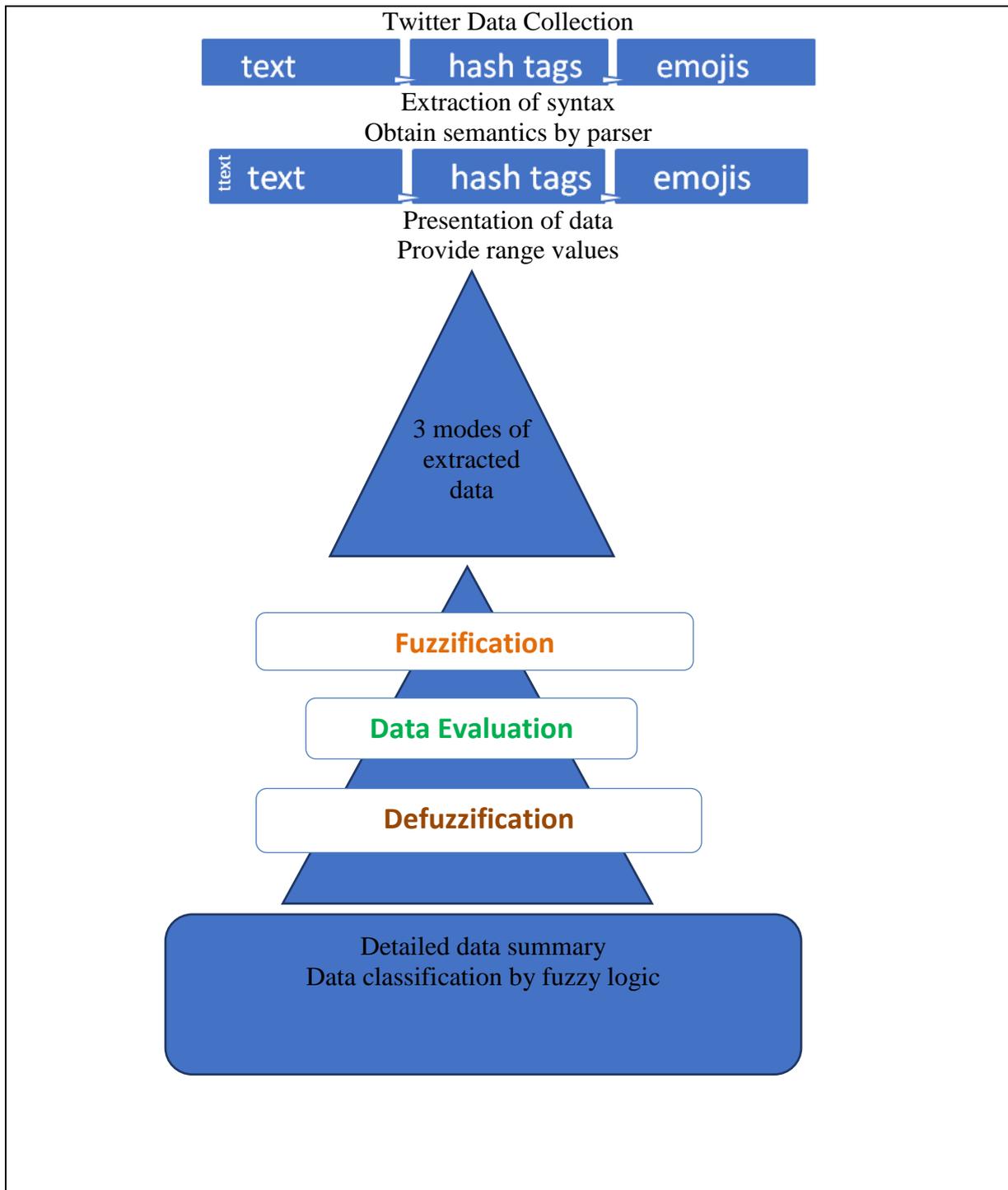
Step 9: Generate the final classification.

Step 10: Print classified labeled value.

Step 11: Stop process.

### 3.1. Detailed process flow state diagram

Crisp I/P → Fuzzification → rule based & fuzzy inference based → defuzzification → crisp O/P



#### 4. Conclusion

In this paper, it is proposed a model with fundamental controlled intelligent algorithm which is used to determine and classify sentiments along with opinions exclusively with twitter data from extraction to classification of any of the product reviews. The crisp input is transformed in to crisp output with fuzzification and defuzzification process. The algorithm is suitable for product reviews which include text, hash tags or emojis etc. Further to provide more accuracy in the data classification, the present algorithm requires inference rules with more ranges.

## 5. References

- [1] Animesh Kar, Deba P. Mandal, (2011), "Finding Opinion Strength Using Fuzzy Logic on Web Reviews", *International Journal of Engineering and Industries*, Vol. 2, No. 1, March 2010. Pp. 37-43.
- [2] Bakhtawar Seerat, Farouque Azam (2012), "Opinion Mining: Issues and Challenges (A Survey)", *International Journal of Computer Applications* (0975 – 8887) Volume 49–No.9, July 2012 42.
- [3] Guohong Fu, Xin Wang (2010), 'Chinese Sentence-Level Sentiment Classification Based on Fuzzy Sets', *Coling 2010: Poster Volume*, Pages 312-319, Beijing, August 2010.
- [4] S. Padmaga & Prof. S. Sane Fatima, "Opinion Mining & SA, An Assessment of people's belief - A survey", *International Journal of AD hoc, Sensor & Computing (IJASUE)* vol4, feb-2013.
- [5] G. Angulakshmi & D R. Mancika Chezian, "An Analysis on Opinion Mining- Techniques & Tools", *IJARCCCE-2014- Vol 3 Issue 7*.
- [6] Ghose A, "Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics", *Knowledge & data Engineering, IEEE Transaction on* (Vol 23, Issue:10)
- [7] Vinodhini G., RM. Chandrasekaran (2012), "Sentiment Analysis and Opinion Mining: A Survey", *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 2, Issue 6, June 2012, ISSN: 2277 128X, Available online at: [www.ijarcsse.com](http://www.ijarcsse.com).
- [8] K. Indhuja and P.C. Reghu Raj, "Fuzzy Logic Based Sentiment Analysis of Product Review Documents", *Proceedings of 1st International Conference on Computational Systems and Communications*, pp. 18-22, 2014.
- [9] E.E. Luneva, P.I. Banokin, A.A. Yefremov and T. Tiropanis, "Method of Evaluation of Social Network User Sentiments Based on Fuzzy Logic", *International Journal of Key Engineering Materials*, Vol. 685, pp. 847-851, 2016.
- [10] Mauro Dragoni, Andrea G. B. Tettamanzi and Celia Da Costa Pereira, "Propagating and Aggregating Fuzzy Polarities for Concept-Level Sentiment Analysis", *Cognitive Computation*, Vol. 7, No. 2, pp. 186-197, 2015.
- [11] V. Reshma and Anamma John, "Aspect Based Summarization of Reviews Using Naive Bayesian Classifier and Fuzzy Logic", *Proceedings of IEEE International Conference on Control, Communication and Computing*, pp. 617-621, 2015.
- [12] Jorge Roperio and Ariel Gomez, "A Fuzzy Logic Intelligent Agent for Information Extraction Introducing a New Fuzzy Logic-based Term Weighting Scheme", *Expert Systems with Applications*, Vol. 39, No. 4, pp. 4567-458, 2012.
- [13] B.N. Supriya, V. Kallimani, S. Prakash and C.B. Akki, "Twitter Sentiment Analysis Using Binary Classification Technique", *Proceedings of International Conference on Nature of Computation and Communication*, pp. 391-396, 2016.

- [14] "Hybrid Controller based Intelligent Speed Control of Induction Motor", Vinod Kumar, R.R. Joshi, <http://jatit.org/volumes/research-papers/Intelligent Speed Control 3 1.pdf>
- [15] "Fuzzy Logic Systems", James Vernon, <http://www.control-systemsprinciples.co.uk/whitepapers/fuzzy-logic-systems.pdf> [6] "Cahier Technique n0 191, Fuzzy Logic", F. Chevie, F. Guey, <http://www2.schneiderelectric.com/documents/technical-publications/en/shared/automation/automation-informationnetworks/ect191.pdf>
- [16] "Sentiment Analysis of Twitter Data", Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, Rebecca Passonneau, 23 June, 2011
- [17] "Twitter as a Corpus for Sentiment Analysis and Opinion Mining", Alexander Pak, Patrick Paroubek,
- [18] "SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining", Andrea Esuli, Fabrizio Sebastiani.