

# Emotion AI using Sectional Textual Repercussion

<sup>1</sup>A. Kumar, <sup>2</sup>K. Naveen, <sup>3</sup>Dr. B. Rama Subba Reddy

<sup>1,2</sup>Assistant Professor, Computer Science and Engineering, SV College of Engineering, Tirupati.

<sup>3</sup>Professor, Computer Science and Engineering, SV Engineering College, Tirupati.

**Abstract**—The internet is a puddle of raw data in the form of social multimedia or text messages, blogs, tweets, etc. - irregular and connectionless. The data becomes information on if it is pre-processed. A survey from the literature, we recognized that the present Emotional analysis tools are moderate, bulky and computationally heavy which makes the task at hand inefficient. Therefore, to overcome the aforementioned problem of analyzing sentiments efficiently, a new method is proposed in the present paper to drive the task of Sentiment Analysis by exploiting the idea of Partial Textual Entailment. We propose to use Partial Textual Entailment for measuring semantic similarity between the tweets so as to group similar tweets together. The method is anticipated to reduce the burden of sentiment analyzer and makes the processing faster. Moreover, we also propose a modification in an existing method of Partial Textual Entailment which can be further adopted for many Natural Language Processing applications.

**Keywords**— Sentiment Analysis, Textual Entailment, Partial Textual Entailment.

## I. INTRODUCTION

In today's era of the ongoing digital revolution, one can find a plethora of textual content on social websites, microblogging websites, etc. The bite-sized messages allow users to voice opinions, exchange views on current affairs and post reviews - positive or negative - of routinely used products/services. Now a days, corporations also view social websites, microblogging websites, etc. as an important channel to gauge public sentiment for their product/services. Many corporations have teams dedicated to addressing user reactions. This has opened up the domain of holistic technology-driven sentiment analysis.

Levy et al [1] describe sentiment analysis as the study of determining polarity of a sentiment. The task of Sentiment Analysis includes finding the emotion, attitude, view or opinion of a person towards another person, object, product or event [2][3] which has become useful in areas like academics, business and industry. These extracted emotions, views or opinions offer the capability to provide accurate views on various applications in different fields. Social media sites like Facebook or Twitter have become a popular channel to perform analysis on shared opinion. Therefore, it can be said that Sentiment Analysis identifies the sentiment or feelings of a person in a domain. It helps to review a person's satisfaction or dis-satisfaction whereas Opinion Mining[4] is used to identify opinion or views of a person in a domain. It helps to review a person's satisfaction or dis-satisfaction. The task of Sentiment Analysis can be classified on the basis of different classes - it can be analyzed as a positive, negative or neutral

sentiment based on classes [5].

- *Positive Sentiment Analysis*: If the document expresses positive views about an entity, the polarity of the document is said to be positive. Polarity can also be determined for sentences or aspects of sentences. For example: The fragrance of the flowers is good. In this sentence "fragrance" can be considered as an aspect of the "flower" and "good" indicates the positive view about the fragrance of the flower.
- *Negative Sentiment Analysis*: If the document expresses negative views about an entity, the polarity of the document is said to be negative. For example: The toppings on the pizza were rotten. In this sentence the "toppings" are considered as an aspect of the pizza and "rotten" indicates the negative view about the toppings of the pizza.
- *Neutral Sentiment Analysis*: If the sentence or document is unable to give a positive or a negative view, its polarity is said to be neutral. For example: He went to buy some products from the market.

Sentiment Analysis can be further classified on the basis of different granularity of text. It can be classified at document level, sentence level or aspect level [5] which is briefly discussed below

- *Document Level Sentiment Analysis*: It analyses the content of the document as a whole for finding its polarity.
- *Sentence Level Sentiment Analysis*: It takes into account different sentences of a document by further breaking them into smaller chunks of words for analyzing its polarity.
- *Word Level Sentiment Analysis*: It determines the polarity of different words of a sentence with relation to some object or event.

With the ever increasing number of social media posts and tweets, it has become a challenging issue to accomplish the task of analyzing all reviews within reasonable time lines. It has become extremely important to categorize similar views so as to reduce computation burden on the organization. In order to reduce the above said issues, this paper proposes the use of Textual Entailment in the process of Sentiment Analysis. Textual Entailment is defined [6] as a directional relationship between two text snippets denoted as Text (T) and Hypothesis (H) It is a process of checking whether H can be concluded from T or not.

Through the Literature survey, the authors have found that TE has been used for redundancy detection by Lloret et al. [7]. This has motivated the authors in this paper to use it in task of Sentiment Analysis as well. TE is also considered as a notion that is used to check semantic inference between text expressions using syntax-based, semantic-based and hybrid methods [8]. The task of TE has been carried forward using various approaches like the bag of words, vectors and directional tree methods.

During the last few years TE has become a popular topic of interest in the area of Natural Language Processing (NLP) including student response analysis, document summarization, machine translation evaluation, paraphrase recognition etc. [9]. Ideally, the authors are not able to accurately measure the inference using TE because of many reasons like the complexity of sentences and limitations of existing linguistic resources. Therefore, a need of Partial Textual Entailment (PTE) has been felt that measures the level of entailment whether the text fully entails the hypothesis or not. The authors plan to employ Partial Textual Entailment (PTE) to improve the task of Sentiment Analysis in this paper. It is expected to reduce the amount of work done in analyzing the sentiment of text thereby providing a better solution for the task of SA.

The following paper is outlined as follows: the related work in the field of Sentiment Analysis and Textual Entailment has been discussed in Section 2, Partial Textual Entailment is introduced and discussed in Section 3, the proposed methodology for Sentiment Analysis is proposed in Section 4, conclusion and future work are laid down in Section 5.

## II. RELATED WORK

Sentiment Analysis has been performed over the years and different models have come into existence. Classifiers using machine learning have been developed to generate polarity of a sentiment on the document level [10]. An algorithm to analyze the semantic arrangement of phrases using Point-wise Mutual Information (PMI) score has been developed[11]. A number of techniques have been used and have evolved over time for the sentiment analysis on data, video and audio [12][13]. NLP based techniques have been used for customer reviews to provide feature-based summaries [14]. NLP techniques have been used to mine the opinion for a product from the web and generate the polarity of expression [15]. A machine learning algorithm has been proposed to calculate polarity of expression using sentiment analysis on phrase level [16]. An unsupervised system has been developed for the extraction of feature based summaries for customer reviews [14][17]. PMI score has been calculated to conclude if a noun can be used as a feature of a product review. The two scoring methods Latent Semantic Analysis (LSA) and PMI are tested using two different corpora to conclude LSA approach to be more accurate in semantic orientation classification[18].

In this paper the authors propose a framework for determination of the polarity of text using the flexibility of Partial Textual Entailment (PTE). The authors suggest the modification over the existing facets generation method in the proposed framework to reduce computations in Sentiment Analysis..

The next section will discuss limitation of Textual Entailment and introduce Partial Textual Entailment.

## III. EMPLOYING PARTIAL TEXTUAL ENTAILMENT

Recognizing Textual Entailment (RTE) is a binary decision problem i.e. the process indicates if Text (T) entails Hypothesis (H) or not. The process can be said to be rigid and crisp [19] as the process is unable to determine the level at which T entails H [20].

To overcome the aforementioned limitations, Nielsen et al.

[1] introduced the concept of Partial Textual Entailment (PTE) using facets. The paper reports that the concept of PTE has achieved significant improvement over baseline methods like bag of words. The idea has been further enhanced by Levy et al. [1] by using faceted PTE. In the following work, the ordered pair facets are created that are free from having semantic relationship and thereby, achieving better results.

According to Levy et al. [1] Partial Textual Entailment solved the problem of negative Textual Entailment. Text (T) partially entails Hypothesis (H) indicated that a part of hypothesis is entailed by text. For Example:

- a) T: *The Taj Mahal was constructed by Shahjahan.*
- b) H: *Shahjahan constructed the Taj Mahal*
- c) T: *Shahjahan constructed the Taj Mahal in Agra near the Yamuna river*
- d) H: *Shahjahan constructed the Taj Mahal*

Sentence (c) and (d) are an example of Partial Textual Entailment as some of the information like Agra, Yamuna River is missing. Next we are going to introduce Partial Textual Entailment and propose a method to carry out the task of recognizing entailment between text using Partial Textual Entailment.

## INTRODUCTION TO PARTIAL TEXTUAL ENTAILMENT

The foundation of our work is based on recognizing Partial Textual Entailment. To develop a system for the same, it is required to introduce Faceted Partial Textual Entailment [1]. The authors are going to propose a fully automatic system of Faceted Partial Textual Entailment.

The concept of Partial Textual Entailment has been proposed by Nielsen et al. [21]. The concept has been further elaborated by Levy et al. [1] in the form of faceted Partial Textual Entailment[1]. The concept has been proposed as a solution to the rigidity and crispness of the task of Textual Entailment.

T: The main job of teachers is to educate students  
H: The teacher gives lectures to the students

A relationship can be seen in this pair of the sentences between the text(T) and the hypothesis(H) but to say there is entailment between the two would not be correct. It can be said that a relationship still exists between the two and is termed as Partial Textual Entailment i.e. a part of the hypothesis is entailed by the text.

Neilsen et al. [21] have defined the idea of facets that is in the form of ordered pair  $w_1, w_2$  where  $w_1$  and  $w_2$  represent words that belong to nouns, pronouns, verbs or adjectives. The ordered pair have a semantic relationship between each other and also belong to the hypothesis. Levy et al. [1] have extended the idea by using faceted Partial Textual Entailment - by considering the ordered pair without any semantic relationship between each other. Levy et al. [1], has suggested the following three stage framework for PTE.

*A. Decompose the Hypothesis into Facets*

The facets are generated in the form of ordered pairs from the hypothesis.

*B. Determine whether each Facet is Entailed*

The facets of the hypothesis are matched with the facets of the text to determine entailment between them.

*C. Aggregate the Individual Facet Results and Decide on Complete Entailment Accordingly*

The aggregated outcome of absolute matching of facets of hypothesis and text results in complete entailment.

The notion of faceted Partial Textual Entailment has been addressed by using the exact match method, the lexical inference method and the semantic method in literature survey [7]. Next, the authors briefly discuss these approaches.

**Exact Match Method:** All the lemma or tokens of H are contained in the bag of words. If the lemma or tokens of T are also present in the bag of words, then the decision is said to be positive. A positive decision implies T entails H. This approach has been commonly used in a number of Textual Entailment challenges.

**Lexical Inference Method:** The similarity between the ordered pair of T and H is estimated and measured using the Resnik Similarity Measure. A lexical based system has been represented as a model based on word overlap between the text and the hypothesis [22]. The authors have reported that the system has not been efficient and a web-based word similarity method has been given that used lexical edit distance for comparison [23].

**Semantic Inference Method:** Semantic entailment is

the problem of determining if the meaning of a given sentence entails that of another [24]. BIUTEE is a commonly used tool for recognizing entailment [6] where the dependency tree of facets for T and H is generated. The cost of generating H from T using a path between the ordered pair determines entailment between the two.

Using these methods and their different configurations, our methods will be checked for results.

PTE uses facets generation by creating pairs of all the words. This leads to a high number of faceted pairs and computations. The existing methods faced this major limitation. To overcome the following issues a new idea is proposed which will eventually be turned into a working method:

1) The existing methods generate a large number of facets that increases the amount of computation effort. This motivated the authors to the propose solution of using facets.

2) The existing methods of sentiment analysis require a lot of computation to generate a sentiment because the reviews are increasing daily. The integration of heavy linguistic resources like entity recognition, stemmers etc make processing and analysis of text/expressions further slower. To overcome the problem, SA is done using PTE.

The main aim of the present paper is:

1) To propose a modification over the existing method for the creation of facets in Partial Textual Entailment. This is done over the existing method of facets creation by Levy et al.

[1] so that the number of facets can be reduced to provide efficient results and reduced computation.

2) With the reduced number of text/expressions as a result of Partial Textual Entailment, the sentiment analyzer will be provided with fewer expressions. The burden of Sentiment Analyzer is reduced to the analysis of a small set of text instead of the original large set of expressions. This reduces the processing time considerably.

3) Lastly, the total percentage of each kind of sentiment can be determined.

To the best of our knowledge, Partial Textual Entailment has not been used anywhere for sentiment analysis till now. Therefore, the idea is:

*a) To implement the task of faceted Partial Textual Entailment.*

*b) To use faceted Partial Textual Entailment for grouping tweets with similar polarity.*

*c) To apply a theoretical method for sentiment analysis on the tweets extracted from Partial Textual Entailment.*

*d) The method proposed will be further checked with the existing methods.*

The next section discusses the proposed modification for faceted Partial Textual Entailment. It discusses the steps for generation of facets and the process of faceted PTE.

IV. PROPOSED METHODOLOGY FOR SENTIMENT ANALYSIS The authors propose a method to carry out the task of

Sentiment Analysis that has been broken into two subtasks: (a) developing a method for recognizing PTE and (b) developing the method of Sentiment Analysis using Partial Textual Entailment.

A. Developing the method of Recognizing Partial Textual Entailment

The proposed idea is to develop a model to reduce computation due to a large number of facets construction and to generate efficient results. This will be done by removing helping words. The facets will only be constructed using the subject of the sentence as a primary word in the creation of facets. This method is expected to predict better Textual Entailment as the text revolves mainly around the subject of the sentence. Consider,

S: Sentence

X: Set of words in the sentence

S Q: Subjects of sentences

R: Attributes of the sentence

1) Selecting Text and Hypothesis:

The text and hypothesis have to be selected from the reviews. For Example,

Text: Rahul goes for jogging daily. He brought some eggs today.

Hypothesis: Rahul takes good care of his health.

2) Refinement of Text and Hypothesis by removing helping words like "of", "some", "for".

The helping words like stop words, prepositions are removed. Text and hypothesis are modified as

X (T): {Rahul, goes, Jogging, daily, He, brought, eggs, today}

X (H): {Rahul, takes, good, care, his, health}

3) Identification of the Subject where Q represents the subject and R represents the remaining words

The subjects from the sentences are selected and the remaining words are listed down separately.

Q (T): {Rahul, eggs}

R (T): {goes, Jogging, daily, He, brought, to-day}

Q (H): {Rahul}

R (H): {takes, good, care, his, health}

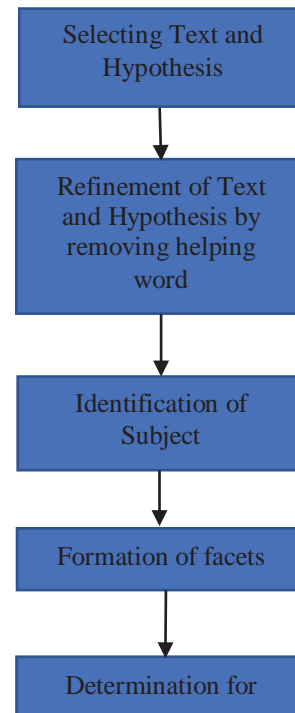


Fig. 1. Block Diagram of Faceted Partial Textual Entailment

4) Formation of facets:

The facets are created by pairing remaining words with the subjects as follows:

Facet (T): Q\*R = ({Rahul, goes}, {Rahul, Jogging}, {eggs, brought} ... Total 12 facets

Facet (H): Q\*R = ({Rahul, takes}, {Rahul, good}, {Rahul, care} ... Total 5 facets

5) Determination of Partial Textual Entailment will be done:

In the last step, the authors have observed that the number of facets gets reduced significantly. The total number of facets generated are equal to 60(12\*5 = 60). This is going to reduce the number of comparisons and thus reduce the computation required. The whole process will be applied to the tweets and the result will be countable non-similar sentences which can be easily passed to the sentiment analyzer to determine the sentiment of the collection of tweets, reviews or expressions. In addition, the number of text inputs relating to a sentiment will determine how strongly people agree or disagree to a topic. The next section discusses the flow of work i.e. an overall architecture of Sentiment Analysis using Partial Textual Entailment.

B. Developing a method for Sentiment Analysis using Partial Textual Entailment

The authors discuss the overall framework for SA using PTE below

1) Recognizing new models or methods for Partial Textual Entailment:

As discussed in Section 4.1, this step will perform the

task of grouping similar tweets so as to reduce the burden of the Sentiment Analyzer resulting in faster processing of reviews. The entailment recognition will be done using BIUTEE tool [6].

2) The result obtained from Step 1 will be passed to the sentiment analyzer to determine the sentiment of the text/expressions.

3) The total number of sentences that entail the hypothesis will then be provided as an estimator of how strong the sentiment is.

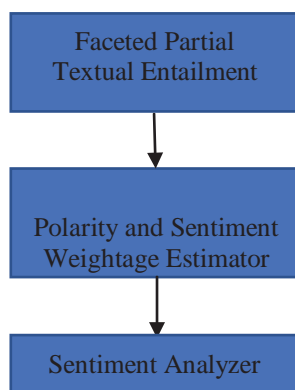


Fig. 2. Overall Architecture of Sentiment Analysis using Faceted Partial Textual Entailment example of a figure caption.

## V. CONCLUSION AND FUTURE WORK

In the present paper the authors have proposed a novel method of Sentiment Analysis using Partial Textual Entailment. To the best of our knowledge, it has been used for the first time in this paper for reducing the computational overhead. The authors plan to implement the proposed method of Partial Textual Entailment which is discussed in Section

IV. In future, the authors also plan to implement method of Sentiment Analysis using the proposed method of Partial Textual Entailment. The authors also plan to use the standard data sets to compute the efficacy of their proposed method in future. The authors plan to use the BIUTEE tool for the task of entailment recognition. The authors also intend to extend this work for different regional languages of the country.

## REFERENCES

- [1] O. Levy, T. Zesch, I. Dagan, I. Gurevych. "Recognizing Partial Textual Entailment" Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Vol 2, pp. 451-455, 2013
- [2] <https://www.lexalytics.com/technology/sentiment>
- [3] H. Kaur, V. Mangat, Nidhi "A survey of sentiment analysis techniques" I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), International Conference on IEEE, 2017.
- [4] I. Guellil, K. Boukhalfa, "Social big data mining: A survey focused on opinion mining and sentiments analysis" Programming and Systems (ISPS), 12th International Symposium on IEEE, 2015
- [5] K. Mouthami, K. Nirmala Devi, V. Murali Bhaskaran. "Sentiment analysis and classification based on textual reviews" Information communication and embedded systems (ICICES), international conference on IEEE, 2013
- [6] A. Stern, and I. Dagan "The BIUTEE research platform for transformation-based Textual Entailment recognition" LiLT (Linguistic Issues in Language Technology), Vol 9, 2014.
- [7] E. Lloret, O. Ferrández, R. Muñoz and M. Palomar, "A Text Summarization Approach under the Influence of Textual Entailment" Natural Language Processing and Information Systems Group, pp: 22- 31, June 2008
- [8] M H. Haggag, Marwa M.A. ELFattah, A M. Ahmed "Different Models and Approaches of Textual Entailment Recognition" International Journal of Computer Applications. Vol 142(1), 2016.
- [9] I. Dagan, B. Dolan, B. Magnini And D. Roth, "Recognizing Textual Entailment: Rational, evaluation and approaches-erratum" Natural Language Engineering, Vol 16(1) pp: 105-105, 2010
- [10] B. Pang, L. Lee, S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques" In Proceedings of the ACL-02 conference on Empirical methods in natural language processing, Vol 10 pp:79-86, 2004.
- [11] P. D. Turney, "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews" In Proceedings of the 40th annual meeting on association for computational linguistics, pp: 417-424.
- [12] E. Cambria, B. Schuller, Y. Xia, Catherine, "New avenues in opinion mining and sentiment analysis" IEEE Intelligent Systems, Vol 28(2) pp: 15-21, 2013.
- [13] M. Wöllmer, F. Wening, T. Knaup, and B. Schuller, C. Sun, K. Sagae and L-P Morency. 2013. YouTube Movie
- [14] M. Hu and B. Liu, "Mining and summarizing customer reviews" In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp: 168-177, 2004.
- [15] K. Dave, S. Lawrence, David, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews" In Proceedings of the 12th international conference on World Wide Web, pp: 519-528, 2003.
- [16] T. Wilson, J. Wiebe, P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis" In Proceedings of the conference on human language technology and empirical methods in natural language processing, pp: 347-354.
- [17] A-M Popescu, O. Etzioni, "Extracting product features and opinions from reviews" In Natural language processing and text mining, Springer, pp: 9-28.
- [18] P. D. Turney, M. L. Littman, "Measuring praise and criticism: Inference of semantic orientation from association" ACM Transactions on Information Systems (TOIS), Vol 21(4) pp: 315-346.
- [19] M. Vita. "Computing Semantic Textual Similarity Based on Partial Textual Entailment".
- [20] <http://ama.liglab.fr/GreenSeminars/2015/Speechs/Diana.pdf>
- [21] R. D Nielsen, W. Ward, and J. H Martin, "Recognizing entailment in intelligent tutoring systems". Natural Language Engineering, Vol 15(4) pp:479-501, 2009
- [22] D. Majumdar, P. Bhattacharyya, "Lexical Based Text Entailment System for Main Task of RTE6" Text Analysis Conference (TAC) Proceedings, 2010.
- [23] R. Adams "Textual Entailment through extended lexical overlap" In Proceedings of the Second PASCAL Challenges Workshop on Recognising Textual Entailment, pp: 128-133, 2006.
- [24] R. S. Braz, V. Punyakanok, D. Roth, M. Sammons. "An inference model for semantic entailment in natural language" Machine Learning Challenges. Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Textual Entailment. Springer, pp: 261- 286, 2006.