Proceedings of the International Conference on Inventive Computing and Informatics (ICICI 2017)
IEEE Xplore Compliant - Part Number: CFP17L34-ART, ISBN: 978-1-5386-4031-9

# *TEXT BASED SENTIMENT ANALYSIS*

Biswarup Nandi[*1], Mousumi Ghanti[*2], Souvik Paul[#3]
[*]*Department of Computer Science and Engineering ,University of Calcutta*
*Kolkata,WB,India*
[1]biswarup1010@hotmail.com
[2]mousumic828@gmail.com
[#]*The Heritage Academy*
*Kolkata,WB,India*
[3]souvikpaul.tiu@gmail.com

***Abstract*— One of the most important parts of running business successfully is analysing customer's opinion and sentiments[1]. In this paper, the paragraph of sentences given by the customer is accepted and after extracting each and every word, they are checked with the stored (database has been maintained here) parts of speech, articles and negative words. After checking against the database, CFG is used to validate proper formation of the sentences. Each sentences are delimited by '.' or '?' or '!'.Emotions[2] are detected as – positive, negative or neutral sentence. There are 3 types of cases- 1. If the paragraph contains more positive sentences than negative, then overall result will be positive. 2. If the number of negative sentence is greater than positive sentence, then the overall result is negative. 3. If there are same numbers of positive and negative sentences in the input paragraph, then the result is neutral and if a sentence has been entered that is a normal statement neither positive nor negative, that will be also considered as neutral.**

***Keywords—Sentiments; CFG; Database; Positive; Negative; Neutral;***

## I. INTRODUCTION

Sentiment analysis is also known as opinion mining. It is extremely useful in business fields, social media etc. When a company sells a product, it is very important to know how customers react about the product for further progress or growth [3] of their business. Not only in case of business, we often see people give opinions as review[4] of movies which is also very important for a film maker to know audiences' responses. Now-a-days the social media such as different types of blogs or various social networks has increased the interest for sentiment analysis. Several research works are now going on sentiment analysis to identify human attitude, reactions or opinions[5].

A sentence is a collection of meaningful words that contains different types of parts of speech (e.g. noun, pronoun, adjective, adverb, verb, preposition, conjunction, interjection) and articles (a, an, the).

A sentence can be formed in various way like Noun+Verb+Not+Adverb+Adjective or Noun+Verb+Adverb+Adjective etc. In this project we have considered few of those and implemented but each and every possible combination can be implemented in this way. For better understanding a transition diagram(Fig 1) is given below (CFG starting with 'Noun' has been considered in the diagram given below ) –

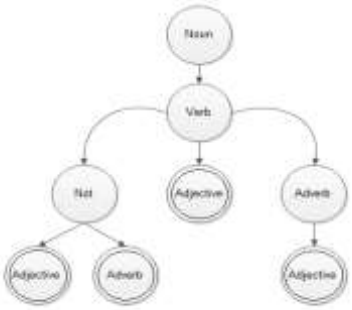


**Fig 1: State Transition Diagram**


**Fig 2: Block Diagram**

The above diagram(Fig 2) is describing the whole process with help of blocks.

## II. RELATED WORK

Walaa Medhat [a,*] Ahmed Hassan [b], Hoda Korashy [b] in their paper discussed about their survey where the main target is to provide a brief idea about SA techniques and some of its related fields. They categorized a large number of articles on sentiment analysis and mentioned those in their paper.

G.Vinodhini*, RM.Chandrasekaran in their survey paper discussed about various techniques of data mining and some of its challenges. They have given some idea about machine learning, role of negation ,semantic orientation etc. Some applications and tools are also discussed there.

Svetlana Kiritchenko ,Xiaodan Zhu ,Saif M. Mohammad in their paper discussed about sentiment analysis of small texts which are informal. They used 2 test suits- one was with tweets and another was with SMS. They collected tweets from the public streaming twitter API and worked on it. The messages those which are not containing words with polarity, they cut those down for simplicity.

## III. PROPOSED WORK

The whole project work has been done in Windows form application using the language 'C#'. User will be asked to enter their opinion or reaction in form of sentences and analyzing those sentences, an output will be generated as positive or negative or neutral. There is an important concept of CFG with which we are validating the proper formation of sentences. After getting the sentences with correct grammar, the counting of polarity[6] (positive, negative or neutral) is started. The algorithm of the process is given below—

### *A. ALGORITHM:*

**Step 1:** Accept a paragraph.
**Step 2:** Sentences are extracted from the paragraph and stored into an array.
**Step 3:** Start a loop for each sentence of the sentence array—
**Step 3.1:** Check if sentence is not NULL, go to step 3.2 else go to step 3.9
**Step 3.2:** Check if the last character of the sentence is '?' or '!' or '.' Go to step 3.3 else go to step 3.8
**Step 3.3:** Spelling check is done with. If the result of spell check of sentence is right go to step 3.4 else go to step 3.7
**Step 3.4**: CFG checking is done (explained later) and if CFG is correct go to step 3.5 else go to step 3.6
**Step 3.5:** Sentence emotion is checked (explained later) in this step. If emotion is positive, then positive_count← positive_count + 1 and if emotion is negative, then negative_count←negative_count+1
**Step 3.6:** Wrong sentence format is detected in this step
**Step 3.7:** In this step it is shown that Spelling of sentence is wrong
**Step 3.8:** This step shows that punctuation is missed.
**Step 3.9:** An error message is generated as the program gets input as NULL sentence.
**Step 4:** If the total count of positive emotion is greater than negative emotion, the result is positive else go to step 5
**Step 5:** If the total count of positive emotion is less than negative emotion, the result is negative else it is neutral.

### *B. EXPLAINED ALGORITHM OF CFG:*

**Step 1:** Extract words from the sentence and store into an array.
**Step 2:** All the words are checked against the stored parts of speech—if that is noun, or pronoun or verb or adjective or article or any negative word like 'not'.
**Step 3**: If the word array length is greater than 5, Check if the order is 'Article+Noun+Verb+Not+Adverb+Adjective' CFG is correct else go to step 3.1
**Step 3.1** If the word array length is greater than 4, Check if the order is 'Noun +Verb +Not + Adverb+Adjective' or 'Pronoun +Verb+ Not+Adverb+Adjective','Article+Noun+Verb+Not+Adjective' or 'Article+Noun +Verb + Adverb +Adjective' CFG is correct else go to step 3.2
**Step 3.2** If the word array length is greater than 3, Check if the order is 'Noun+Verb +Not + Adjective' or 'Noun+ Verb+ Adverb+ Adjective' or 'Pronoun+Verb+Not+Adjective' or 'Pronoun+Verb+Adverb+Adjective' or 'Article+Noun+ Verb+Adjective' CFG is correct else go to step 3.3
**Step 3.3** If the word array length is greater than 2, Check if the order is 'Noun+ Verb+ Adjective' or 'Pronoun+Verb+Adjective' CFG is correct.

### *C. EXPLAINED ALGORITHM OF SENTENCE EMOTION:*

**Step 1:** The extracted words from the sentence are checked against positive and negative words (Stored in database).
**Step 2:** If the words are matched with stored words--
**Step 2.1:** If matched with stored negative words, negative_count <negative_count+1 , else go to step 2.2
**Step 2.2:** If matched with stored positive words, positive_count <- positive_count+1, else go to step 2.3
**Step 2.3:** That means it is neutral.

The whole architecture and processing technique can be explained clearly by flowchart. Here we have designed flowchart in three parts, where the first one (Fig 3) represents the whole technique and other two (Fig 4 & 5) are given to show detailed method of sub modules. The flowcharts of the algorithm are given below—

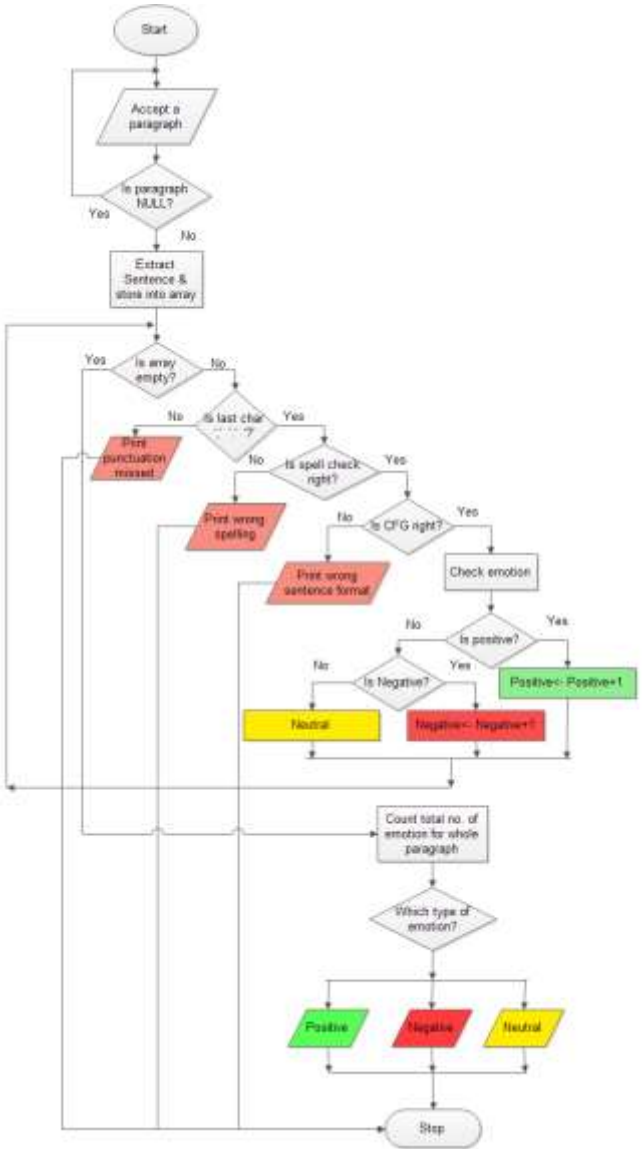


**Fig 3: Flowchart of the whole process**

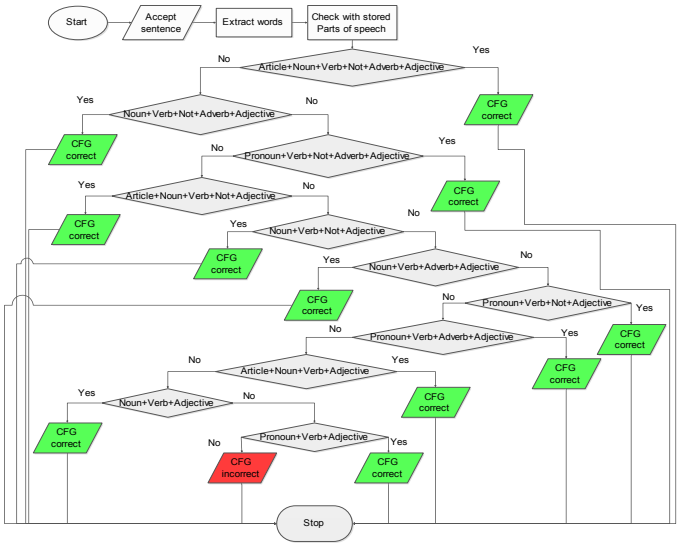


**Fig 4: Flowchart of CFG checking**

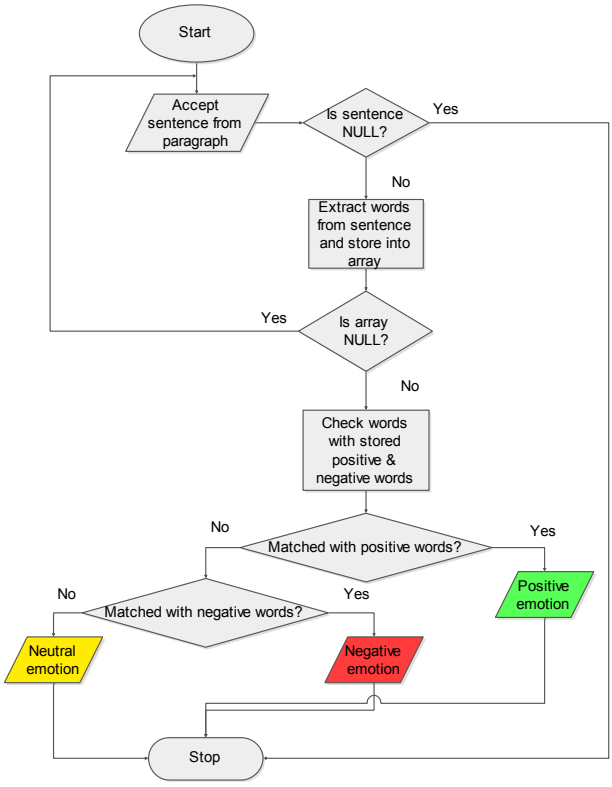


**Fig 5: Flowchart of emotion checker**

DFD (Data Flow Diagram) is the technique to explain each and every module with the small details so that one can implement the method easily. The DFD of our proposed algorithm is given below—

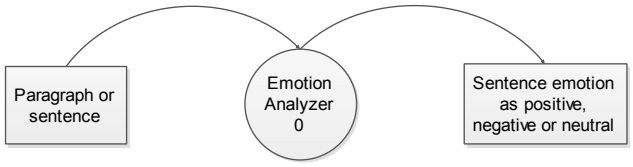


**Fig 6: Level 0 DFD**

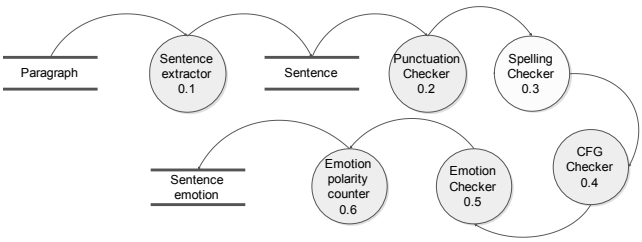


**Fig 7: Level 1 DFD**

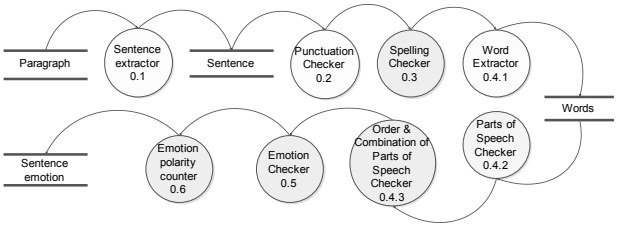


**Fig 8: Level 2 DFD**

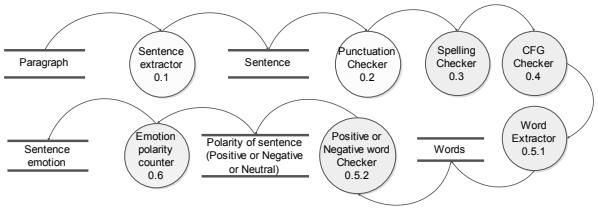


**Fig 9: Level 3 DFD**

## IV. RESULT ANALYSIS

The user have to enter a sentence or a paragraph. The software will then generate a result detecting [7] the emotion which can be either positive or negative or neutral. The three possible are given below--

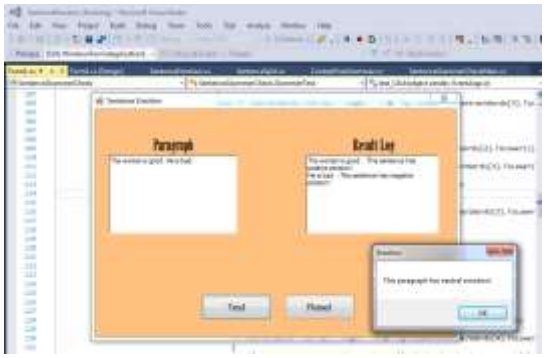

**Fig 10: Result with neutral emotion**

The above snapshot(Fig 10) is output with neutral emotion.

**Fig 11: Result with negative emotion**

The above snapshot(Fig 11) is the output with negative emotion.

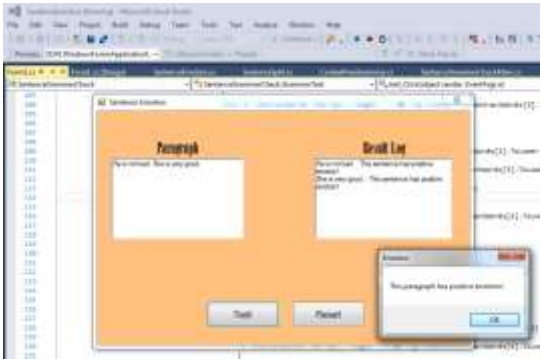

**Fig 12: Result with positive emotion**

The above snapshot(Fig 12) is the output with positive emotion.

## V. CONCLUSION

The topic of this project paper has a vast amount of future scope. Sentiment analysis is very essential for all types of business. Several researches on this topic are going on and it will be really helpful if software is made whose function will be to detect human's sentiment or opinion to run businesses, blogs, and social sites smoothly. We have implemented some of CFGs but there are many CFGs that can be implemented to make the software more realistic and powerful.

In today's world, we need to know people's reaction[8] in every business type works. Without people's opinion we cannot determine whether there is any need for change. Sentiment analysis is required in this space and in this project; we have implemented text based sentence analysis process to know the polarity of the sentences i.e. whether the review given by people is positive or negative or neutral in nature.

## REFERENCE

[1] G.Vinodhini*, RM.Chandrasekaran 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. Page-285

[2] G.Vinodhini*, RM.Chandrasekaran 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. Page-283.

[3] Svetlana Kiritchenko ,Xiaodan Zhu ,Saif M. Mohammad Sentiment Analysis of Short Informal Texts. Journal of Artificial Intelligence Research 50 (2014) 723-762.
Page no-723.

[4] Svetlana Kiritchenko ,Xiaodan Zhu ,Saif M. Mohammad Sentiment Analysis of Short Informal Texts. Journal of Artificial Intelligence Research 50 (2014) 723-762.
Page no-725.

[5] Walaa Medhat [a,*] Ahmed Hassan [b], Hoda Korashy [b] Sentiment analysis algorithms and applications:A survey . Ain Shams Engineering Journal.
Page no-1094.

[6] G.Vinodhini*, RM.Chandrasekaran 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. Page-286.

[7] Svetlana Kiritchenko ,Xiaodan Zhu ,Saif M. Mohammad Sentiment Analysis of Short Informal Texts. Journal of Artificial Intelligence Research 50 (2014) 723-762.
Page no-727.

[8] Walaa Medhat [a,*] Ahmed Hassan [b], Hoda Korashy [b] Sentiment analysis algorithms and applications:A survey . Ain Shams Engineering Journal.
Page no-1098.