Using Natural Language Processing to Improve Customer Experience

Dr. R.Balu
Faculty, Department of Computer Applications
School of Computer Science and Engineering, Bharathiar University, Coimbatore – 641 046
Email: rvkbalu@yahoo.co.in
+91 98420 12361

Abstract — ***An innumerable of customer service channels exist today, such as social media, email, chat services, call centers, and voice mail. There are so many ways that a customer can interact with a business and it is important to take them all into account. The aim of this research is to understand the sentiments of the feedbacks or opinion of the general public that they provide on the social media or one feedback forums about a particular topic. Organisations can take pre-emptive marketing decision based on these analysis in order to improve the performance of their product.***

Keywords — NLP, Sentiment Analysis, Customer Experience, Social Media, Big Data

I. INTRODUCTION

Natural Language Processing is an area of computer science and Artificial Intelligence concerned with the interactions between computers and Natural Languages in particular how to program computers to fruitfully process large amounts of natural language data [4].

The challenges in natural language processing involves speech recognition, natural language understanding and natural language generation. The development of NLP applications is challenging because computers traditionally require humans to "speak" to them in a programming language that is precise, unambiguous and highly structured, or through a limited number of clearly enunciated voice commands [2].

Human speech, however, is not always precise -- it is often ambiguous and the linguistic structure can depend on many complex variables, including slang, regional dialects and social context. Many of the notable early successes occurred in the field of machine translation. These systems were able to take advantage of existing multilingual textual corpora. However, most other systems depended on corpora specifically developed for the tasks implemented by these systems, which was (and often continues to be) a major limitation in the success of these systems. As a result, a great deal of research has gone into methods of more effectively learning from limited amounts of data [8].

With the amount of data that is available on the Social media and with the corporates it becomes very difficult to understand the pretext of the customer with regards to the product or the Service. Eventually these opinions are never even considered. The opinions are read through and not understood in terms of the changes in the requirement of products and services offered.

II. CUSTOMER EXPERIENCE

Today most of the companies have implemented customer experience programs in some form or another. Some programs may be even considered mature enough to understand the customer better beyond the simple listening of the customer and go to the extent of taking action to customer’s requirements and feedbacks taking action delivering insights across the organization [1].

In order to achieve this, there is one area which they need to resolve efficiently and systematically: how to deal with unstructured data and analyze text from various sources in a meaningful and measurable way. In most of the cases the brands mostly focus on structured data. It is also easy to analyze data from survey forums.

In order to understand the voice of the customer correctly, you need to listen to the voices in all forms. This data may be available in structured and unstructured form. In general customer experience can be captured from the following sources:

- Enterprise Feedback software – provided as part of the corporate CRM software
- Market Research agencies – appointed to capture customer feedback from the customers
- Social Media – such as twitter, Facebook, LinkedIn etc

All such sources of data need to be tapped to understand the customer better. A customer experience program along with a robust text analytics solution is critical to make sound business decisions distributed to the right people at the right time. The potential of unstructured data and translate it to business value day after day will show the value of a True Leader.

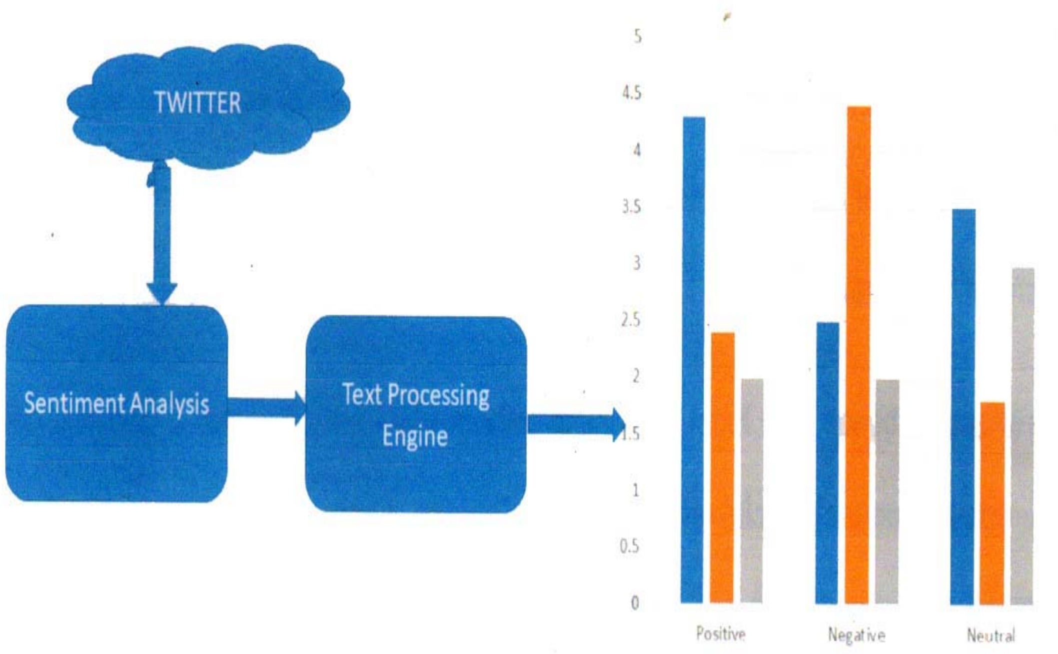


Fig. 2.1 Process of the data from the social media website

A. Interpretation of Data

Once the data is collected from both the solicited and unsolicited sources, the more challenging aspect of this whole process. The use of Natural Language processing (NLP) has proved to be more efficient, reliable, faster and easier to analyze the data and has proven to be cost efficient too. The use of text analytics can map 100% of the meaning of any textual content to reveal customer sentiments, concerns, intentions, and ideas, clearly and reliably [5].

Initially the analysis was made on a case to case basis using Word Cloud, Bubble chart and timeline analysis of the text analysed.

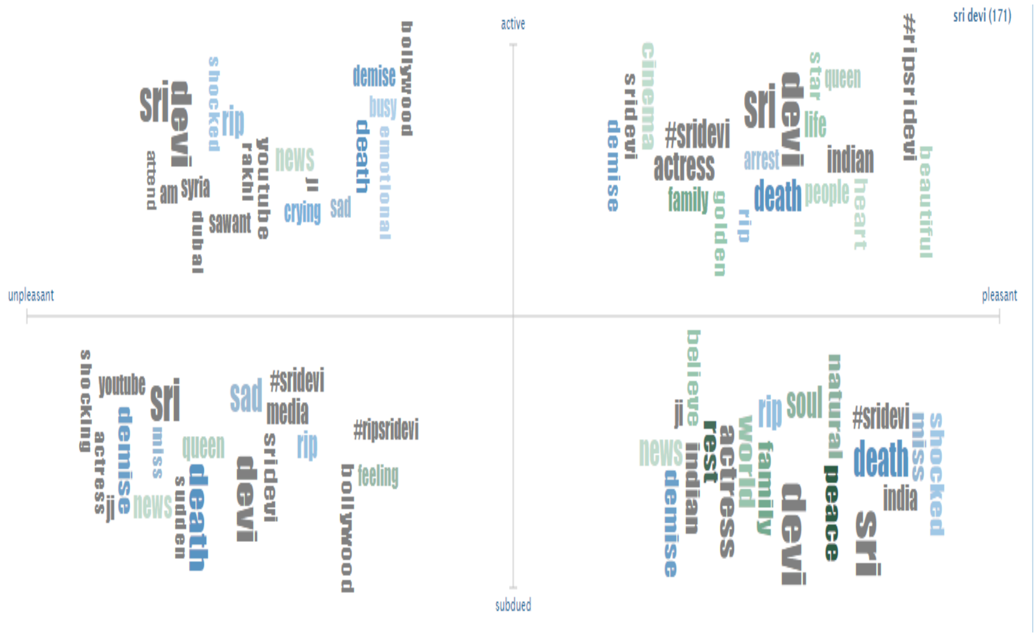


Fig. 2.2 Word Cloud

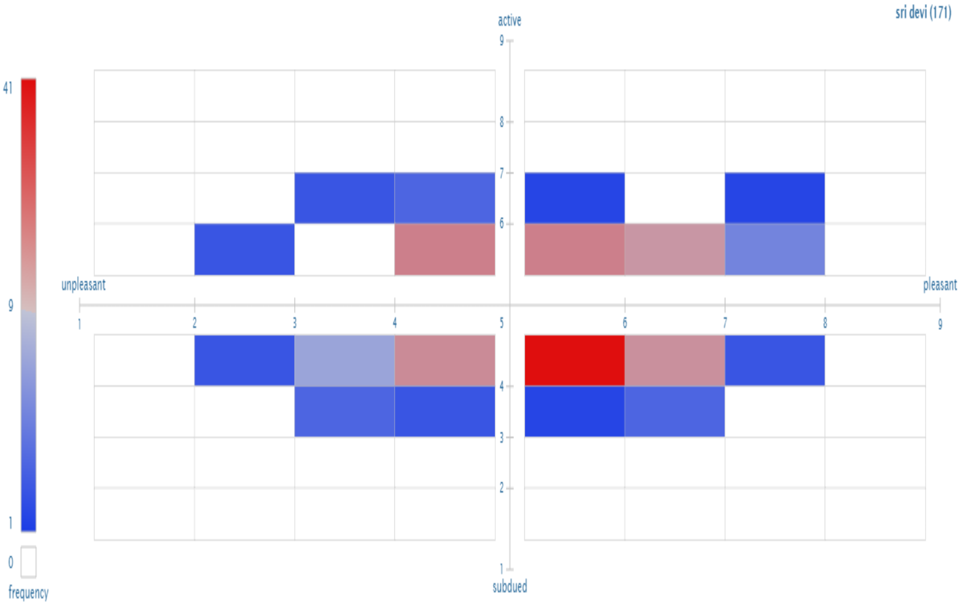


Fig. 2.3 Heat Map

III. Association of Text

In order to analyse text, the text has to be mined to find the text relations between text broadly classified as *paradigmatic relation* and *syntagmatic relation*.

The elements A and B are said to have paradigmatic relation if they can be substituted one for each other. This means that the two words are in the same semantic class, or syntactic class [6].

In general, they can be replaced one by the other without affecting the understanding of the sentence, which means that the result would still be a valid sentence. In the case of a syntagmatic relation, on the other hand, the two words can be combined with each other. Therefore, the elements A and B can be combined with each other in a sentence, since they are semantically related.

In another perspective, the relations can be seen as: (1) relations that occur in similar locations relative to the neighbors in the sequence (paradigmatic relation) or; (2) relations concerning co-occurrent elements that tend to show up in the same sequence (syntagmatic relation). These two basic relations of words are complimentary. For discovering paradigmatic relation, one can assume that words that have high context similarity also have paradigmatic relation. So, the context similarity of each word must be computed. To discover syntagmatic relation, one must search for words with high co-occurrences but relatively low individual occurrences. The justification is that those words tend to occur together. To compute the syntagmatic relation, one must count of how many times two words occur together in a context (a sentence, a paragraph, or even a document). Then a comparison should be made between the co-occurrences and their individual occurrences [6].

IV. Paradigmatic Relations

The idea of discovering paradigmatic relations is to look at the context of each word and try to compute the similarity of those contexts. This can be done in two steps: a. by formally representing the context and; b. by defining a similarity function. The context contains in general lot of words usually regarded as bag of words [3].

The similarity function is, in general, a combination of similarities on different contexts. Thinking in a bag of words is a useful mean for representation of vectors in a vector space model. The subjacent idea for this approach is to define each word in the vocabulary as one of the dimensions of the high dimensional space. Since there are N words in the vocabulary, then, there are also N dimensions in the space model. So, the context of a word, w1, can be represented as a vector d1, and a different word w2, by another context d2. The paradigmatic relation between the two words can then be measured computing the similarity of the two vectors. Therefore, by representing the context in the vector space model, the problem of paradigmatical relation discovery is converted into the problem of computing similarity of the vectors. For referring each vector we use the expression (1):

$$d1= (x1,...,xN) \text{ where each } xi \text{ is given by } \frac{xi=c(wi,d1)}{d1}.$$

where c(wi,d1)represents the total count of word w1in pseudo document d1and |d1|is the total amount of words in d1.

Regarding the computation of similarity, there are several approaches developed for information retrieval that can be adapted to text mining. One of the approaches, is to try to match the similarity of context based on the Expected Overlap of Words in Context (EOWC) method. The idea is to represent a context by a word vector where each word has a weight equal to the probability that a randomly picked word wi, from the document vector, is the specific word wi. In other words, xiis defined as the normalized count of word wiin the context, and this can be interpreted as the probability of randomly pick this word from the document dj. Since these are normalized frequencies, the sum of xi0sis one, which means the vector is the in fact a probability of words distribution.

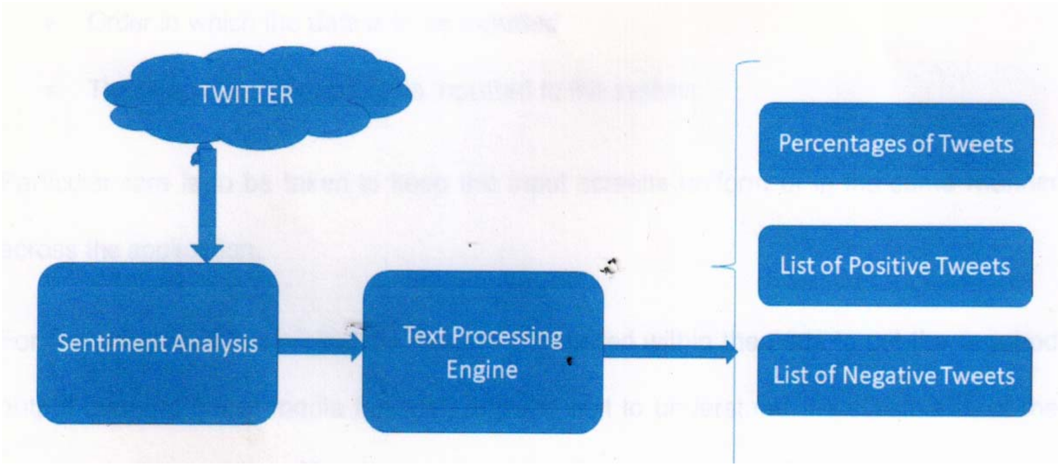


Fig. 4.1 Process of the data from the social media

According this method each context is represented by a vector that specifies the probability of each word in the context. Consequently, the similarity is defined (2) as the dot product of the two vectors

$$Sim(d1,d2) = d1 \cdot d2 = x1y1 + \ldots + xnyn = \sum_{i=1}^{N} x_i y_i$$

With this similarity function one can compute the probability of two randomly words from the two contexts being identical with EOWC approach measuring the overlap of words in the contexts. The EOWC method however, has two problems, namely: 1- it favors matching frequent terms over matching distinct terms; 2- it treats every word equally, meaning that even a common word would contribute equally as others more relevant to the current content. Retrieval heuristics, used in the text retrieval domain, can be used to solve these problems. To address the first problem, a sublinear transformation named Term Frequency (TF) is used, instead the raw frequency count of the terms, to represent the context. In the TF transformation, denoted by T F(w,d), the raw count of a word is converted into a weight that reflects the belief about the importance of the word. An implementation of the transformation in (3), called a BM25 transformation, was used in information retrieval to solve the same problem of overemphasizing a frequent word. This transformation

$$T\,F(w,d) = \frac{(k+1)x}{x+k}$$

where k∈[0,+∞[is a parameter and xis the raw count of a word, has an upper bound of k+1, which puts a constraint on high frequency. To solve the second problem, one must penalize popular terms and put more weight on rare terms. The heuristic (4) is used in text retrieval is called Inverse Document Frequency (IDF) term weighting. Document frequency means the count of the total number of documents that contain a particular word. The IDF measure is defined as a logarithm function of the document frequency

$$IDF(W) = \log \frac{(M+1)}{k}$$

where k is the document frequency and Mis the total number of documents in the collection. The IDF function gives a higher value for lower k, which means that it rewards a rare term. It reaches the maximum value on log(M+1), for a very rareterm that occurs just once in the context. The lowest value of IDF, close to zero, is when k reaches its maximum of M.TF and IDF heuristics are used to improve the similarity function for paradigmatic relation mining. The document vector is defined as containing elements rep-resenting normalized BM25 values. The new weight reflects now, the frequency of occurrence of the word in the context. In the document vector (1), each xi is now given by (5):

$$x_i = \frac{BM25(w_i, d_1)}{\sum_{j=1}^{N} BM25(w_j, d_1)}$$

The weight of each word is normalized by the sum of the weights of all the words. This ensures that all the xi's will sum to 1 in the vector that represents now the word distribution. The formula in (6) allows to define a document vector giving to the high frequency terms a lower weight. This helps to control the influence of the high frequency terms. So, in (5) the weight computed for each word x_i in document d_1.

RT @nramind: Dr V. Shanta, legendary cancer specialist, Chair of the Cancer Institute (W.I.A.) , Chennai, and Padma Vibhushan and Magsaysay...

RT @pankajsuper30: All 3 States of North East Showing Sign of Saffronization in #ExitPoll .... Best one will be kicking out Commies @cpimsp...

RT @pankajsuper30: Great Decision of @PMOIndia to handover #NPAs above Rs 50 Cr + of PSBs to CBI

RT @pushpendrakum: Punjab National Bank Rs 11,000-crore fraud? Why you are not doing legal bitcoin in india: https://t.co/DzeAkAkkw6 via @Y...

The recent fall has been triggered by the Punjab National Bank fraud, estimated to be worth over Rs 114 billion https://t.co/nEDaDxexMf

RT @ShashiTharoor: Among the best lines going around on the @PunjabNational Bank Financial Crisis:

"These days, Punjab National Bank's bal...

@Tata_Crucible Ans) Punjab National Bank @Indiapnb

#FromTheQuizmaster #NaviMumbai round using #TCCQ18 & #win

Fig. 4.2 Positive Tweets

```
Negative tweets:
-------------------
RT @GautamGambhir: A 1971 war hero Major Ravi Prakash Nayyar is down wid spine injury 4 last 7 yrs.He's admitted to Indian Spinal injuries…
RT @ANI: #PNBFraudCase: Last night, CBI arrested a General Manager (GM) rank officer of Punjab National Bank, Rajesh Jindal, who was the Br…
```

RT @GautamGambhir: A 1971 war hero Major Ravi Prakash Nayyar is down wid spine injury 4 last 7 yrs.He's admitted to Indian Spinal injuries...

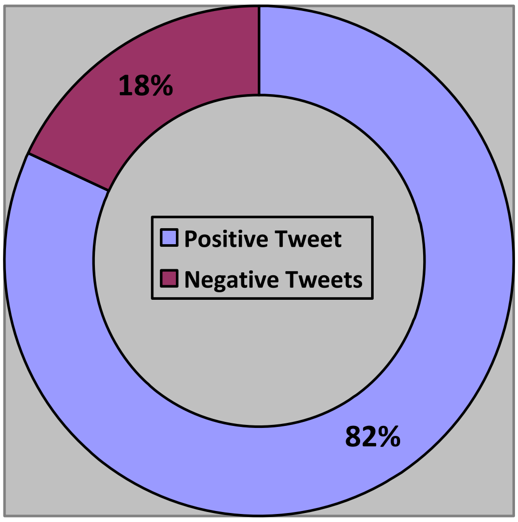
RT @ANI: #PNBFraudCase: Last night, CBI arrested a General Manager (GM) rank officer of Punjab National Bank, Rajesh Jindal, who was the Br...

RT @firstpost: The alleged fraud comes amid #CBI's probe into recent complaints of massive #frauds in the Punjab National Bank, Bank of Bar...

Fig. 4.3 Negative Tweets

```
F:\Projects\Python\SentimentAnalysis>python SentimentAnalysis.py
Positive tweets percentage: 37.5 %
Negative tweets percentage: 8.333333333333334 %
```

S.No.	Positive Tweets	Negative Tweets
1	37.5	8.3



V. CONCLUSION

The use of sentiment analysis is primarily used to understand the feelings of the customers towards the services or products offered to the customers. The completion of this research completely satisfies the requirements of understanding the sentiments of the customers.

The current systems can be expanded to a web based application based on the customer's needs. There by authors can use other third party tools to deliver the graphical representation for the research work. This research authors have used the social media as a methodology to connect and extract the data required for analysis. It is in the normal process of the business to receive the feedback and not act upon it. There are no statistics maintained in any organization to understand the overall acceptance of the product of service. However this application work would in the long run help all the organization to understand the product very well as well as the feedbacks of the customers and thereby help in providing a better product or service to the end customer. This analysis results in a set of improvements to be implemented in the future version of the research work.

REFERENCES

[1] Blake Morgan, "Five Trends Shaping The Future Of Customer Experience In 2018", https://www.forbes.com/sites/blakemorgan/2017/12/05/five-trends- shaping-the-future-of-customer-experience-in-2018/#286e7f9c2d9c Dec 5, 2017

[2] Diksha Khurana, Aditya Koli, Kiran Khatter and Sukhdev Singh, "Natural Language Processing: State of The Art, Current Trends and Challenges", https://arxiv.org/ftp/arxiv/ papers/1708/ 1708.05148.pdf

[3] Ga¨el Dias, Rumen Moraliyski, Jo˜ao Cordeiro, "Automatic Discovery of Word Semantic Relations using Paraphrase Alignment and Distributional Lexical Semantics Analysis", Natural Language Engineering, Volume 1, Page No. 1–30. 2010 Cambridge University Press.

[4] Jocelyn D'Souza, A dive into Natural Language Processing, https://medium.com/greyatom/a-dive-into-natural-language-processing-103ae9b0a588Oliveira, T.A., Kitsos, C.P., Oliveira, A., Grilo, L.M, "Recent Studies on Risk Analysis and Statistical Modeling", Springer International Publishing, 2018

[5] MarketandMarkets , "Customer Experience Management & Voice of the Customer Analytics Market Size and Forecasts " (2012-2017) by 12/2012.

[6] Oliveira, T.A., Kitsos, C.P., Oliveira, A., Grilo, L.M, "Recent Studies on Risk Analysis and Statistical Modeling", Springer International Publishing, 2018

[7] Vishant Kumar, Rajkumar Singh Rathore, "A Review on Natural Language Processing", 2016 IJEDR | Volume 4, Issue 3 | ISSN: 2321 – 9939