WR Based Opinion Mining on Traffic Sentiment Analysis on Social Media

Karan Bhandari[1], Sumit Keskar[2], Mohammed Haseeb[3], Neelash Raina[4]

[1] *Student,Computer Enggineering PDEA's COEM, Maharastra, India*
[2] *Student,Computer Enggineering PDEA's COEM, Maharastra, India*
[3] *Student,Computer Enggineering PDEA's COEM, Maharastra, India*
[4] *Student,Computer Enggineering PDEA's COEM, Maharastra, India*

ABSTRACT

Sentiment analysis or opinion mining is a computing device finding out process in which class of the human's sentiments, feelings, opinions and so forth in the shape of constructive, horrible or impartial remarks underlying the text. The social media is normally growing technological know-how that may upload and using it appreciably. In this social media similar to face eBook, twitter, on-line discussion board and different internet, customers usually use it and offer their response and pointers for any regular speedy. There is various application of sentiment assessment and plenty of researchers have check on those features but there are no more reviews on transportation technique, for guard, efficient transportations. consequently to reduce the visitors associated troubles, the internet site on-line site visitors sentiment evaluation (TSA).This survey will try and attention on sentiment evaluation strategies, associated paintings for automated net records crawling, one in all a type levels of SA, subjectivity magnificence, a few pc getting to know techniques on the idea in their usage and significance for the assessment, assessment of Sentiment classifications and its current developments and the long-time research instructions within the difficulty of web page site visitors Sentiment assessment. With the booming of social media, sentiment analysis has advanced hastily in latest years. However, only some opinions desirous about the region of transportation, which did now not meet the stringent necessities of protect, performance, and know-how alternate of sensible transportation techniques (ITSs). Our paintings will support the development of TSA and its purposes.

Keyword : - *Rule base, Sentimental analysis, Sentimental base, Web-based*

---

1. INTRODUCTION

To develop a web monitoring system to automatically retrieve tweets related to traffic, and by using that traffic related tweets, extract the potential safety topics (e.g., traffic accidents, road flooding), calculate sentiment polarity by using unsupervised approach of sentiment analysis.

1.1 Literature Survey

By undertaking a literature review it is able to critically summarize the current knowledge in the area under investigation, identifying any strengths and weaknesses in previous work, so it helps to identify them in research and thus eliminate the potential weaknesses, whilst bringing to the fore the potential strengths. In addition, a good and full literature search will provide the context. It provides an up-to-date understanding of the subject and its significance to practice. It identifies the methods used in previous research on the topic;
It helps to work out how to answer the questions - and indeed, what questions need to be asked. It provides comparisons for research findings.

1.2 Natural Language Processing And Sentiment Analysis

Current human-centric paradigm of information technology places user experience and human-computer interaction in the centre of attention of developers in small and large companies. The trend is led by such software giants as

Google and Microsoft. The most natural way information can be exchanged between humans and computers, as well as between humans themselves, is natural language the language we use for everyday communication. There is a great variety of scenarios where computer's ability to deal with human language can play crucial role in information technology. An example is machine translation, which enables computer-mediated exchange of information, written or spoken, between people speaking different languages. This is especially important given a huge misbalance between languages in which information is produced and in which the users would prefer, or just be able, to consume it: while the main body of information in Internet is produced in English or a few major languages, the main body of Internet users are not native speakers of those languages or even don't understand them.
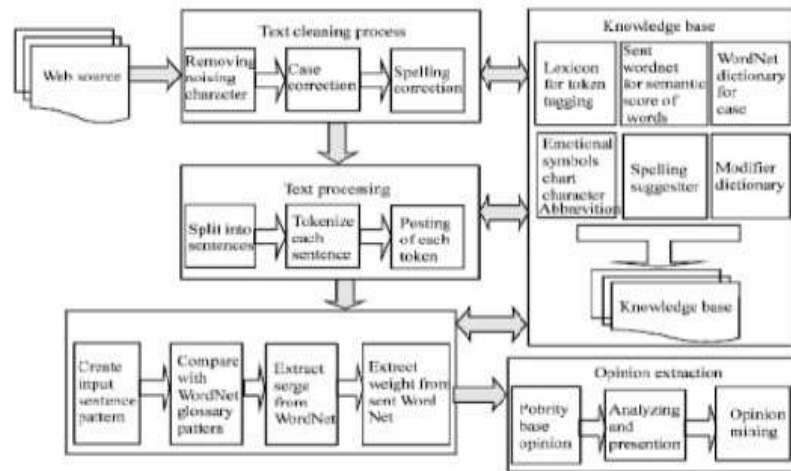

Fig -1: Sentiment analysis of the proposed system

1.2 Problem Statement

- develop a web monitoring system to automatically retrieve tweets related to traffic, and by using that traffic related tweets, extract the potential safety topics (e.g., traffic accidents, road flooding), calculate sentiment polarity by using unsupervised approach of sentiment analysis. The potential users includes transportation engineers (e.g. early identification of safety bottleneck), Transportation planner (e.g. Adjustment of transportation policies in response to public sentiment and opinion) and public user (e.g. improved routine to avoid potential unsafe regions). Such techniques provide citizen oriented government model where the priorities and services should be provided according to the citizens.

- There are various researches in sentiment analysis and its related areas but there are no more studies on transportations, hence there is lack of efficiency and safety. Hence to reduce the traffic related problems, we proposes the traffic sentiment analysis (TSA).

2. GOALS AND OBJECTIVES

- The goal of system is to develop a web monitoring system to automatically retrieve tweets related to traffic, and by using that traffic related tweets, extract the potential safety topics (e.g., traffic accidents, road flooding) calculate sentiment polarity by using unsupervised approach of sentiment analysis.
- Also provide and calculate this real time information for needed one.
- It will help to reduce traffic related injuries and fatalities, it would be helpful to monitor traffic in real time in order to quickly identify regions and activities that have the potential to become a risk to public safety.

3. ADVANTAGES:

1. The results produced via machine learning techniques are quite good in comparison to the human generated baselines.

2. Perform features based on unigrams (with negation tagging) and Unigram presence information perform well

3. Improve detection of sentiments

4. DISADVANTAGES:

1.   Human generated baseline is not worked up to the expectation.

2. Topic based categorization is not performed.

3. Applied on for movie data.

5. ARCHITECTURE AND PROCESSES

5.1 TSA Architecture

Previous studies on Chinese texts have devoted consider- able efforts on architectural design.
Che et al. designed the architecture of the language technology platform (LTP), an integrated Chinese processing platform including a suite of high-performance natural language processing (NLP) modules and relevant corpora. They achieved plausible results in several relevant evaluations, particularly for syntactic and semantic parsing modules . Li et al. designed the architecture of
sentiment analysis application in the financial domain on the basis of morphemes .

A rule-based approach is adopted here to address the distinct challenges posed by the Chinese data set. Fig. 2 illustrated the architecture of   TSA; the architecture is based on the tackling process; and its main components, including 1) web data col-
lection, 2) preprocessing, 3) extraction of subjects and objects, 4) extraction of sentiment properties, 5) sentiment calculation and classification, 6) evaluation or applications, and 7) feed- back, improve the construction of the sentiment, rule, and TSA object bases.
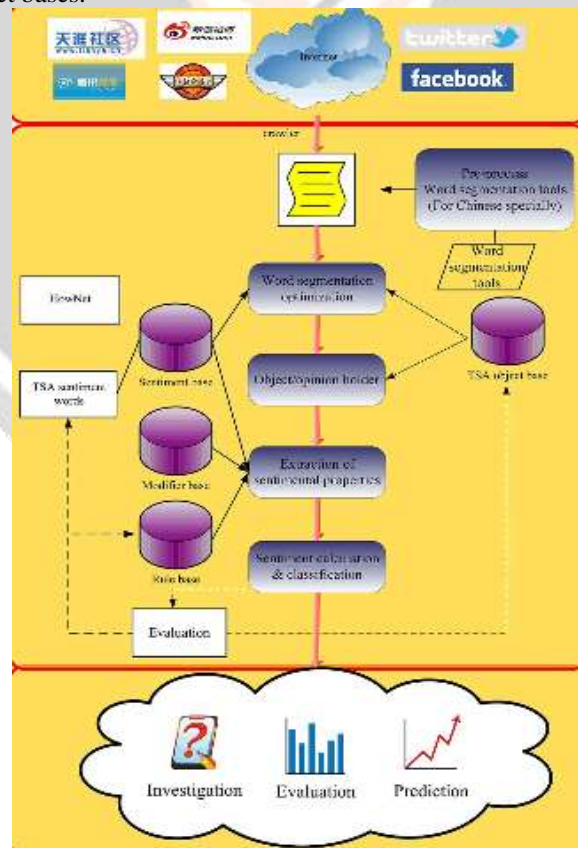

Fig -2: Architecture of the rule-based TSA
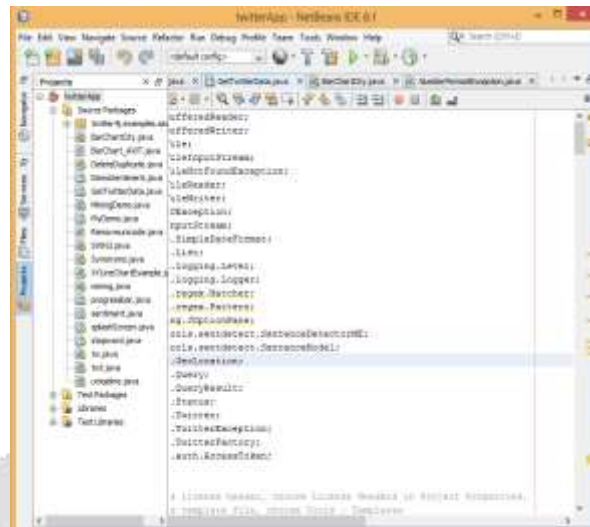
4. Result Analysis

Fig -3: Initial Window

When we run the GetTwitterData file the new window opens as shown in Figure 3.

In the search textbox we require to enter specific keyword. And in the date textbox we require to select time period from specific date to another date. After that choose one city, then it returns the tweets regarding to that keywords between that selected time period and for that selected city. After retrieving the data when we press the remove button then it remove the unnecessary data like RT, https, @, Co/, #, etc.
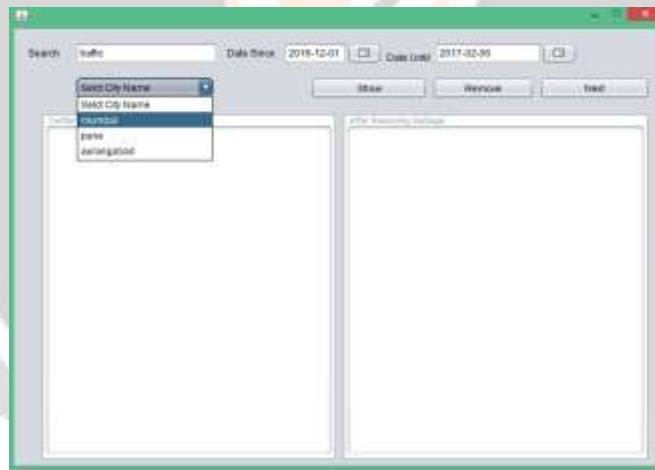

Fig -4: . After removing unnecessary data it removes duplicate lines from the retrieved tweets
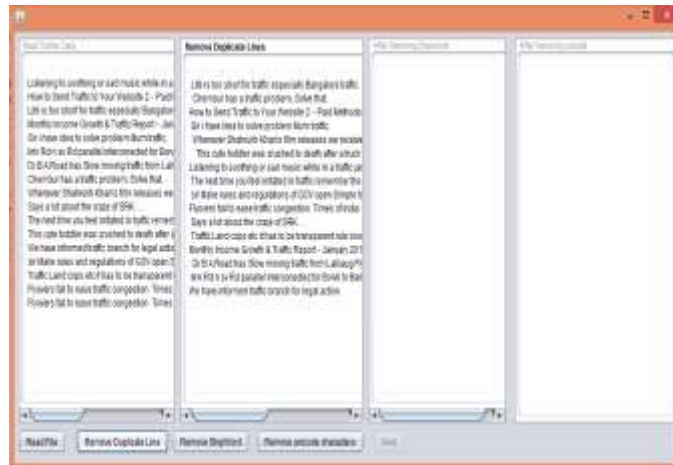
Fig -5: After that when we press on the remove stopwords it removes the stopwords from that file
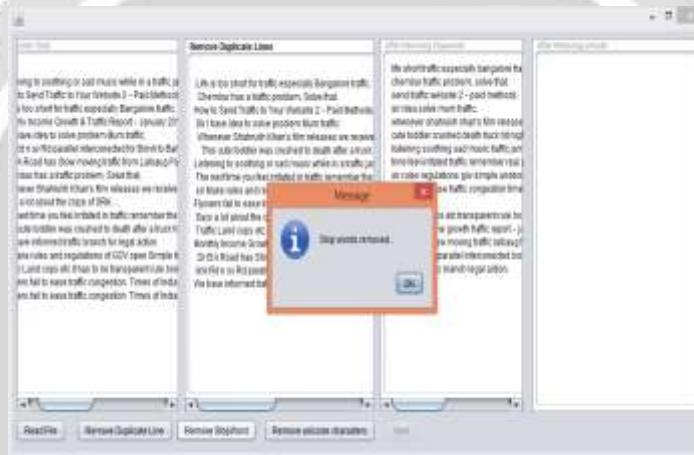

Fig -6: When we press on the remove unicode characters it remove the characters like ♥ ,✓,☀,★, ↑,
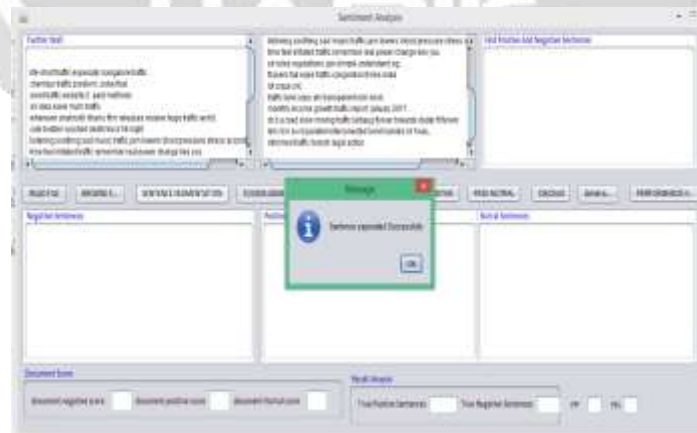

Fig -7: After that tokens get generated

5. CONCLUSION

- We proposed the distributed deduplication systems to improve the reliability of data while achieving the confidentiality of the users' outsourced data without an encryption mechanism. Four constructions were proposed to support fine-grained block-level data deduplication. The security of tag consistency and integrity

were achieved.We implemented our deduplication systems using the Ramp secret sharing scheme and demonstrated that it incurs small encoding/decoding overhead compared to the network transmission overhead in regular upload/download operations.

- In this work, we have identified a new privacy challenge during data accessing in the cloud computing to achieve privacy-preserving access authority sharing. Authentication is established to guarantee data confidentiality and data integrity. Data anonymity is achieved since
- the wrapped values are exchanged during transmission. User privacy is enhanced by anonymous access requests to privately inform the cloud server about the users' access desires. Forward security is realized by the session identifiers to prevent the session correlation. It indicates that the proposed scheme is possibly applied for enhanced privacy preservation in cloud applications.

6. REFERENCES

[1] F. Y. Wang, "Social computing: Concepts, contents, and methods," Int. J. Intell. Control Syst. , vol. 9, no. 2, pp. 91–96, 2004.

[2] F.-Y. Wang, R. Lu, and D. Zeng, "Artificial intelligence in China," IEEE Intell. Syst. , vol. 23, no. 6, pp. 24–25, Nov./Dec. 2008.

[3] Jin Li, Xiaofeng Chen, Xinyi Huang, Shaohua Tang and Yang Xiang Senior Member, IEEE and Mohammad Mehedi Hassan Member, IEEE and Abdulhameed Alelaiwi Member, IEEE "Secure Distributed Deduplication Systems with Improved Reliability"IEEE Transactions on Computers Volume: PP Year: 2015

[4] Miss Prachi D. Thakar, Dr. D.G.Harkut "Cloud based Hybrid Model for Authorized Deduplication" International Journal of Application or Innovation in Engineering & Management (IJAIEM). May 2015

[5] T. Sivashakthi1, Dr. N Prabakaran "A Survey on Storage Techniques in Cloud Computing" International Journal of Emerging Technology and Advanced Engineering 2013