https://doi.org/ 10.33472/AFJBS.6.10.2024.3860-3869

African Journal of Biological Sciences
Journal homepage: http://www.afjbs.com

Research Paper Open Access

A Sentence Level Sentiment Analysis Model for Improved Content Generation in Social Network

Dr.Kumar Shwetabh, Assistant Professor, Faculty of CS & IT, Kalinga University, Naya Raipur, Chhattisgarh, India.
Mail ID:ku.kumarshwetabh@kalingauniversity.ac.in
ORCID ID: 0009-0006-5751-352X
Nikita Pathrotkar, Assistant Professor, Faculty of CS & IT, Kalinga University, Naya Raipur, Chhattisgarh, India.
Mail ID:ku.nikitapathrotkar@kalingauniversity.ac.in
ORCID ID: 0009-0009-4986-8752

Article History

Volume 6,Issue 10, 2024

Received:13 Apr 2024

Accepted : 05 May 2024

doi: 10.33472/AFJBS.6.10.2024.3860-3869

Abstract:
Content generation has been identifies as a key challenge in a social network. As the social network is being in a peak in the human society, generating content according to the sentiment of any user is a challenging task. There exist number of approaches in content generation and sentiment analysis; they use the tags, messages posted and page visited as the key feature in identifying the sentiment of the user towards content generation. Such methods suffer to achieve higher performance in sentiment analysis as well as in content generation. To handle this issue, a Sentence Level Sentiment Analysis Model (SLSAM) is presented in this article. The proposed method maintains the traces of different social network and the users of the environment. First, the method collects the conversational traces of any user. The traces are preprocessed to eliminate the meaningless words. With the subject taxonomy available, the method estimates Sentimental Topical Measure (STM) and Sentimental Emotion Measure (SEM). Based on these values, the method identifies the sentiment of the person and generates content for the user by measuring Topical content support (TCS) for different contents of the class. The proposed method improves the performance of content generation and sentiment analysis.
Index Terms:Social Network, Sentiment Analysis, Content Generation, SLSAM, STM, SEM, TCS.

1. Introduction:

The social network is the most dominant environment being used by various peoples of the world. The social network users share many things between groups or in a public forum. It has begun an open forum where the people can post their opinion about any entity which has been visited and replied by variety of peoples. As the contents and posts of various users are shared among number of peoples, it is being used to perform different real time problems like public opinion, public voting and so on.

On the other side, the same social network content managers populate variety of contents for their users. Even though there exist number of contents, the methods populates only limited one related to the interest of the user. For example, if the user is interested on viewing the content about the upcoming movies, then the social network would populate and shares content about the new arrival movies and related events on their page. This is being performed based on the number of times the user has explored and conversed with others. However, the methods uses such frequency of conversation and repeated conversation topics as the key to produce content on their page.

The social network user would have number of interest and the content managers should populate the content according to the interest of the user. The existing method identifies the interest or sentiment of the user based on the terms present in the conversational text features. Such methods are not suitable for producing content on the social network. It is necessary to consider the sentence features in identifying the interest of the user. With this consideration, an Efficient Sentence Leven Sentiment Analysis Model (SLSAM) is presented in this article.

The SLSAM model is focused on identifying the sentiment of the user in sentence level. The conversation would have number of sentiments and emotions, by considering the terms of the sentence; the proposed method finds the interest of the user to produce content to the user. The method estimates Sentimental Topical Measure (STM) and Sentimental Emotion Measure (SEM). Based on these values, the method identifies the sentiment of the person and generates content for the user by measuring Topical content support (TCS) for different contents of the class. The detailed approach is briefed in this section.

2. Related Works:

The methods related to sentiment analysis and content generation is analyzed in this section.

A social relations-guided multi-attention networks (SRGMANs) is presented in [1], which consider the visual features in region and object level to measure the correlation between the social images to identify the sentiment. Based on these features, the emotion is extracted to perform sentiment analysis. An attention-based heterogeneous relational model is presented in [2], which uses progressive dual attention module to capture the relation between the images towards sentiment analysis.

A Gated Fusion Semantic Relation (GFSR) network is presented to explore the semantic relations in social network [3], which identifies inter relation among visual and textual features to perform sentiment analysis.

A multi model sentiment analysis scheme is presented in [4], which analyze the sentiment in social internet of things according to the hierarchical attentions and CSAT-TCN With MBM Network. A lengthened words based sentiment analysis is presented in [5], which uses a lexicon based system works by concerning the lightened words and computes senti scores towards different emotions to perform sentiment analysis. A gabor CNN model is presented in [6], which performs visual sentiment analysis on social media data in cloud environment. The method uses two variants of GCNN model which perform pooling and normalization in different strategy. A machine learning sentiment analysis model is presented in [7], which uses monkeypox outbreak data to analyze the public sentiment on recent monkeypox outbreak.

Multi-level correlation mining framework is presented in [8], which identifies the correlation among the modalities of high and low level. The method uses three level of analysis like linguistic, acoustic, and visual tasks to label the sentiment by designing self-supervised label generation module (SLGM) to support sentiment analysis.

An integrated machine learning based sentiment analysis model is presented in [9], which provide three key aspects to process and analyse the growing digital news articles. An aspect based sentiment analysis model (ASAM) is presented in [10], which performs restaurant survival prediction with customer generated content. .

Sentiment analysis on chinese students is presented in [11], which perform opinion mining on LMOOC, and works according to the comments provided. A GAN based synthetic text generation towards sentiment analysis is presented in [12], which uses different opinion data and feedback data sets for synthetic text generation using CatGAN and SentiGAN models. Different approaches of hate speech detection and sentiment analysis is analyzed in [13], which analyze the challenges faced by various deep learning models. An internal threat detection based sentiment analysis model is presented in [14], which uses Natural Language ToolKit (NLTK) to analyze the sentiment of each preprocessed email and produces a graph. With the use of singular value decomposition (SVD), the matrix is decomposed into base networks, and we analyze the change in importance of important base networks over time to detect the company's internal conditions. An Arabic dialect sentiment analysis model is presented in [15], which analyze the performance of various sentiment analysis model in detail.

3. Sentence Level Sentiment Analysis Model (SLSAM):

The proposed method maintains the traces of different social network and the users of the environment. First, the method collects the conversational traces of any user. The traces are preprocessed to eliminate the meaningless words. With the subject taxonomy available, the method estimates Sentimental Topical Measure (STM) and

Sentimental Emotion Measure (SEM). Based on these values, the method computes the value of Sentiment Class Value (SCV). Based on the value of SCV, the method identifies the sentiment of the person and generates content for the user by measuring Topical content support (TCS) for different contents of the class. Based on the TCS value, the method selects the content to be populated on the user page. The working procedure is described in detail in this part.

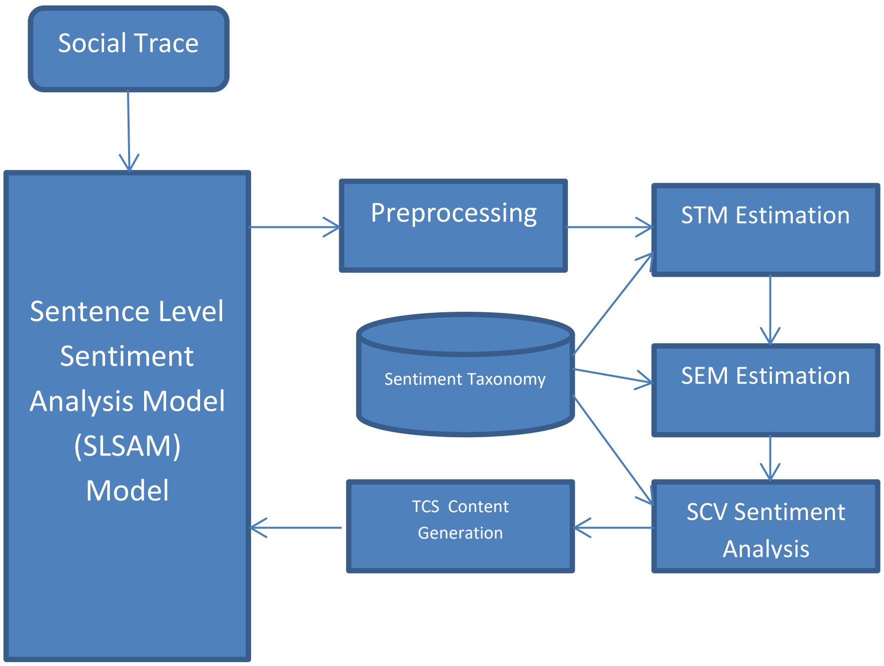

Figure 1: Architecture of Proposed SLSAM Model

The functional architecture of SLAM sentiment analysis model is presented in Figure 1, and the functions are detailed in this part.

Preprocessing:

The conversational text has been given as input to this module which split the text into number of sentences. Further, the method extracts the terms by splitting to produce term set at each sentence. Now, for each sentence term set, the method identifies the stop words and eliminates them. Further, the terms are applied with stemming and part of speech tagging. Finally, from each term set, the method selects only nouns as result to support sentiment analysis.

Algorithm:

Given: Conversational Text CT, Stop word list Swl.

Obtain: Set of Term set STs.

Start

Read CT and Swl.

Initialize STS.

Sentence set ss = Split(CT, “.”)

```
        For each sentence s
                Term set Tes = Split (s,” “)
                For each term T
                        If Swl∈ T then
                                Tes = Tes ∩ T
                        Else
                                T = Stemming (T)
                                Tag = POSTagging(T)
                                If Tag==Noun then
                                        Leave
                                Else
                                        Tes = Tes∩ T
                                End
                        End
                End
                STS = STS ∪ Tes
        End
Stop
```

The preprocessing algorithm finds the sentences in the conversation and identifies set of terms from each sentence to support the analysis.

STM Estimation:

The sentimental topical measure is the value which represents the class of sentiment topic present in any sentiment. The conversational text would have number of sentiments and topics. It is necessary to identify the topic of the text, so that STM is measured. To measure the topical measure, the method uses the sentiment taxonomy as the key and computes the STM value towards number of topical classes. Estimated set of STM are used to measure the SCV value towards sentiment analysis.

Algorithm:

Given: Term Set Tes, Sentiment Taxonomy ST, Class C.

Obtain: STM.

```
Start
        Read Tes, ST, C.
        Compute STM.
```

$$STM = \frac{\sum_{i=1}^{Size(Tes)} \sum_{j=1}^{size(ST)} Count(ST(j).Topic == C \ \&\& \ ST(j).Term == Tes(i))}{Size(Tes)} \times \frac{\sum_{i=1}^{Size(Tes)} \sum_{j=1}^{size(ST)} Count(ST(j).Topic == C \ \&\& \ ST(j).Term \equiv Tes(i))}{Size(Tes)}$$

Stop

The STM estimation algorithm counts the number of topical terms completely equal to the terms in the term set and number of topical terms approximately equal to the terms of the set. Estimated value of STM is used to perform sentiment analysis.

SEM Estimation:

The Sentiment Emotion measure represent the emotional similarity of terms present in the term set obtained. It has been measured by counting the number emotional terms present in the term set for the total terms. Estimated value of SEM is used to perform sentiment analysis.

Algorithm:

Given: Term Set Tes, Sentiment Taxonomy ST, Class C.

Obtain: SEM.

Start

Read Tes, ST, C.

Compute SEM.

$$SEM = \frac{\sum_{i=1}^{size(Tes)} Count(Tes(i) \in ST(C))}{Size(Tes)}$$

Stop

The sentiment emotional measure is measured based on the occurrence of sentimental terms in the term set with total term size. Measured SEM value has been used to perform sentiment analysis.

SCV Sentiment Analysis:

The sentiment of the conversation is measure which represents the sentiment similarity of the conversation with number of sentiment classes. To perform this, the method reads the conversation text and applies preprocessing to obtain the set of term sets from different sentences available. For each term set, the method computes STM and SEM values. Using the STM and SeM value, the method computes the value of SCV. Based on the value of SCV, the method identifies the sentiment class with maximum SCV value as result.

Algorithm:

Given: Conversation text CT, Sentiment Taxonomy ST.

Obtain : Sentiment class Sc.
Start
Read CT and ST.
Term set sets Tss = Preprocessing (CT)
For each term set ts
For each sentiment class sc
Compute STM = STM-Estimation (Ts, ST, sc)
Compute SEM = SEM-Estimation (Ts, ST, sc)
Compute SCV = STM×SEM
End
End
Class sc = Choose the sentiment class with maximum SCV.
Stop

The above pseudo code represents the working of sentiment analysis and identifies the sentiment class of the conversation text given.

TCS Content Generation:

The proposed content generation algorithm computes Topical content support (TCS) towards various contents of the sentiment identified. To perform this, the method computes TCS towards various content available on the class. It is measured based on the number of topical terms present in the content with the number of terms available. Based on that the method identifies set of content as result.

Algorithm:
Given: Sentiment class sc, Content Set CS.
Obtain: result
Start
Read Sc and Cs.
For each content c

Compute Tcs = $\frac{\overset{size(ST)}{\underset{i=1}{Count(ST(i)\in c)}}}{size(ST)}$

If Tcs>Th then
Add to result R = $\sum(c \in R) \cup c$
end
End
Stop

The content generation algorithm computes the TCS value for various content and identifies the set of content as result for the user.

4. Results and Discussion:

The proposed SLAM sentiment analysis and content generation algorithm has been implemented with advanced java. The method has been evaluated for its performance with Twitter data set. Obtained results are compared with the results of other approaches.

Parameter	Value
Tool Used	Advanced java
Data set	Twitter data set
Number of sentiment class	5
Number of tweets	1 million
Number of users	500

Table 1: Experimental details.

The experimental details used for the performance evaluation is presented in Table 1. The results obtained are compared with the result of others.

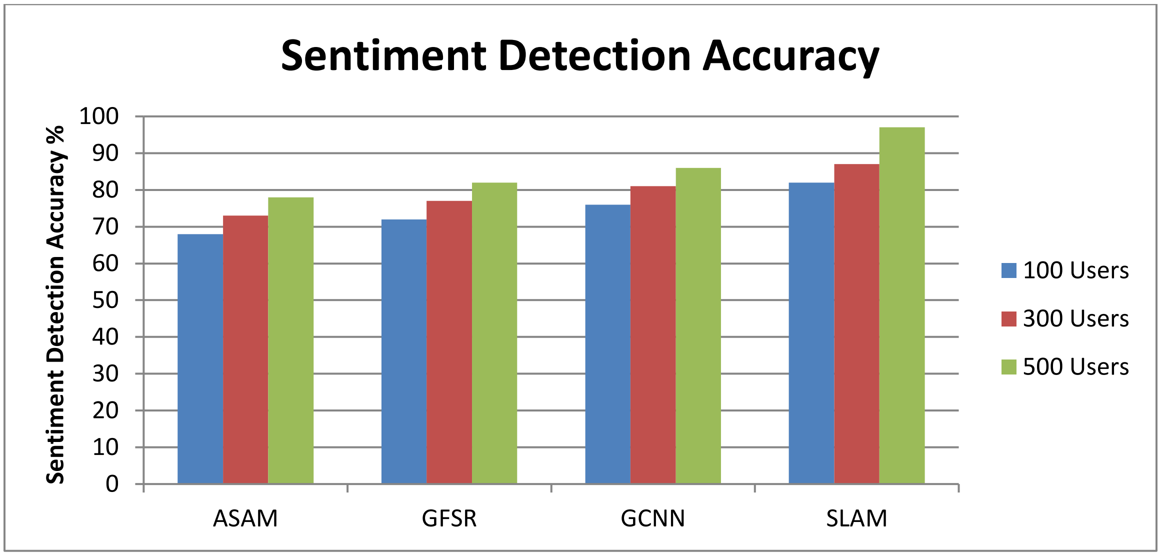

Figure 1: Sentiment Detection Accuracy %

The accuracy of methods in detecting the sentiment is measured for various number of users and presented in Figure 1. The proposed SLAM method introduces higher accuracy than other methods.

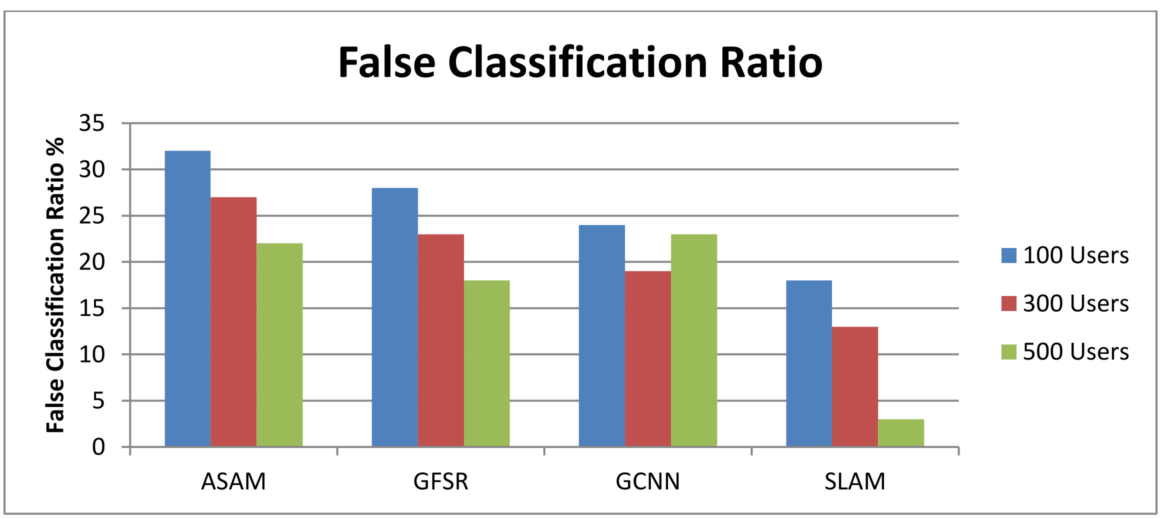

Figure 3: Analysis on False classification ratio

The ratio of false classification produced by different methods on sentiment detection is measured and compared in Figure 3. The proposed SLAM method introduces less false ratio than other methods.

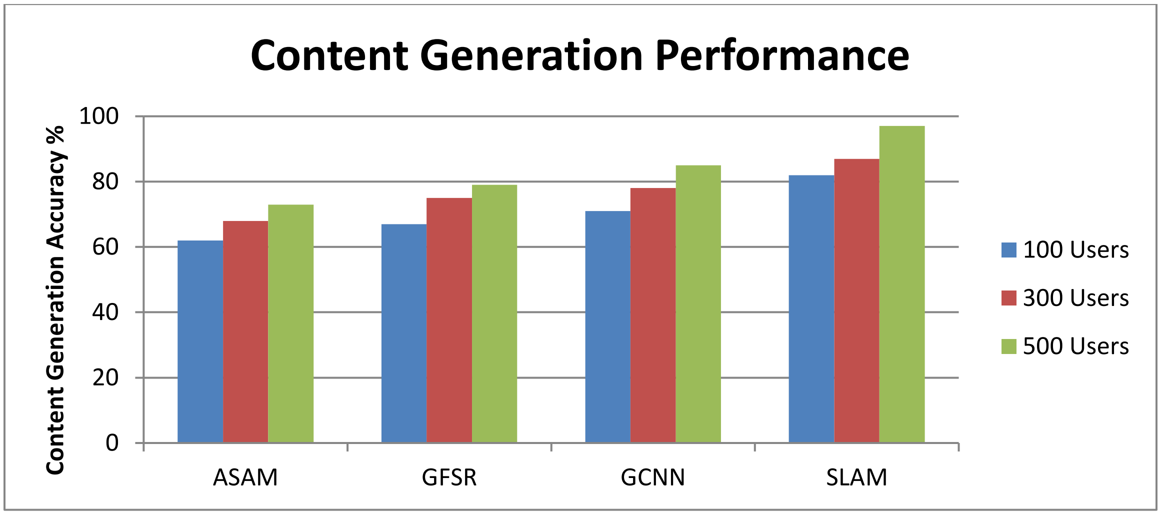

Figure 4: Analysis on Content Generation Accuracy %

The content generation accuracy is measured and compared with the results of other methods. The proposed SLAM model improves the accuracy of content generation then other methods.

5. Conclusion:

This paper presented a Sentence Level Sentiment Analysis model towards sentiment analysis and content generation. The traces are preprocessed to eliminate the meaningless words. With the subject taxonomy available, the method estimates Sentimental Topical Measure (STM) and Sentimental Emotion Measure (SEM). Based on these values, the method computes the value of Sentiment Class Value (SCV). Based on the value of SCV, the method identifies the sentiment of the person and generates content for the user by measuring Topical content support (TCS) for different contents of the class. Based on the TCS value, the method selects the content to be populated on the user page.

REFERENCES:

1. J. Xu, Z. Li, F. Huang, C. Li and P. S. Yu, "Visual Sentiment Analysis With Social Relations-Guided Multiattention Networks," IEEE (TC), Volume. 52, Number 6, pp. 4472-4484, 2022.
2. J. Xu, Z. Li, F. Huang, C. Li and P. S. Yu, "Social Image Sentiment Analysis by Exploiting Multimodal Content and Heterogeneous Relations," IEEE (TII), Volume. 17, Number 4, pp. 2974-2982, 2021.
3. J. Zeng, J. Zhou and C. Huang, "Exploring Semantic Relations for Social Media Sentiment Analysis," IEEE (ACM/ TAS&LP), Volume. 31, pp. 2382-2394, 2023.

4. G. Xiao et al., "Multimodality Sentiment Analysis in Social Internet of Things Based on Hierarchical Attentions and CSAT-TCN With MBM Network," IEEE (ITJ), Volume. 8, Number 16, pp. 12748-12757, 2021.

5. A. Kukkar, R. Mohana, A. Sharma, A. Nayyar and M. A. Shah, "Improving Sentiment Analysis in Social Media by Handling Lengthened Words," in IEEE Access, Volume. 11, pp. 9775-9788, 2023.

6. S. Urolagin, J. Nayak and U. R. Acharya, "Gabor CNN Based Intelligent System for Visual Sentiment Analysis of Social Media Data on Cloud Environment," in IEEE Access, Volume. 10, pp. 132455-132471, 2022.

7. S. Bengesi, T. Oladunni, R. Olusegun and H. Audu, "A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets," in IEEE Access, Volume 11, pp. 11811-11826, 2023.

8. Zuhe LI, “Multi-level correlation mining framework with self-supervised label generation for multimodal sentiment analysis”, ELSEVIER (IS), Volume 99, Number 101891, 2023.

9. Jeelani Ahmed,” Classification, detection and sentiment analysis using machine learning over next generation communication platforms”, ELSEVIER (M&M), Volume 98, Number 104795, 2023.

10. Hengyun LI, “Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews”, ELSEVIER ™, Volume 96, Number 104707, 2023.

11. Jian – E Pang, “Mining opinions on LMOOCs: Sentiment and content analyses of Chinese students’ comments in discussion forums”, ELSEVIER (S), Volume 109, Number 102879, 2023.

12. Ali Shariq Imran, “The impact of synthetic text generation for sentiment analysis using GAN based models”, ELSEVIER (EIJ), Volume 23, Issue 3, PP 547-557, 2022.

13. Malliga Subramaniyan, “A survey on hate speech detection and sentiment analysis using machine learning and deep learning models”, ELSEVIER (AEJ), Volume 80, PP 110-121, 2023.

14. Xueyuan Wen,” An approach to internal threats detection based on sentiment analysis and network analysis”, ELSEVIER (JIS&A), Volume 77, Number 103557, 2023.

15. Yassir Matrane,” A systematic literature review of Arabic dialect sentiment analysis”, ELSEVIER (JKSU-C&IS), Volume 35, Issue 6, Number 101570, 2023.