

A Framework for Sentiment Analysis of Telugu Tweets



G. Balakrishna Priya, M. Usha Rani

Abstract: Now a day Social Media like Facebook, twitter and Instagram is major Sources for people to share their emotions based on the current situations in society. By knowing the interesting patterns in it, a government/appropriate person for that situation can take good and useful decisions. Sentiment analysis is a method where people can extract the useful information from the text like the emotions (happy, sad, and neutral) of people. Much research work has been underdoing in the area of sentiment analysis. Among that work the Machine learning and Deep learning approaches plays a maximum role. Existing works on sentiment analysis is going in the English language. In this paper, proposed a novel framework that specifically designed to do sentiment analysis of the text data, that available in the telugu language. The proposed framework was integrated with the word embedding model Word2Vec, language translator and deep learning approaches like Recurrent Neural Network and Navie base algorithms to collect and analyse the sentiment in tweeter data that present in telugu language. The results shows effective in terms of accuracy, precision and specificity.

Keywords: Sentiment Analysis, Framework, Twitter data, Telugu and Deep learning.

I. INTRODUCTION

Sentiment analysis is a method to identify the polarity in the text document. The polarity may be a negative, positive and neutral. With the development of technologies like web, the huge textual data is available in the internet, which consists of the historical, news, science and political from the major social media websites like twitter and face book [1]. To determine the thoughts of a person opinion mining or sentiment analysis is used. This can be done in sentence level, document level and future levels in terms of positive, negative and Neutral [2]. Sentiment analysis has very popular due to the web. The current customer behaviour can be obtained to the business people by collecting the massive amount of information from the social media like Facebook, twitter and LinkedIn in the form of the tweets, posts and surveys [3]. Market Analyst will develop the products by analysing the reviews of the product given by the Customer, that which helps to know the strategy of the market. To do this the Natural language processing, machine learning and deep learning approaches can be used [4]. Twitter is the major source to every individual to express and respond to the current situations in the world.

Revised Manuscript Received on August 27, 2020.

* Correspondence Author

G. Balakrishna Priya*, Research Scholar, Department of Computer Science, Sri Padmavathi Mahila Viswa Vidyalayam, Tirupati (Andhra Pradesh), India. E-mail: gbkrisnapriya18@gmail.com

Prof. M. Usha Rani, Department of Computer Science, Sri Padmavathi Mahila Viswa Vidyalayam, Tirupati (Andhra Pradesh), India. E-mail: musha_rohan@yahoo.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

So this will help to the people to overcome some disaster situations and can do better for the society by understanding the individual behaviour [5]. This paper is focused on the sentiment analysis of the telugu language tweets. Here proposed a framework to do the sentiment analysis and the results were obtained. This paper is organized as follows, in section 2 the start of art were discussed, in section 3 the proposed system were explained, in section 4 the results and discussion were done and finally in section 5 the conclusion and future work was explained.

II. STATE OF ART

Automated sentiment analysis was done by using the deep learning approaches like LSTM, GRU and BiLSTM and obtained the results terms of the accuracy, Precision, Recall, and Specificity [2]. Sentiment Analysis on Tamil tweets and Saudi dialect was done by using the deep learning algorithms like LSTM, and BiLSTM. Here the results was shown in the terms of the accuracy and F-score, 80% of training dataset and 20% of the testing data was taken to perform the experiment [3,6,7]. Sentiment analysis was performed by using the Natural language processing and deep learning approaches like RNN and Navie bays algorithms. Here tweets are collected from the twitter by using the API keys of the twitter account and obtained the results in terms of the Accuracy, precision, F-Score and Recall with the polarities like negative, positive and neutral [4]. Supervised machine learning algorithms was used to analysis the tweets in the Turkish language and Saudi dialect. The algorithms are support vector machine and Random forest algorithms. The results were shown in terms of the accuracy [5, 7]. Sentiment analysis for the tweets that available in the languages other than English was reviewed and specified the various challenges involved in it [8]. Sentiment analysis was done for the region language kannada based on the domain knowledge and by using decision tree classifier [9]. Experiment analysis was done for the word embed tool Word2vec in the bangala language [10].

III. PROPOSED FRAMEWORK

Here proposed a Framework to perform sentiment analysis of a telugu language tweets that obtain from a tweeter account. This framework facilitates the users to analyse the telugu language and find the sentiment analysis by using the user defied deep learning algorithms. The framework consists of a steps Tweets, Word embed tool, Language Translator and deep learning algorithms.

A Framework for Sentiment Analysis of Telugu Tweets

In the Tweets steps, the telugu tweets that obtain from the tweeter account is added to the framework that which user want to find the sentiment analysis. In the Word embed tool step by using any word embed tools like wordtvec is used to represents the words in vector format to identify the meaning in it. Then the language translator tool is used to

convert tweets from telugu language to the English language and finally by using the deep learning algorithms like RNN and Naviebays algorithms the sentiment analysis was identified in the tweets that uploaded in this proposed framework.

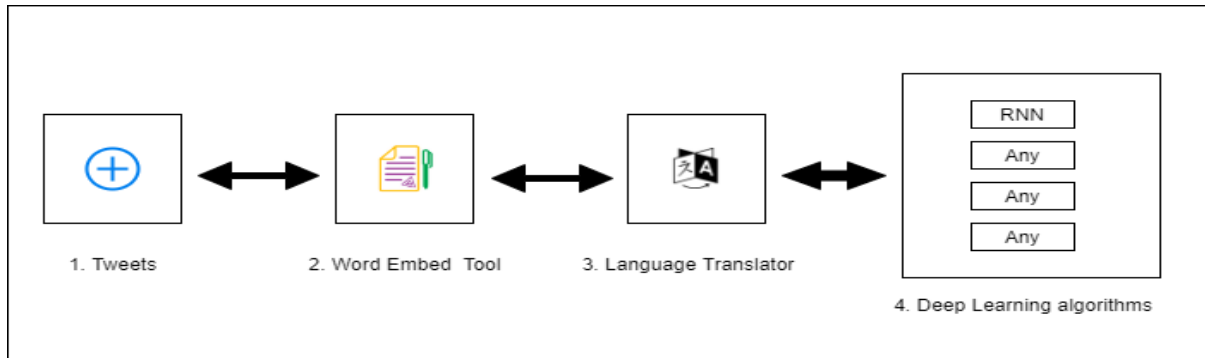


Figure 3.1 Proposed Framework to perform Sentiment Analysis on Tweets available in Telugu Language

IV. RESULTS AND DISCUSSIONS

Here telugu tweets are taken manually and prepared as dataset and the details of dataset are tabulated in Table 4.1.

Table 4.1

Dataset	Training data			Testing data
	Negative	Positive	Neutral	
Telugu Twitter data	456	396	687	756

The collected data is pre-processed and then applied the word embedding tool and then the tweets are translated and applied the deep learning algorithms to the sentiment analysis. For the applied dataset the results was shown in the Table 4.2 and Figure 4.1 in terms of Accuracy, Specificity and Precision.

Table 4.2

	Accuracy	Specificity	Precision
Proposed Framework	80.45%	76.57%	82.46%

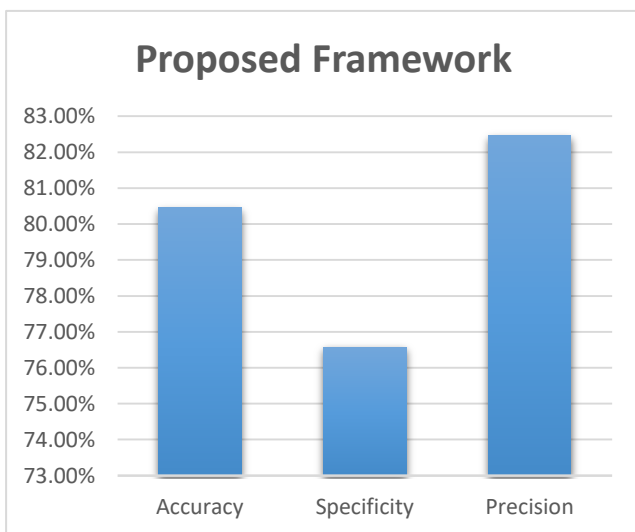


Figure 4.1 Results of Proposed Framework

The proposed framework is tested with the dataset that collected manually and by applying the deep learning

approaches like RNN and Navie bays algorithm the results shows efficiency in terms of the Accuracy, Specificity and Precision.

V. CONCLUSION AND FUTURE WORK

Here a novel framework to find the sentiment analysis of the telugu tweets was proposed and implemented and obtained results. The results shows the efficiency in terms of the Accuracy, specificity and Precision for the dataset taken manually and applied to the framework after pre-processing. Currently the frame work has the deep learning algorithms like RNN and Navie bays algorithms, in future to this framework other algorithms can be added and the dataset size can be increased.

REFERENCES

1. Abd El-Jawad, Mohammed H., Rania Hodhod, and Yasser MK Omar. "Sentiment analysis of social media networks using machine learning." 2018 14th international computer engineering conference (ICENCO). IEEE, 2018.
2. Cheng, Li-Chen, and Song-Lin Tsai. "Deep learning for automated sentiment analysis of social media." Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2019.
3. Anbukkarasi, S., and S. Varadhaganapathy. "Analyzing Sentiment in Tamil Tweets using Deep Neural Network." 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2020.
4. Goel, Vikas, Amit Kr Gupta, and Narendra Kumar. "Sentiment Analysis of Multilingual Twitter Data using Natural Language Processing." 2018 8th International Conference on Communication Systems and Network Technologies (CSNT). IEEE, 2018.
5. Demirci, GözdeMerve, Şeref Recep Keskin, and GülüstanDoğan. "Sentiment Analysis in Turkish with Deep Learning." 2019 IEEE International Conference on Big Data (Big Data). IEEE, 2019.
6. Monika, R., S. Deivalakshmi, and B. Janet. "Sentiment Analysis of US Airlines Tweets Using LSTM/RNN." 2019 IEEE 9th International Conference on Advanced Computing (IACC). IEEE, 2019.
7. Alahmary, Rahma M., Hmood Z. Al-Dossari, and Ahmed Z. Emam. "Sentiment analysis of Saudi dialect using deep learning techniques." 2019 International Conference on Electronics, Information, and Communication (ICEIC). IEEE, 2019.
8. Djatmiko, Fahim, RidiFerdiana, and Muhammad Faris. "A review of sentiment analysis for non-English language." 2019 International Conference of Artificial Intelligence and Information Technology (ICAII). IEEE, 2019.

9. Rohini, V., Merin Thomas, and C. A. Latha. "Domain based sentiment analysis in regional Language-Kannada using machine learning algorithm." 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). IEEE, 2016.
10. Sumit, SakhawatHosain, et al. "Exploring word embedding for bangla sentiment analysis." 2018 International Conference on Bangla Speech and Language Processing (ICBSLP). IEEE, 2018.

AUTHORS PROFILE



G. Balakrishna Priya, is a Research Scholar in the Dept. of Computer Science at Sri Padmavathi MahilaVisvavidyalayam, Tirupati, Andhra Pradesh, India. She has Pursured done Bachelors and Master's degree in Computer Science, Sri Venkateswara University, Tirupati. And she has done M.Phil in the same university. She has published more than 5 articles

in reputed journals.



Prof. M. Usha Rani has working at Dept. of Computer Science, Sri Padmavathi MahilaVisvavidyalayam, Tirupati, Andhra Pradesh, India. She completed her bachelors, Masters and Ph.D in the Sri Venkateswara University, Tirupati. She guided 10+ PhD students, 30+ PG projects and 55+

UG projects. She has published 20+ papers in the reputed journals and presented 30+ conferences nationally and internally. She has filed 3+ patents and published and waiting for the examinations to get grant. He has 2 edited books and 2 authored books. Her interested areas are Sentiment analysis, IoT and Big data analytics.