

Knowledge Based Text & Image Sentiment Detection from Selected Words Using Fuzzy Entropy

¹N.Sushma ²N.Hanumantha Rao

¹M.Tech Scholar, Department of CSE, RVR & JC College of Engineering, AP, India.

²Assistant Professor, RVR & JC College of Engineering, AP, India.

Abstract -Sentiment analysis is the way towards assessing human feelings, conclusions and audits communicated in text to recognize the author's psychological standpoint towards a specific occasion, subject, item, administration, and so forth. Contribute towards conclusion cognizance is troublesome. Recreating this errand of shortlisting of words by human eyewitnesses is trying because of multifaceted nature of human brain's preparing. The utilization of fuzzy entropy is proposed in our work as an imaginative advance to tap feeling remainders of online film audits. We have proposed a novel methodology of shortlisting of words that help in opinion insight utilizing a mix of fuzzy entropy, k-means clustering and notion vocabulary SentiWordNet. Not quite the same as past estimation research, this paper separates assumption into three classes dependent on three-way choice thoughts—in particular, positive, negative and impartial—by leading opinion investigation on client audit data. Assumption Analysis and one best in class technique for regulated AI. Our technique can be adjusted to any vocabulary and any dataset. The trials on benchmark datasets yield better for our methodology when contrasted with the cutting edge.

Keywords: Sentiment analysis, Fuzzy entropy, K-means clustering, Lexicon SentiWordNet

I. INTRODUCTION

Sentiment Analysis is a difficult exploration issue particularly via web-based networking media. Clients can uninhibitedly communicate their perspectives, conclusions and sentiments on various drifting occasions, subjects, and so forth through web based life posts. These presents need on be broke down to comprehend what assessment is passed on through these posts. Slant Analysis, additionally alluded as

feeling AI, includes dissecting sees from the composed content to comprehend and measure human feelings. The online life permits overall clients to interface and associate with one another and express suppositions on general subjects. Social Sentiment Analysis can be utilized to improve client assistance and promoting and furthermore fills in as a proportion of web-based social networking execution. Understanding human mental state and concluding the manner in which people reason and pass on data in their language is critical to develop human-machine cooperation. Ongoing investigations uncover that there exists a connection among perception and feeling; they are related. Supposition examination is the way toward assessing human feelings, conclusions, audits communicated in text to identify the essayist's psychological standpoint towards a specific occasion, theme, item, administration, and so on and dole out an applicable assumption.

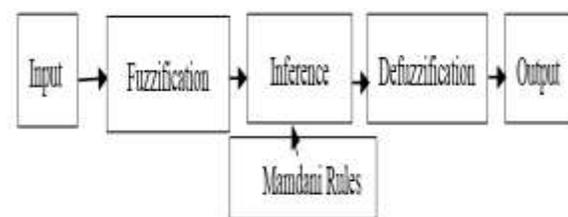


Fig.1: The framework of using a fuzzy logic-based model

Twitter is the well-known online life and brags of a deferential 255 million dynamic month to month clients. A portion of difficulties in examining tweets are: utilization of casual language, short structures, contractions, overwhelming utilization of emoji's and slangs. Twitter, otherwise called microblogging, has restricted size of tweets that makes it hard to register extremity. This procedure utilizes natural language processing (NLP), joined with AI arrangement

methods. NLP is a scope of computational strategies for the programmed examination and portrayal of human language. Word level ways to deal with NLP are initial move towards understanding normal language. NLP methods can be applied at various levels record, sentence, fine-grained, cross-area, cross lingual with directed, unaided and probabilistic generative model-based methodologies.

II. RELATED WORK

A review of natural language processing techniques for opinion mining systems [1]

In this paper, S.Sun et al, introduced an audit of Natural Language Processing (NLP) procedures for supposition mining. In the first place, they presented general NLP methods which are required for text preprocessing. Second, they examined methodologies of feeling digging for various levels and circumstances. At that point they presented relative sentiment mining and profound learning approaches for feeling mining. Supposition rundown and propelled points are presented later.

A Cognition Based Attention Model for Sentiment Analysis [2]

In this work, N.Fang et al, proposed a novel consideration model prepared by insight grounded eye-following information. A perusing forecast model is first fabricated utilizing eye-following information as needy information and different highlights in setting as free information. The anticipated perusing time is then used to construct a cognition based attention (CBA) layer for neural opinion examination. As a far reaching model, They can catch considerations of words in sentences just as sentences in archives. Distinctive consideration instruments can likewise be joined to catch different parts of considerations. Assessments show CBA based technique beats cutting edge nearby setting based consideration strategies essentially. This carries knowledge to how discernment grounded information can be brought into NLP undertakings.

Evaluating the Capability of Text Understanding Based on Cognitive Principle [3]

As indicated by human psychological procedure of text understanding, Text Cognitive Function is proposed to portray capacity of text understanding, which can be communicated by an expansion of different separated watchwords and relations among catchphrases. Text Energy and Text Information are characterized as two qualities in text understanding so as to quantify gained printed data and peruser expended vitality in text understanding procedure, which can be communicated by coefficients of elements of Text Cognitive Function. As per the intellectual efficient rule, the proportion of Text Information to Text Energy is utilized as a standard of assessing ability of text understanding. Investigations check that consequences of the proposed technique comprise with intellectual encounters.

III. FRAMEWORK

Quick development of web has prompted making of gigantic measure of text information, which is unreservedly accessible on web. Estimation examination with NLP can be applied to this information. Be that as it may, in this huge measure of text information, not all words are noteworthy. A few words are a higher priority than others for registering the assumption and these get progressively visual consideration. Consideration models are coordinated with feeling examination to feature the watchwords in text. Perusing time of individual words is likewise a marker of semantic setting in text. Numerous consideration models are constructed utilizing data inserted in text in nearby setting. Profound learning-based calculations have additionally raised the exhibition of opinion examination.

This work endeavors to make a scaffold between human brain's handling of words during perusing and NLP based conclusion investigation framework. We propose a supposition examination framework dependent on LSTM (Long Short-Term Memory) with the use of fuzzy logic for shortlisting words. The fuzzy measure- fuzzy entropy is utilized to extricate intellectual words which is taken care of into LSTM for preparing. These psychological words develop a visual consideration model that copies the manner in which human brain sees text.

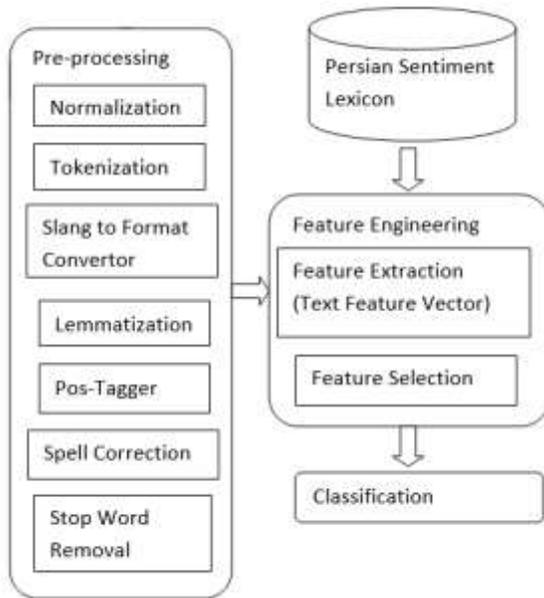


Fig.2: Framework of sentiment detection

evaluations words for feeling remainders utilizing fuzzy entropy. K-means clustering is utilized to isolate the significant words versus non-huge words and the critical words are at last prepared utilizing LSTM classifier. The shortlisted words in text are extricated in a programmed way with no human intercession. Our model imitates the manner in which the human cerebrum thinks, by concentrating on noteworthy words as it were. This segment examines the technique embraced in this paper, beginning with fuzzy logic, with regards to slant examination.

To find significant words and to detect sentiment author is using 3 techniques.

Technique 1 (Calculate positive & negative score):

First sentence will be break into tokens and then assign POS (part of speech) to each token and then extract NOUN, ADJECTIVE, VERB and ADVERB words and then look those words inside SENTIWORD-NET database. If those words available in SENTIWORD-NET database then we will calculate positive and negative score of those words.

Technique2 (Apply fuzzy entropy):

In this technique we will apply fuzzy entropy on scores and remove all those words which has high fuzzy entropy score. As high score means word is less significant. Below code use to calculate entropy.

Technique3 (Apply k-means clustering):

Calculated entropy scores will be passed to k-means clustering and k-means create two clusters where one cluster contains low entropy values and other cluster contains high entropy values. In this step high entropy values will be removed or segregate from low entropy values. All low entropy values contains high significant words. Below code used to calculate kmeans clustering

IV. EXPERIMENTAL RESULTS

In this paper we are describing the concept to detect sentiments from movie reviews dataset called IMDB. This paper propose the concept for sentiment cognition (learning) by shortlisting important or significant words from reviews. All existing techniques never concentrate on obtaining significant words to detect sentiments. As we know instead of reading large text humans prefers to read significant word to understand any topics as significant words require less energy and help in ease understanding of topic. By seeing significant words one person can easily understand emotion of writer.



Fig.3: LSTM model screen

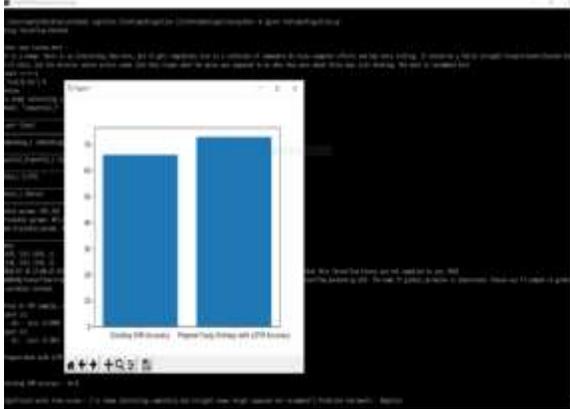


Fig.4: Accuracy graph

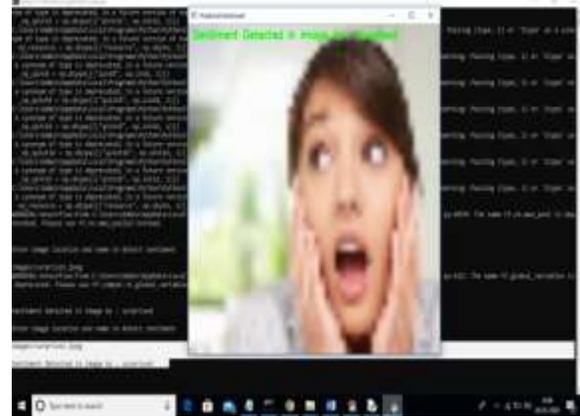


Fig.6: Prediction result

V. EXTENSION

In this project as extension work we are analysing input facial expression images to detect sentiments. To detect sentiment we trained CNN model with all possible facial expressions and then used OPENCV python API to detect faces from images. Detected faces will be input to CNN model to predict expression from it.

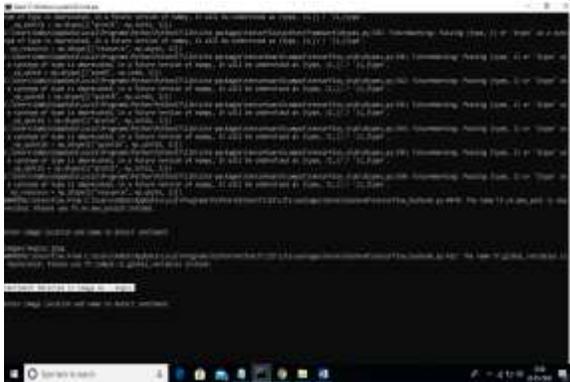


Fig.5: Sentiment prediction

VI. CONCLUSION

In this paper, we have proposed a model that extracts out the high sentiment cognition words from text in an automated way using fuzzy entropy measure for fuzzy reasoning. The fuzzy scores are extracted from SentiWordNet lexicon that has been compiled with the help of human annotations. We are interested in the cluster with low fuzzy entropy values since they correspond to the important words which contribute in evaluation of the sentiment of text. Our approach emulates the way human mind works by glancing at only a few key important words in text, to determine the polarity of review. These shortlisted words are trained using a LSTM network. It is observed that our model has achieved the highest accuracy as compared to other methods.

ACKNOWLEDGEMENT

The authors are greatly acknowledged DST-FIST(Govt. of India)for funding to setting the research computing facilities at RVR&JC College of Engineering.

REFERENCES

- [1] B. Pang, and L. Lee. "Opinion mining and sentiment analysis." Foundations and Trends® in Information Retrieval 2, no. 1–2 (2008): 1-135.
- [2] S. Sun, C. Luo, and J. Chen. "A review of natural language processing techniques for opinion mining systems." Information Fusion 36 (2017): 10- 25.

[3] Y. Long, Q. Lu, R. Xiang, M. Li, and C. Huang. "A cognition based attention model for sentiment analysis." In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pp. 462-471. 2017.

[4] N. Fang, X. Luo, and W. Xu. "Evaluating the capability of text understanding based on cognitive principle." In Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on, pp. 1110- 1114. IEEE, 2008.

[5] V. Demberg, and F. Keller. "Data from eye-tracking corpora as evidence for theories of syntactic processing complexity." Cognition 109, no. 2 (2008): 193-210.

[6] S. Susan, and J. Keshari. "Finding significant keywords for document databases by two-phase Maximum Entropy Partitioning." Pattern Recognition Letters 125 (2019): 195-205.

[7] D. Tang, B. Qin, and T. Liu. "Learning semantic representations of users and products for document level sentiment classification." In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), vol. 1, pp. 1014-1023. 2015.

[8] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy. "Hierarchical attention networks for document classification." In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1480- 1489. 2016.

[9] H. Chen, M. Sun, C. Tu, Y. Lin, and Z. Liu. "Neural sentiment classification with user and product attention." In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 1650-1659. 2016.

[10] L. Gui, R. Xu, Y. He, Q. Lu, and Z. Wei. "Intersubjectivity and Sentiment: From Language to Knowledge." In IJCAI, pp. 2789-2795. 2016.