ANALYSIS OF AUTOMATIC EMOTION RECOGNITION ON AUDIO-TEXTUAL DATA

UGC CARE APPROVED

ABSTRACT

Social media has rapidly developed in recent years, providing people with spontaneous possibilities to share their thoughts, opinions, and attitudes about numerous things in text, audio, and video formats. An effective platform has been developed by the World Wide Web 2.0, and many new avenues for interpersonal communication have been made available. Online forums where users can discuss different viewpoints, express their feelings and opinions, and read and post reviews of products, movies, and other media. The amount of data on social media is always expanding. Big data refers to the quantity, pace, and variety of this social media data. The user's thoughts, views, and attitudes on social media reflect their emotional state. For both, the government and businesses, analyzing the emotions portrayed by these social media data has yielded rich, insightful information. This paper offers the research relevant to this issue in order to accomplish that objective.

Keywords - *Emotion Detection, Sentiment analysis, Natural Language Processing, Opinion mining.*

Introduction

Sentiment analysis is one of the most extensively anticipated uses of machine learning and natural language processing (NLP) (ML). With the introduction of Web 2.0, this sector has experienced tremendous growth. Astonishing opportunities for individuals to express their views, opinions, visions, and attitudes about a product, person, or incident on a different aspect of life every day have been created by the accelerated evolution of the internet, which has given rise to websites, blogs, forums, news portals, and social networking sites. User emotion is present in these perspectives and opinions of people. Government, industry, and individuals can all greatly benefit from analyzing this feeling. Because of this, decision-making is a frontier for sentiment analysis. It can be used in a variety of applications, including feedback analysis and recommendations.

The internet offers a vast environment that allows individuals to interact and express themselves at any time using a variety of media, including text, speech, and other modalities. This increases the amount, speed, and variety of material in social media. Such a tremendous amount of information would be a waste of time, space, and effort to collect and analyze. Therefore, it appears to be quite challenging to satisfy the needs of an emotion identification system with just one emotion recognition model. To enhance the accuracy and performance of the emotion identification system, a multi-model system is therefore required. To predict the emotion expressed by the user through text and speech, a variety of machine learning techniques, including Support Vector Machine, Naive Bayes Classifier, Maximum Entropy, etc., and Deep Learning Techniques are employed. The power of data is being unlocked by business entities, the government, organizations, and individuals. In order to make rational and financially sound decisions, they examine the data to evaluate and comprehend client interests, needs, and behavior.

The main goal of this research article is to examine

A. SHUNMUGA SUNDARI
Part-Time Ph.D. Research Scholar (internal),
PG and Research Department of Computer Science, Rani Anna Government College for Women, Tirunelveli, Affiliated to Manonmaniam Sundaranar University, Abhishekapatti, Tirunelveli, Tamilnadu, India.

Dr. R. SHENBAGAVALLI
Assistant Professor (Research Supervisor)
PG and Research Department of Computer Science, Rani Anna Government College for Women, Tirunelveli, Tamilnadu, India.

and investigate several sentiment analysis methods that are used to infer emotions from user-depicted text and speech.

Literature Review

When upmarket and elegant products are available online, people tend to base their decisions more on emotional factors like brand, distinctiveness, and prestige than on rational factors like technical, functional, and cost [1]. If a buyer purchases a product and is emotionally delighted with it, he will also insist that another customer purchase the same thing. The fundamental objective of emotional marketing is to stir up the customer's emotions to persuade them to choose a specific brand or product to boost sales. The emotion detector's crucial objective is to give clients emotional assurance about a good or service [2].

In order to create a pleasant environment that provides a platform for offering only good news, this study analyses the news articles filters out the bad articles and provides only positive articles [3]. They initially performed a classification, followed by a feature extraction, to put this concept into action. For classification, the methods Naive Bayes, Support Vector Machine, and Maximum Entropy are employed. A model is then created utilizing the training data after the categorization. The new dataset is classified using it after it has been produced using the constructed model.

People now utilize a common technique known as "Sentiment analysis over Twitter data" to keep track of the public's perception of their brands and corporate entities. Due to its noise, Twitter data presents a significant difficulty for sentiment analysis. The removal of stop words [4] using pre-compiled collections of stop words or using dynamic stop-word recognition using more complex techniques is a well-known method to minimize the noise in text. The effectiveness of stopping stop-words in sentiment categorization is examined in this research to see if it helps or hinders it.

The purpose of this research [5] is to explore and model the connection between online documents and user-generated social emotions. This problem is known as social affective text mining. They hope to determine the relationship between affective phrases and social emotions in order to predict the mood of a text. The emotion term model, topic model, and emotion topic model are the methods used in this paper to do this.

The hybrid method outlined in this study [6] aims to capture the emotion from the communication medium. Humans can convey their emotions in a variety of ways, including words, pictures, facial expressions, and more. This work only addresses textual data because of the limited requirements for textual data. In order to identify the mood in the text, this study employs both machine learning-based methods and keyword-based methods.

The model for language learning sentiment analysis using a naive Bayes classifier was proposed in Paper [7]. As sample data in this work, Facebook status is used. The proposed approach predicts whether an emotion is expressed in a sentence at the sentence level, whether it is positive, negative, or neutral.

A novel method for emotion estimate from the user-entered text on social networking sites was put out by the author in the paper [8]. The author created a method for creating visual representations that are based on the text's emotional content.

This paper [9] proposed a method for automatically identifying the text's emotional state, which may then be used to depict facial expressions. Here, sample data comes from a corpus of children's stories. Using supervised machine learning, this system successfully assigns the children's stories to one of the predetermined emotion classes.

This paper uses Linguistic Inquiry and Word Count (LIWC) to analyze the emotion of a sentence as positive or negative [10]. This method uses exclamation marks and words to predict the emotion as positive and uses affective words to state it as a negative emotion.

This paper examines whether or not an emotion is present in the statement at first. Additionally, it defines whether the sentence is positive or negative. The sentence is finally divided into eight separate emotion classifications. The only issue they encountered was that a sentence's emotion prediction was not dependent on context [11].

This paper [12] tackles a variety of communication-related issues. It examines automatic dialogue systems that make use of prosodic elements to identify vocal emotions and finds that their predictions are inaccurate. The suggested method, named "Monitoring of User State Emotion," combines prosodic elements and knowledge sources to increase prediction accuracy and produce the best long-lasting model of communication problems.

A real-time CNN model is abandoned [13] to identify the emotion in speech. The suggested model correctly and quickly describes the emotion. It divides the predicted feeling into the three basic emotional categories of "Angry," "Happy," and "Sad." Future research will focus on accelerating the suggested model's ability to detect emotion in speech for applications involving real-time human-machine interaction.

This paper [14] focuses on the identification of emotions in multi-speaker speeches. They developed a graph-based convolutional neural network for conversations, known as ConGCN. Conversational language presents a special challenge in modeling context-sensitive dependency for emotion recognition compared to non-conversational text. When more speakers are participating in a conversation, it is more difficult to identify emotions than it is in two-speaker conversations. This is a result of the unique personality and speech characteristics that each speaker possesses, which have a significant influence on how emotions are expressed. Because of this, context-sensitive reliance is important, but it can be challenging to accurately model for one another's reliance on words for emotion detection in multi-speaker conversations.

This Paper [15] propounds a methodology to classify the emotions in speech independent of the speaker. Classify the emotion it has two levels of steps. At first, using acoustic features it classifies the emotion into six labels. For classification, it uses the Gaussian Mixture Model (GMM). Then it finds out the highest-rated emotion among all others by using the Sequential Floating Forward Selection (SFFS) algorithm.

To identify human emotion using massive data, this study [16] presented Deep learning aided semantic text analysis (DLSTA). In this study, DLSTA is used as a technique for locating human emotions in vast amounts of textual data. Word embedding is frequently used in natural language processing for a range of tasks, such as question-answering, emotion interpretation, and machine translation.

Text emotion identification may identify emotions, and it involves categorizing data directly using deep learning and natural language processing techniques. The suggested DLSTA algorithm accurately detects emotion in text data.

A novel neural network structure called SENN [17] is proposed in this paper to identify emotions in text. The SENN (Semantic-Emotion Neural Network) is a new neural network design that uses trained word representations to utilize either semantic/syntactic or emotional information. Writings today take many different forms, which include social media posts, microblogs, news stories, customer feedback, and so on. Text mining can be a useful tool for analyzing the content of these brief texts to find and disclose a number of properties, including emotions. Experimental results show that the SENN model surpasses most baseline methodologies and cutting-edge techniques.

An innovative ensemble classifier is presented in this paper [18] to extract emotions from textual contexts. The proposed model outperforms its competitors in separating emotions from noisy content. This study's architecture for emotion recognition through ensemble learning uses cognitive clues to distinguish between writers of brief messages. It also proposes a regression approach to identify latent author clues in short texts. Also proposed and developed was a multi-channel method for feature extraction based on attentional mechanisms and emotion lexicons. The preferred vectors are recovered using this approach, which fuses various embedding models. This work proposes a cognitively conscious aggregation function for the ensemble model, which aggregates the output of different base classifiers.

Conclusion

This paper has summarized the numerous studies on emotion prediction in textual and audio data using deep and machine learning approaches. Additionally, it evaluates each technique's effectiveness and effectiveness gaps in identifying the user's emotions from textual and speech data. It has been shown after examining all of these experiments that deep learning techniques perform better at properly predicting emotions than machine learning techniques. In a future research, we can present a fresh approach to determining emotion more accurately in the textual and audio domains.

References

1. *J. Kaur and J. R. Saini, "Emotion Detection and Sentiment Analysis in Text Corpus: A Differential Study with Informal and Formal Writing Styles", International Journal of Computer Applications, vol.101, issue 9, pp. 1-9, September 2014.*
2. *H. Binali, Chen W., V. Potdar, "Computational Approaches for Emotion Detection in Text",4th IEEE International Conference on Digital Ecosystems and Technologies, pp. 172-177, 2010.*
3. *Doddi, K.S. Haribhakta, M.Y.V. Kulkarni, P. (2014)"Sentiment Classification of News Articles", Kiran Shriniwas Doddi et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3), pp. 4621-4623*
4. *H. Saif, Y. He, H. Alani, and M. Fernandez, "On stopwords, ?ltering, and data sparsity for sentiment analysis of Twitter," in The International Conference on Language Resources and Evaluation, 2014.*
5. *K.Thendral, S.Chitrakala, "Emotion Recognition System for Affective Text", ISSN (Online): 2347 - 2812, Volume-2, Issue -11,12 2014*
6. *Yong-soo seol and Dong-joo kim, "Emotion recognition from text using Knowledge-based ANN", ITC-CSCC 2008*
7. *Christos Troussas, Maria Virvou," Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning", 2012.*
8. *Tejasvini Patil and Sachin Patil," Automatic generation of emotions for social networking websites using text mining" in proceedings of IEEE-2012*
9. *R. A. Calix, S. A. Mallepudi, B. C. B. Chen, and G. M. Knapp. Emotion recognition in text for 3-d facial expression rendering, 2010.*
10. *Hancock, J. T., Landrigan, C., Silver, C., "Expressing emotion in text-based communication", Proceedings of the SIGCHI conference on human factors in computing systems, ACM, San Jose, California, USA, pp. 929-932, April 2007.*
11. *Ghazi, D., Inkpen, D., Szpakowicz," Hierarchical versus flat classification of emotions in the text", Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text, Association for Computational Linguistics, vol. 3 issues 2, pp. 140-146, June 2010.*
12. *Batliner A, Fischer K, Huber R, Spilker J, North E (2003) How to find trouble in communication. Speech Commun 40: 117–143*
13. *Bertero, D., and Fung, P. 2017. A ?rst look into a convolutional neural network for speech emotion detection. In 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP), 5115– 5119. IEEE.*
14. *D. Zhang, L. Wu, C. Sun, S. Li, Q. Zhu, G. Zhou, Modeling both context-and speaker-sensitive dependence for emotion detection in multi-speaker conversations, in Proceedings of the Twenty-Eighth International Joint Conference on Arti?cial Intelligence, IJCAI), 2019, pp. 5415–5421.*
15. *Atassi H, Esposito A (2008) A speaker-independent approach to the classification of emotional vocal expressions. In: Proceedings of 20th IEEE international conference on tools with artificial intelligence, pp 147-152.*
16. *Jia Guo, "Deep learning approach to text analysis for human emotion detection from big data", DOI:10.1515/jisys-2022-0001, 2022.*
17. *E. Batbaatar, M. Li, K. Ryu, Semantic-emotion neural network for emotion recognition from text, IEEE Access 7 (2019) 111866–111878.*
18. *Sara Kamran, Raziyeh Zall, Mohammad Reza Kangavari, Saeid Hosseini, Sana Rahmani, and Wen Hua, EmoDNN: Understanding emotions from short texts through a deep neural network ensemble, arXiv:2106.01706v1 [cs.LG] 3 Jun 2021.*