JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Sentimental Analysis And Emotion Detection Using ML, DL And Random Under Sampling

[1]Prof. Santosh Kale, [2]Lalit Malviya, [3]Yash Mahajan, [4]Yash Desai, [5]Aayush Patil

[1]M.E.Computer Engineering,
[1]NBN Sinhgad Technical Institutes Campus, Pune, India
[2]Computer Engineering,
[2]NBN Sinhgad Technical Institutes Campus, Pune, India
[3]Computer Engineering,
[3]NBN Sinhgad Technical Institutes Campus, Pune, India
[4]Computer Engineering,
[4]NBN Sinhgad Technical Institutes Campus, Pune, India
[5]Computer Engineering,
[5]NBN Sinhgad Technical Institutes Campus, Pune, India

Abstract : The proliferation of textual data across digital platforms has given rise to the need for automated sentiment analysis and emotion detection. This project addresses this demand by proposing a comprehensive solution that leverages both traditional machine learning (ML) techniques and state-of-the-art deep learning (DL) architectures, while also mitigating class imbalance through random under sampling. In today's data-driven world, understanding the sentiments and emotions expressed in textual data is of paramount importance. This project presents a comprehensive solution to the challenges of sentiment analysis and emotion detection by harnessing the power of machine learning (ML), deep learning (DL), and strategic random under sampling. The explosion of textual data on digital platforms necessitates automated methods for sentiment analysis, involving the classification of text into positive, negative, or neutral sentiments. Furthermore, emotion detection requires the identification of specific emotions, such as happiness, anger, and sadness, within this text. These tasks are compounded by imbalanced datasets where some classes are significantly underrepresented. Accurate sentiment analysis and emotion detection have profound implications across various domains, including brand management, customer satisfaction analysis, and mental health monitoring. The models developed in this project will empower decision-makers with insights derived from textual data, enhancing the quality of user experiences and informed decision-making.

Keywords : Machine Learning , Deep Learning , Sentiment Analysis , Data Preprocessing , Convolutional Neural Networks , Haar Cascade , Emotion Detection.

I. INTRODUCTION

Sentiment analysis involves determining whether a piece of text, speech, or video conveys a positive, negative, or neutral sentiment. On the other hand, emotion detection delves deeper, aiming to identify specific emotional states, such as happiness, anger, sadness, or surprise. These techniques find applications in a wide range of domains, including marketing, customer feedback analysis, mental health monitoring, and entertainment. Sentiment analysis and emotion detection, powered by machine learning techniques, offer a means to unlock valuable insights from these rich sources of information. Sentiment analysis and emotion detection are widespread and diverse. They have revolutionized customer service by automatically assessing customer feedback and reviews. They're fundamental in tracking public opinion on social and political issues, enabling policymakers and businesses to make informed decisions. They are also invaluable in the field of market research, helping companies understand consumer behavior, preferences, and reactions.

II. LITERATURE SURVEY

1. Paper Name : Multi-Classifier Interactive Learning for Ambiguous Speech Emotion Recognition.
   Author : Ying Zhou, Xuefeng Liang (B) , Yu Gu, Yifei Yin, Longshan Yao.

Abstract : The combination of speech recognition and speech emotion recognition can improve the feedback efficiency and the quality of service. Thus, the speech emotion recognition has been attracted much attention in both industry and academic. Since emotions existing in an entire utterance may have varied probabilities, speech emotion is likely to be ambiguous, which poses great challenges to recognition tasks. However, previous studies commonly assigned a singlelabel or multi-label to each utterance in certain. Therefore, their algorithms result in low accuracies because of the inappropriate representation. Inspired by the optimally interacting theory, we address the ambiguous speech emotions by proposing a novel multi-classifier interactive learning (MCIL) method. In MCIL, multiple different classifiers first mimic several individuals, who have inconsistent cognitions of ambiguous emotions, and construct new ambiguous labels (the emotion probability distribution). Then, they are retrained with the new labels to interact with their cognitions.

2. Paper Name : Improved Cross-Corpus Speech Emotion Recognition Using Deep Local Domain Adaptation.
Author : ZHAO Huijuan1,2, YE Ning3 , and WANG Ruchuan.
Abstract : Due to the scarcity of high-quality labeled speech emotion data, it is natural to apply transfer learning to emotion recognition. However, transfer learningbased speech emotion recognition becomes more challenging because of the complexity and ambiguity of emotion. Domain adaptation based on maximum mean discrepancy considers marginal alignment of source domain and target domain, but not pay regard to class prior distribution in both domains, which results in the reduction of transfer efficiency. In order to address the problem, this study proposes a novel cross-corpus speech emotion recognition framework based on local domain adaption. A category-grained discrepancy is used to evaluate the distance between two relevant domains. According to research findings, the generalization ability of the model is enhanced by using the local adaptive method. Compared with global adaptive and non-adaptive methods, the effectiveness of cross-corpus speech emotion recognition is significantly improved.

3. Paper Name : Multi-View Speech Emotion Recognition Via Collective Relation Construction.
Author : Mixiao Hou , Zheng Zhang.
Abstract : The discriminative knowledge from speech for effective emotion recognition may come from multiple physical properties such as energy spectrum, frequency, prosody, which could be collected as multi-view representations. However, the current works fail to fully explore the underlying interactive relations among multiple speech representations for emotion recognition. In this paper, we propose a novel Collective Multi-view Relation Network (CMRN) to exploit the intrinsic characteristics of multi-view speech representations for discriminative speech emotion recognition. Generally, the proposed CMRN consists of three sub-networks, i.e., view-specific attention network, multi-view shared attention network and collective relation network. Specifically, the view-specific attention network is designed to excavate the distinguishable view-specific features deduced from the original speech. By contrast, the multi-view shared attention network is conceived to capture the collaborative knowledge from multiple views.

4. Paper Name : Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model.
Author : NAILA ASLAM 1,FURQAN RUSTAM 2, ERNESTO LEE 3.
Abstract : The cryptocurrency market has been developed at an unprecedented speed over the past few years. Cryptocurrency works similar to standard currency, however, virtual payments are made for goods and services without the intervention of any central authority. Although cryptocurrency ensures legitimate and unique transactions by utilizing cryptographic methods, this industry is still in its inception and serious concerns have been raised about its use. Analysis of the sentiments about cryptocurrency is highly desirable to provide a holistic view of peoples' perceptions. In this regard, this study performs both sentiment analysis and emotion detection using the tweets related to the cryptocurrency which are widely used for predicting the market prices of cryptocurrency. For increasing the efficacy of the analysis, a deep learning ensemble model LSTM-GRU is proposed that combines two recurrent neural networks applications including long short term memory (LSTM) and gated recurrent unit (GRU).

III. Methodology

A. CNN

A Convolutional Neural Network (CNN) is a type of deep neural network primarily designed for processing and analyzing visual data, including images and videos. CNNs have gained immense popularity in various applications, from computer vision tasks to natural language processing.

B. Haar Cascade

Cascading classifiers are trained with positive sample views of a particular object and arbitrary negative images of the same size. The Haar Cascade is a machine learning object detection algorithm used to identify objects or regions of interest in images or video.

IV. RELEVANT MATHEMATICS ASSOCIATED WITH THE PROJECT

System Description:

- Input: Live Camera/Audio
- Output: Person Sad/ Happy
- Let S be the Whole system S= I,P,O
- I-input
- P-procedure
- O-output
- Input(I) I=Live Camera/Audio
- Where,
- Procedure (P),
- P=I, Using I System perform operations and calculate the prediction
- Output(O)-O=Person Sad / Happy

V. Major Constraints

Our software has many quality attribute that are given below:-

1. Adaptability: This software is adaptable by all users.
2. Availability: This software is freely available to all users. The availability ofthe software is easy for everyone.
3. Maintainability: After the deployment of the project if any error occurs then itcan be easily maintained by the software development.
4. Reliability: The performance of the software is better which will increase thereliabilityof the Software.
5. User Friendliness: Since, the software is a GUI application; the output generated is much user friendly in its behavior.
6. Integrity: Integrity refers to the extent to which access to software or data byunauthorized persons can be controlled
7. Security: Users are authenticated using many security phases so reliable security is provided.
8. Testability: The software will be tested considering all the aspects.

VI. Motivation Of The Project

The project aims to develop a machine learning model that can analyze and detect sentiments and emotions in social media posts, such as tweets or Facebook updates. This tool could be used to monitor public sentiment and emotional trends about a particular topic, brand, or event in real-time.
Gather a large dataset of social media posts containing text and associated labels for sentiment and emotions.
Clean and preprocess the data by removing special characters, emojis, and irrelevant information.
Sentiment analysis can be applied to monitor public sentiment on political issues, social causes, and current events.
Emotion detection can be used in healthcare to monitor and assess patients' emotional states. VII.
Sentiment analysis assists in tracking customer sentiment in e-commerce reviews and chat conversations with customer support.

VIII. METHODOLOGIES OF PROBLEM SOLVING AND EFFICIENCY ISSUES

Gather a dataset containing text samples labeled with sentiment scores or emotion categories. Perform stemming or lemmatization to reduce words to their base forms. Extract relevant features from the text data. For sentiment analysis, you may consider features like the sentiment of specific words or phrases. Split your data into training, validation, and test sets. Assess your model's performance on the test set to understand its generalization capabilities. Regularly update your model with fresh data to adapt to evolving language and sentiment patterns.

Figure 3.1: Login Page

A. In figure 3.1. On the login page, initial account creation is the first step. Subsequently, users can sign in by inputting the previously established username and password.

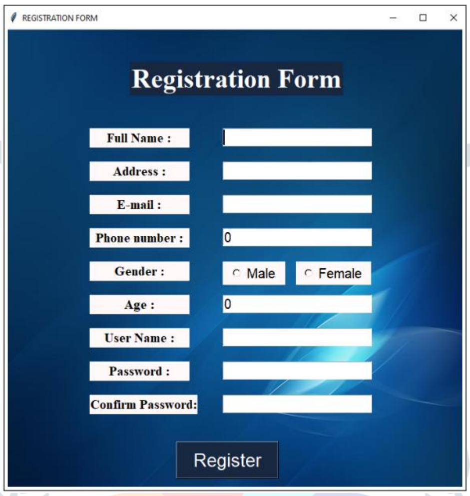

Figure 3.2: Registration Page

B. In Figure 3.2. On the registration page, we enter our personal details to create our account.

Figure 3.3: GUI Page

C. In figure 3.3. On our main GUI page, users can log in, register, or exit the application.

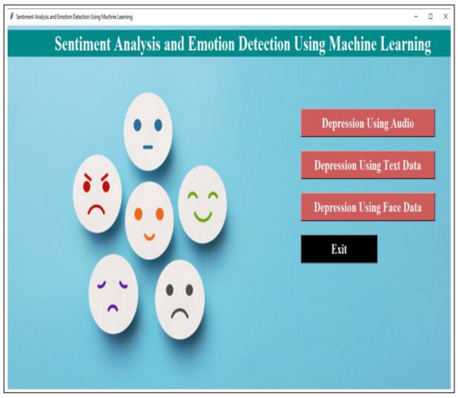

Figure 3.4: Master

D. On the Master page, we analyze text, audio, and facial expressions to predict a person's emotions, including happiness, sadness, and anger.

Figure 3.5: Output Page 1 (Emotion detection through Face Recognition)

E. This face recognition system page predicts your facial expression and then gives the desired output of your predicted face.

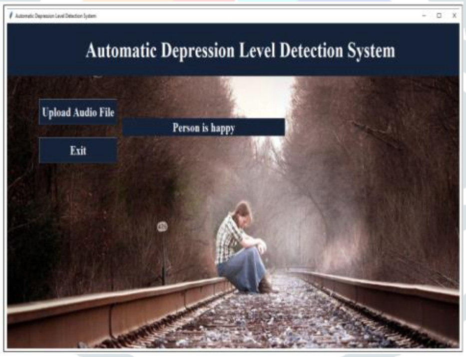

Figure 3.6: Output Page 2 (Emotion detection through Audio)

F. This automatic depression level detection system predicts your audio that whatever you speak and by that audio the system pedicts your emotion detection.

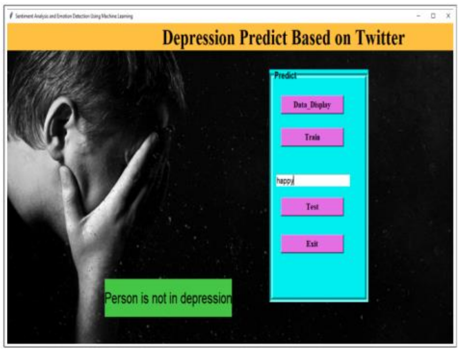

Figure 3.7: Output Page 3 (Emotion detection through Text)

G. By this page whatever you write your feelings in words it will predict that words and will give the desired output.

IX. SCENARIO IN WHICH MULTI-CORE, EMBEDDED AND DISTRIBUTED COMPUTING USED

Multi-core computing allows you to leverage the power of multiple CPU cores to speed up data preprocessing and model training. Utilize multi-threading or multiprocessing in deep learning libraries like TensorFlow or PyTorch to train models concurrently on different cores.

You want to deploy sentiment analysis and emotion detection models on resourceconstrained devices, such as IoT sensors or edge devices, to perform local analysis without relying on cloud servers. Implement quantization and model compression techniques to reduce the model's size and memory requirements.

You need to scale sentiment analysis and emotion detection to process a massive volume of data or serve a large number of users.Deploy the sentiment analysis and emotion detection models on a cluster of machines, and use load balancers to distribute incoming requests.

X. OUTCOME

1. Sentiment analysis can be applied to customer reviews, social media comments, and survey responses to understand how customers feel about a product, service, or brand.

2. Emotion detection can help companies track and manage their brand's reputation by identifying trends in positive and negative sentiment.

3. Sentiment analysis can be used in market research to gain insights into consumer preferences and market trends.

XI. Application

Companies use sentiment analysis to track and analyze social media posts and comments to gauge public opinion about their products, services, or brand. Businesses can automatically analyze customer reviews, surveys, and feedback forms to identify positive and negative sentiments.

XII. Goals and objectives

Sentiment analysis has numerous real-world applications, including analyzing customer reviews, social media comments, and news articles to understand public opinion, market sentiment, or customer satisfaction.
To emotion detection and sentiment analysis are used in market research to gauge public opinion about a product, service, or a new concept.
To social media platforms generate massive amounts of data. To sentiment analysis can be used to automatically categorize customer feedback and support requests, allowing companies to prioritize and address critical issues more effectively.

XIII. Statement of scope

Develop and train machine learning models for sentiment analysis.
Create models for emotion detection, which goes beyond sentiment analysis by identifying specific emotions or more complex emotional states.

XIV. Risk Analysis

The risks for the Project can be analyzed within the constraints of time and quality

ID	Risk Description	Probability	Impact		
			Schedule	Quality	Overall
1	Description 1	Low	Low	High	High
2	Description 2	Low	Low	High	High

Figure XIII.1: Risk Table

Probability	Value	Description
High	Probability of occurrence is	$> 75\%$
Medium	Probability of occurrence is	$26 - 75\%$
Low	Probability of occurrence is	$< 25\%$

Figure XIII.2: Probability Definitions

Impact	Value	Description
Very high	$> 10\%$	Schedule impact or Unacceptable quality
High	$5 - 10\%$	Schedule impact or Some parts of the project have low quality
Medium	$< 5\%$	Schedule impact or Barely noticeable degradation in quality Low Impact on schedule or Quality can be incorporated

Figure XIII.3: Risk Impact Definitions

XV. State Transition Diagram

State Transition Diagram Fig.XIV.1 A state diagram, also known as a state machine diagram, is a type of UML (Unified Modeling Language) diagram used to represent the behavior of a system or an entity in response to events. State diagrams are particularly useful for modeling the lifecycle of an object, the behavior of a software system, or the states of a finite-state machine.

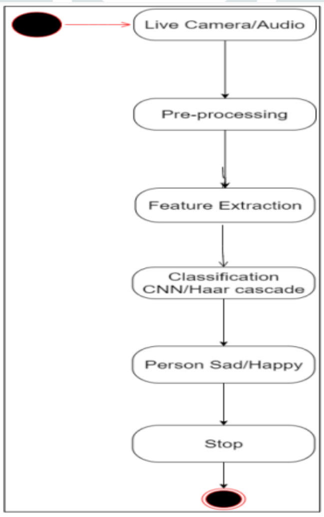

Figure XIV.1: State Transition Diagram

XVI. Result

This segment contains the findings as well as discussion. By laying out the computer hardware as well as software set up used for the testing. Later, this discuss numerous assessment methods and performance of our model in relation to them. We used a variety of performance measurements, including precision, recall, Fmeasure and also compared different ML classifiers. Sentiment Analysis findings are very affected by a number of things, including data pretreatment. Another critical component is the choice of a classification algorithm to train and test the Twitter data. This examined the data with a variety of classifiers, including SVM, naive Bayes, and others, to determine the best classifier. XGB classifier outperformed other classifiers with respect to accuracy.

XVII. Conclusion

Sentiment Analysis and Emotion Detection are crucial applications of machine learning that provide valuable insights into understanding human emotions, opinions, and attitudes. These techniques have found wide-ranging applications across various domains, including social media monitoring, customer feedback analysis, market research, and healthcare.Sentiment analysis and emotion detection using machine learning have become indispensable tools for businesses and organizations to gain insights into customer opinions and emotional states. These techniques have evolved significantly, thanks to advancements in machine learning and NLP. However, there is still room for improvement, particularly in handling complex language nuances and context. In summary, sentiment analysis and emotion detection using machine learning have a profound impact on various aspects of human life and business. They continue to be an active area of research, with ongoing efforts to enhance the accuracy and applicability of these methods. As technology advances, we can expect even more sophisticated and context-aware sentiment analysis and emotion detection systems to emerge.

XVIII. References

[1]. X. Ma, Z. Wu, J. Jia, M. Xu, H. Meng, And L. Cai, “Speech Emotion Recognition With Emotion-Pair Based Framework Considering Emotion Distribution Information In Dimensional Emotion Space,” In Proc. Interspeech, 2017, Pp. 1238–1242.

[2]. T. L. Nwe, S.W. Foo, And L. C. De Silva, “Speech Emotion Recognition Using Hiddenmarkovmodels,” Speechcommun., Vol. 41, No. 4, Pp. 603–623, 2003.

[3]. B.-H. Su, C.-M. Chang, Y.-S. Lin, And C.-C. Lee, “Improving Speech Emotion Recognition Using Graph Attentive Bi-Directional Gated Recurrent Unit Network,” In Proc. Interspeech, 2020, Pp. 506–510.

[4]. S. Mao, P. Ching, And T. Lee, “Deep Learning Of Segment-Level Feature Representation With Multiple Instance Learning For Utterance-Level Speech Emotion Recognition,” In Proc. Interspeech, 2019, Pp. 1686–1690.

[5]. B. Liang, H. Su, L. Gui, E. Cambria, And R. Xu, “Aspect-Based Sentiment Analysis Via Affective Knowledge Enhanced Graph Convolutional Networks,” Knowl.-Based Syst., Vol. 235, 2022, Art. No. 107643.

[6]. M. Hasan, E. Rundensteiner, and E. Agu, “Automatic emotion detection in text streams by analyzing Twitter data,” Int. J. Data Sci. Anal., vol. 7, no. 1, pp. 35–51, Feb. 2019.

[7]. S. Sharifirad, B. Jafarpour, and S. Matwin, “How is your mood when writing sexist tweets? Detecting the emotion type and intensity of emotion using natural language processing techniques,” 2019, arXiv:1902.03089.

[8]. F. Calefato, F. Lanubile, and N. Novielli, “EmoTxt: A toolkit for emotion recognition from text,” in Proc. 7th Int. Conf. Affect. Comput. Intell. Interact. Workshops Demos (ACIIW), Oct. 2017, pp. 79- 80.

[9]. S. A. Salam and R. Gupta, “Emotion detection and recognition from text using machine learning,” Int. J. Comput. Sci. Eng., vol. 6, no. 6, pp. 341–345, Jun. 2018.

[10]. H. Xu, H. Zhang, K. Han, Y. Wang, Y. Peng and X. Li, ”Learning alignment for multimodal emotion recognition from speech”, Proc. INTERSPEECH, pp. 3569.3573, 2019.