

DEPRESSION DETECTION BY ANALYZING SOCIAL MEDIA POST OF USER: A REVIEW

¹Gayatri Shinde, ²Priyanka Date, ³Swati Hire, ⁴Punam Sanap,

^{1,2,3,4}Student, ⁵Professor, ^{1,2,3,4,5}Computer Engineering,

^{1,2,3,4,5}Matoshri College of Engineering and Research Centre, Nashik, Maharashtra, India.

Abstract. Nowadays the problem of early depression detection is one of the most essential withinside the concern of psychology. Mental health issues are widely not unusual to place as one of the most prominent health stressful conditions withinside the world, with over 3 hundred million humans currently affected by depression alone. With huge volumes of man or woman-generated records on social networking platforms, researchers are growing variety the use of gadgets gaining know-how to determine whether or not or now no longer this content material cloth can be used to find out highbrow health problems in clients. Depression is a disorder that has been a superb concern in our society and has been continuously a heating concern relying on researchers withinside the world. Despite the huge quantity of assessments on know-how character moods together with depression, anxiety, and strain supported hobby logs collected thru pervasive computing devices like smartphones, foretelling depressed moods continues to be an open question. Social networks assessment is widely executed to address this problem. In this paper, we have got were given proposed a depression assessment and a suicidal ideation detection system, for predicting the suicidal acts that supported the extent of depression. The present examination aims to make the maximum tool for getting to know techniques for detecting a possible depressed Social Media man or woman in his/her Posts. For this purpose, we knowledgeable and tested classifiers to differentiate whether or not or now no longer someone is depressed or now not the use of competencies extracted from his/her sports activities withinside the posts. kind tool algorithms are used to train and classify it in Different tiers of depression on a scale of 0-100%. Also, records end up collected withinside the form of posts and have been categorized into whether or not or now no longer the most effective that tweeted is in depression or now not the use of kind algorithms of Machine Learning In this way Predictive method for early detection of depression or exceptional highbrow illnesses. This examination's number one contribution is the exploration of a network of competencies and its impact on detecting the Depression degree. This examination aims to increase a deep getting to know the model to categorize clients with depression via a couple of instances getting to know, that would study from man or woman-degree labels to find out post-degree labels. By combining every possibility of posts label category, it can generate temporal posting profiles that would then be used to categorize clients with depression.

This paper shows that there are smooth versions in posting patterns amongst clients with depression and non-depression, that's represented thru the combined opportunity of posts label category. In this study, the tool gaining know-how is used to system the scrapped records collected from social media clients' posts. Natural Language Processing (NLP), categorized the use of the BERT set of regulations to find out depression probably in a greater on hand and inexperienced way

Keywords: Machine Learning, NLP, BERT Algorithm, Depression, Classification, Social Media Post.

1 Introduction

1.1 Introduction:

Nowadays the trouble of early depression detection is one of the most essential withinside the location of psychology. Depression is furthermore a mean highbrow hassle. In today's world, the stresses of lifestyles activities in one's life should grow the opportunities for depression. Over 350 million humans globally are bothered via way of means of depression, which is about 5% of the entire population. Close to 8,00,000 humans die due to suicide each year and its miles are statistically the second most important motive for lack of lifestyle among humans 15–29 years old. At the same time, the major type of suicide is associated with depression. Recent researches display that depression is also the number one motive for disability and severe somatic diseases. The proliferation of the internet and communication technologies, specifically online social networks have rejuvenated how humans have interacted and talked with each distinct electronically. The programs which include Facebook, Twitter, Instagram, and alike now do not handiest host the written and multimedia content but moreover offer their clients particular their feelings, emotions, and sentiments about a topic, issue, or a hassle online. On one hand, this is remarkable for clients of social networking websites to overtly and freely make contributions and respond to any issue counted online; on the other hand, it creates opportunities for humans running withinside the health area to get the notion of what might be happening at the highbrow united states of America of someone who reacted to a topic in a selected way. To provide such a notion, tool gaining knowledge of techniques should possibly offer some particular features that would assist in analysing the particular patterns hidden in online communication and method them to show the highbrow united states of America (which include 'happiness', 'sadness', 'anger', 'anxiety, depression) among social networks' clients. Moreover, there is a growing body of literature addressing the characteristic of social networks in the form of social relationships which include cut-up relationships, highbrow illness ('depression', 'anxiety', 'bipolar' etc.), smoking, and consuming relapse, sexual harassment and for suicide ideation. Younger adults, racial/ethnic minorities, crucial staff, and unpaid man or woman caregivers counselled

having intimate disproportionately worse recognition outcomes, raised substance use, and elevated volatile cerebration. Youth is printed as age fifteen to twenty-four years, and it includes middle and past due adolescence. It's characterized by the manner of approach of contemporary adjustments in physical, psychological, and social dimensions. For healthy growth and development, teens want to have the handiest way of happiness, love, action, and independence and to have a cause in life. Throughout this herbal method stage, many kinds of behaviour square diploma advanced which could motive each normalcy or recognition malady. Depression can motivate the affected person to undergo drastically and perform poorly at work, at college, and in social activities. Irrespective of what you nearly simply did presently on your cell phone or PC, it's apparent that social media modified right into concern. Did you seize up with friends on Facebook, put up snapshots of your cat, or video of your totter taking walks for 1st time on Instagram? likely a Twitter link delivered to you here. These days parents tend to unique their emotions, opinions companion degrees reveal their regular lives through the boom of social media structures like Twitter, Facebook, and In. Instagram. These expressions square tiers typically through snapshots, videos, and general posts. In this study, we intend to analyse Social Media posts to come across any factors that would mirror the depression of relevant Social Media clients. Various tools for gaining knowledge of techniques are employed for such purposes. Considering the essential factor purpose of this study, the following are subsequent research disturbing conditions addressed withinside the paper. We will be predisposed to tend to aim to utilize tools gaining knowledge of techniques and algorithms for depression detection on social media posts of clients.

1.2 NLP (Natural Language Processing):

The art work described in this paper belongs to the vicinity of Natural Language Processing (NLP) and text class in particular. The origins of text class duties can be decided in early research to mechanically categorize documents based mostly on a statistical assessment of unique clue terms in 1961. Later, similar research wants to bring about rule-based text class systems like CONSTRUE in 1990, and finally, the field started to shift increasingly to machine getting to know algorithms throughout the year 2000. In addition to text categorization, machine getting to know turns into moreover using stress in ONE-OF-A-KIND text-based duties like sentiment assessment, which is focused on extracting reviews and sentiment from text documents. It turns into first applied in combination with the machine getting to know to discover high-quality or horrific reviews in movie critiques and end up then extended to one-of-a-kind assessment domains, further to virtually unique areas like social media monitoring and favored assessment of client attitudes. More recently, deep getting to know has been carried out for text class in addition to its greater common region usages withinside the image class. State-of-the-art work consequences in several text-based duties could, for example, be done thru transfer getting to know techniques

like Universal Language Model Fine-tuning (ULM Fit) and the Google research task Bidirectional Encoder Representations from Transformers (BERT) for the training of language representations, which includes ULM Fit and several one-of-a-kind techniques. The code of BERT and several pre-professional models.

2 Literature Review:

Instrumental opportunities of studying the conduct of customers in social networks are actively developing. In particular, strategies of computational linguistics are efficiently utilized in studying the posts from social media.

1) A records-analytic-primarily based totally version to hit upon melancholy of any individual is proposed withinside the paper. The records are gathered from the customers' posts on famous social media websites: Twitter and Facebook. In this research, device studying is used to manner the scrapped records gathered from SNS (Social Networking Sites) customers. Natural Language Processing (NLP), labelled the usage of Support Vector Machine (SVM) and Naïve Bayes set of rules to hit upon melancholy probably in an extra handy and greenway. [1]

2) The research employs Natural Language Processing (NLP) strategies to increase a melancholy detection set of rules for the Thai language on Facebook in which human beings use it as a device for sharing opinions, feelings, and existence events. [2]

3) The fitness tweets are analysed for Depression, Anxiety from the blended tweets via way of means of the usage of Multinomial Naive Bayes and Support Vector Regression (SVR) Algorithm as a classifier in paper [3].

4) In the paper, researchers gift a way to discover the melancholy degree of someone via way of means of looking at and extracting feelings from the text, the usage of emotion theories, device studying strategies, and herbal language processing strategies on unique social media platforms. [4]

5) The paper, pursuits to use herbal language processing on Twitter feeds for engaging in emotion evaluation specializing in melancholy. Individual tweets are labeled as impartial or negative, primarily based totally on a curated phrase listing to hit upon melancholy tendencies. In the manner of sophistication prediction, a guide vector device and Naive-Bayes classifier had been used. The consequences had been offered the usage of the number one category metrics inclusive of F1-score, accuracy, and confusion matrix. [5]

6) The paper, proposes a melancholy evaluation and suicidal ideation detection system, for predicting suicidal acts primarily based totally on the extent of melancholy. Real-time records changed into gathered

withinside the shape of Tweets and Questionnaires. Then, category device algorithms are used to teach and classify it in 5 degrees of melancholy relying on severity. [6]

7) Yates et al. used a neural community version to show the dangers of self-damage and melancholy primarily based totally on posts from Reddit and Twitter and confirmed the excessive accuracy of this diagnostic approach. The authors suggest that proposed strategies may be used for large-scale research of intellectual fitness in addition to for scientific treatment. [8]

8) O'Dea et al. tested that Twitter is gradually researched as a way for spotting mental well-being status, inclusive of melancholy and suicidality withinside the population. Their research found out that it's miles workable to understand the extent of fear amongst suicide-associated tweets, using each human coder and a programmed device classifier.[10]

There is a severe and developing variety of methodologies and strategies for detection of the melancholy degree from the posts on Social Media networks. In our study, we consolidate a technical description of strategies implemented for melancholy identity the usage of the Natural Language Processing approach labeled the usage of the BERT set of rules to hit upon melancholy. The framework is created from Data pre-processing step, the Feature extraction step following the Machine Learning classifiers, the Feature evaluation of the records, and the Experimental Results

3 Problem Statement:

Depression has been verified to have an impact at the language of individuals. To expand a software to investigate and hit upon depression of social media posts of clients thru tool learning techniques. This assignment pursuits to use natural language processing, tool learning techniques, and neural network architectures to build, tune, and have a look at models that classify social media Post of Users as "depressed" or "non-depressed."

4 Objective:

The targets are as follows:

1. System will constantly hold on tracking the posts and chats of users. And if it detects the bad notion sort of conduct then the device will mechanically put up the high-quality put up on his/her wall primarily based totally on the extent of depression.
2. Help the individual to pop out of depression

5 Methodology:

Machine Learning Classification Techniques used for the mode

1. BERT Algorithm:

“BERT stands for Bidirectional Encoder Representations from Transformers. It is designed to pre-educate deep bidirectional representations from the unlabeled textual content with the aid of using collectively conditioning on each left and proper context. As a result, the pre-skilled BERT version may be fine-tuned with simply one extra output layer to create today's fashions for a huge variety of NLP tasks.” We advanced a Bidirectional Encoder Representations from Transformers (BERT)-primarily based totally version, that's a brand-new language illustration version as defined in. As the call suggests, it changed into designed to pre-educate deep bidirectional representations that may be fine-tuned with an extra output layer. For this project, this outputs layer - a pooled output - changed into used for the binary type of the comments. From the various pre-skilled fashions available, we selected the English-language uncased (all lowercase earlier than tokenization) version of BERT, as case records isn't always specially essential to the venture of social media remark type

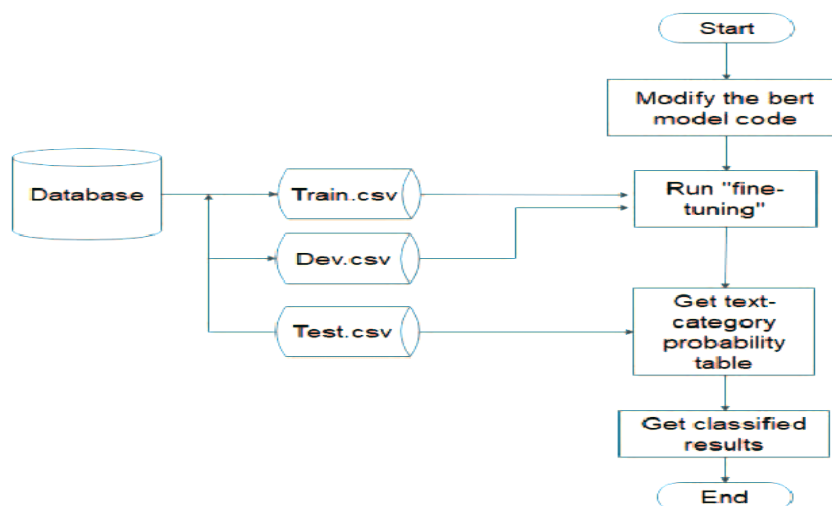


Fig.BERT Algorithm

6 Mathematical Model:

Relevant Mathematics Associated with The Project:

System Description:

S= I, O, F, DD, NDD, Failure, Success

Where,

S=System

I= Input

O=Output

F=Failure

S=Success

I is Input of system

Input I = set of Inputs

Where,

I= {Users Social media posts}

F is Function of system

F = set of Function

Where,

F1= {Input Dataset}

F2= {Json to CSV Conversion}

F3={Pre-processing}

F4={Cleaning}

F5= {Train test split}

F6= {Sentiment Dictionary}

F7= {Classifier (BERT Algorithm)}

F8={Tokenization}

is Output of system

Output O1= {Depression detection}

Success Conditions: Product working Smoothly. depression detection successfully.

Failure Conditions: if internet connection Unavailable.

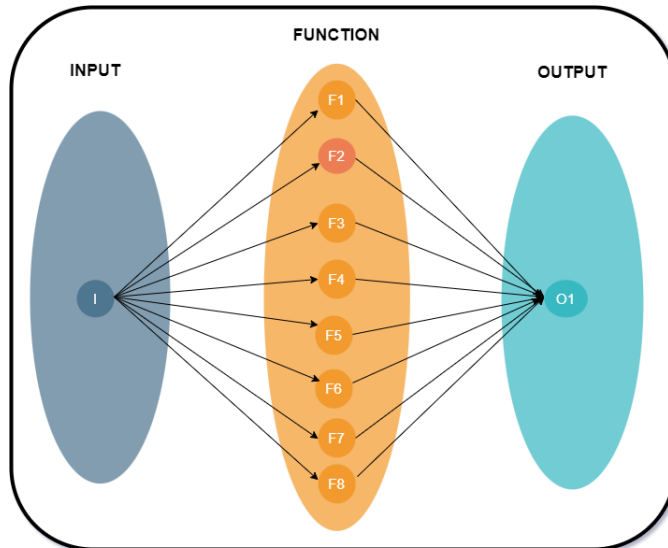


Fig.Venn Diagram

Where,

$I = \{\text{Users Social media posts}\}$

$F1 = \{\text{Input Dataset}\}$

$F2 = \{\text{Json to CSV Conversion}\}$

$F3 = \{\text{Pre-processing}\}$

$F4 = \{\text{Cleaning}\}$

$F5 = \{\text{Train test split}\}$

$F6 = \{\text{Sentiment Dictionary}\}$

$F7 = \{\text{Classifier (BERT Algorithm)}\}$

$F8 = \{\text{Tokenization}\}$

7 System Architecture:

Depression is a severe task in private and public fitness. One of the predominant answers to this hassle is an in-depth look at an individual's conduct attributes. These attributes are to be had on diverse social networking websites together with Facebook, Twitter, Instagram, etc. Social networking platform is the pleasant manner to recognize a person's conduct, wondering style, mood, egoistic networks, evaluations, etc. The use of social networking websites is growing in particular the various younger generation. The humans on social media-specific their feelings, day by day sports, evaluations approximately diverse topics, etc. So social networking websites are used as screening equipment to are expecting melancholy

stages. These social networking systems provide a person's experiences, evaluations, socialization, personality. The advanced technique of analysis of the affected person isn't always so applicable however the use of consumer-generated content material on social media posts enables one to be expecting the intellectual fitness stages and melancholy of a specific individual. Our challenge goal is to extract facts from social media posts and via way of means of having a clean knowledge of a person's behavioral attributes and tried questionnaires, melancholy stages of the consumer are predicted. A quantitative look is carried out to teach and check diverse device mastering classifiers to decide whether or not a social media submit of the consumer is depressed, from posts initiated via way of means of the consumer or his/her sports on social media. The following parent illustrates the melancholy detection of the use of the pastime and content material capabilities type model. First, all tweets for depressed and non-depressed accounts, in addition to facts of consumer accounts and sports together with the number of followers, the quantity of following, time of posts, the number of mentions, and quantity of reposts, are retrieved. Next, all posts of an account are assembled in a single file. Text pre-processing is carried out on all documents. First, a corpus is created and posts in every file are tokenized. BERT Classification Algorithm may be used.

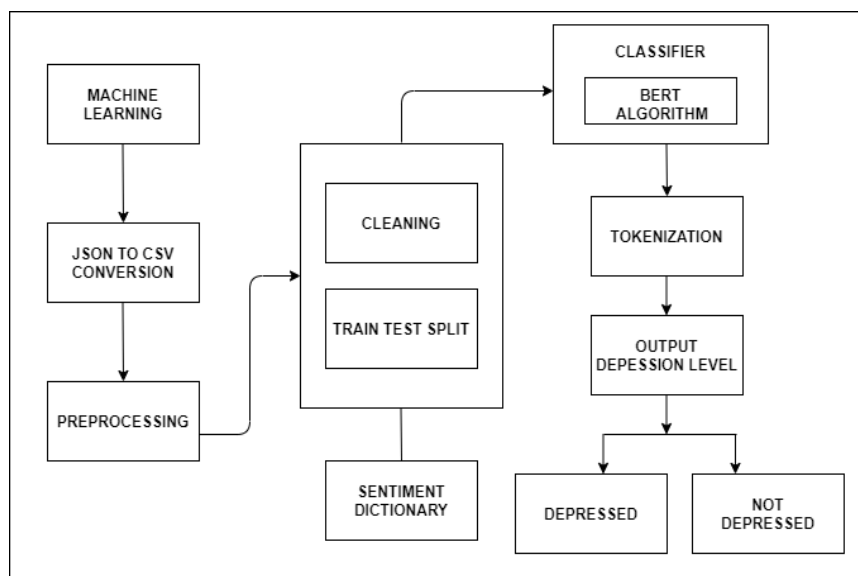


Fig.System Architecture

8 Existing System:

The already present device affords a far smooth flowing manner to decide the melancholy stage of customers the use of the Naïve Bayes algorithm. The extraction of textual information is finished through the extraction elegance from Facebook with the assist of the Facebook Graph API. After

extraction, the information is pre-processed. The lacking or repetitive attributes are looked after in pre-processing. Techniques like tokenization, decrease case conversion, and phrase stemming and phrases elimination are used for Pre-processing of information. In the proposed device, consistent with customers' Facebook submit version can discover whether or not he/she is depressed or not. But simplest reading posts won't deliver correct outcomes so we additionally examine the feedback through the consumer and his buddies and his chats also are analysed because the consumer will sincerely proportion his melancholy together along with his friend. On foundation of those analyses, the customers may be labelled as burdened and non-burdened.

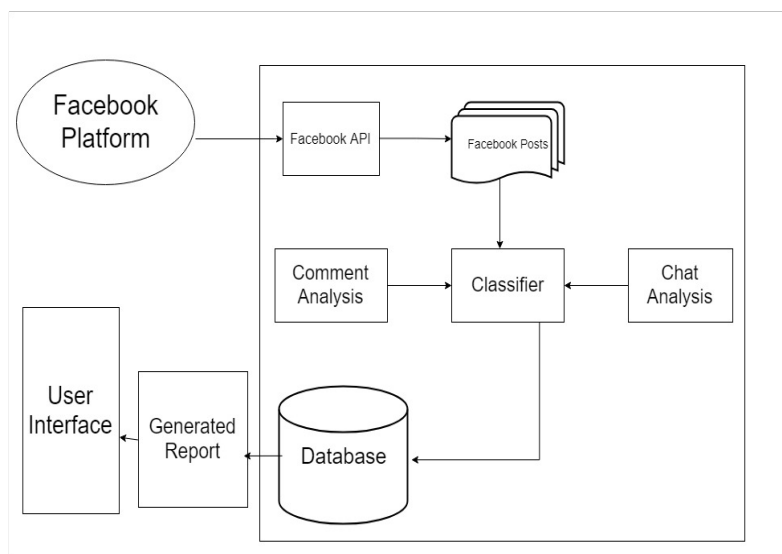


Fig.Exisiting System

9 Conclusion:

The proposed system may help the suspected purchaser to save his/her life, through manner of approach of know-how in advance whether or not or now no longer the customer is depressed or maybe the system will deliver some motivational posts to the customer based mostly on the volume of his depression. We end the system can be very useful in today's global wherein most humans don't have time to fulfill our friends, percent their thoughts and feelings as we applied in older days due to busy schedules. So, our system plays a important role proper right here to avoid any unwanted human loss. The system will inform their family individuals or spouses and kids approximately the situation of a depressed man or woman. So that very own family or buddy circle will help the man or woman to come out of depression.

10 References:

- [1] N. A. Asad, M. A. Mahmud Pranto, S. Afreen and M. M. Islam, "Depression Detection by Analyzing Social Media Posts of User," 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), Dhaka, Bangladesh, 2019, pp. 13- 17, doi: 10.1109/SPICSCON48833.2019.9065101.
- [2] K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot and Y. Kaewpitakkun, "Facebook Social Media for Depression Detection in the Thai Community," 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), Nakhonpathom, 2018, pp. 1-6, doi: 10.1109/JCSSE.2018.8457362.
- [3] P. Arora and P. Arora, "Mining Twitter Data for Depression Detection," 2019 International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2019, pp. 186-189, doi: 10.1109/ICSC45622.2019.8938353.
- [4] A. U. Hassan, J. Hussain, M. Hussain, M. Sadiq and S. Lee, "Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression," 2017 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, 2017, pp. 138- 140, doi: 10.1109/ICTC.2017.8190959.
- [5] M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 858-862, doi: 10.1109/ISS1.2017.8389299.
- [6] S. Jain, S. P. Narayan, R. K. Dewang, U. Bhartiya, N. Meena and V. Kumar, "A Machine Learning based Depression Analysis and Suicidal Ideation Detection System using Questionnaires and Twitter," 2019 IEEE Students Conference on Engineering and Systems (SCES), Allahabad, India, 2019, pp. 1-6, doi: 10.1109/SCES46477.2019.8977211.
- [7] B. Yalamanchili, N. S. Kota, M. S. Abbaraju, V. S. S. Nadella and S. V. Alluri, "Real-time Acoustic based Depression Detection using Machine Learning Techniques," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-6, doi: 10.1109/icETITE47903.2020.394.
- [8] Yates, A., Cohan, A., and Goharian, N.: Depression and self-harm risk assessment in online forums. arXiv preprint arXiv:1709.01848 (2017).
- [9] Seabrook, E.M., Kern, M.L., Fulcher, B.D., and Rickard, N.S.: Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. Journal of Medical Internet Research 20 (5), e168 (2018).

- [10] O'Dea B, et al. Detecting suicidality on Twitter. Internet Interv. 2015;2(2):183–188. doi: 10.1016/j.invent.2015.03.005.
- [11] S. Grover, A. Dutt, and A. Avasthi, “An overview of indian research in depression,” Indian journal of psychiatry, vol. 52, no. Suppl1, p. S178, 2010.
- [12] M. Reddy, “Depression: the disorder and the burden,” Indian journal of psychological medicine, vol. 32, no. 1, p. 1, 2010.
- [13] “Depression.” [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/depression>
- [14] M. M. Aldarwish and H. F. Ahmad, “Predicting depression levels using social media posts,” in 2017 IEEE 13th international Symposium on Autonomous decentralized system (ISADS). IEEE, 2017, pp. 277–280.
- [15] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, and S. Zafeiriou, “End-to-end multimodal emotion recognition using deep neural networks,” IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 8, pp. 1301–1309, 2017.