JOURNAL OF DYNAMICS AND CONTROL

VOLUME 9 ISSUE 1: 48 - 61

INVESTIGATING EMOTION REGULATION DIFFICULTIES IN INDIVIDUALS WITH MENTAL HEALTH DISORDERS USING ADVANCED AUDIO ANALYSIS TECHNIQUES

Ashish Kumar Pandey, Tanisha Jain, Eram Fatma, Harshita Raj

Department of CSE, Dr. Rammanohar Lohia Avadh University, Ayodhya, U.P., India,

 DOI: https://doi.org/10.71058/jodac.v9i1006

INVESTIGATING EMOTION REGULATION DIFFICULTIES IN INDIVIDUALS WITH MENTAL HEALTH DISORDERS USING ADVANCED AUDIO ANALYSIS TECHNIQUES

Ashish Kumar Pandey[1*], Tanisha Jain[2], Eram Fatma[3], Harshita Raj[4]

[1,2,3,4] *Department of CSE, Dr. Rammanohar Lohia Avadh University, Ayodhya, U.P., India,*

[1] ashishkumarpandey@rmlau.ac.in, [2] jaintanisha706@gmail.com , [3]eramfatma1309@gmail.com, [4]harshitaraj499@gmail.com

Abstract: *Emotion regulation is essential for psychological well-being, particularly in individuals with Major Depressive Disorder (MDD) and Post-Traumatic Stress s Disorder (PTSD), who often struggle with managing their emotions. Traditional methods like self-reports and behavioral observations have limitations in fully capturing the complexities of emotion regulation. This research work addresses these gaps by employing advanced audio analysis techniques to explore the difficulties in emotion regulation experienced by individuals with mental health disorders. The proposed model uses advanced audio analysis techniques on audio data from participant-interviewer interactions and self-reported questionnaires. Logistic regression classifiers are employed to predict emotion regulation difficulties and shows strong correlations between DERS subscales and severity of disorders. The model's robustness is validated through Receiver Operating Characteristic (ROC) curve analysis and by calculating Area Under the Curve (AUC) values for each subscale of the Difficulties in Emotion Regulation Scale (DERS).*

Keywords: *Major Depressive Disorder, Post-Traumatic Stress Disorder, Audio analysis, Machine Learning, Difficulties in Emotion Regulation Scale, Non-intrusive assessment*

1. Introduction

The integration of advanced technologies in emotion regulation research signifies a significant step forward, offering insights into the dynamics of emotional experiences that traditional methods may overlook. One notable approach integrates audio analysis, enabling researchers to capture a wide array of information such as verbal expressions, non-verbal cues, and physiological responses. This multimodal approach contributes to significant understanding of emotion regulation within the context of mental health disorders, potentially offering a non-intrusive and cost-effective diagnostic tool that could transform mental health assessment and intervention strategies.

By analyzing both verbal and self-reported data, researchers can develop more targeted therapeutic approaches, leading to improved mental health outcomes. Previous studies have established associations between subscales of DERS questionnaire [1], degrees of Major Depressive Disorder [1], PTSD [2]. Emotions triggered in response to scenarios vary in intensity and polarity based on an individual's capacity for emotional regulation. Research indicates that the results may vary due to unbalanced ratio of male and female [3], [4], highlighting the importance of addressing this bias for fair and accurate diagnosis across different demographics.

Emotion regulation, as stated by Ross Thompson, mentions the processes, both internal as well as external, that monitor, evaluate, and modify emotional responses to achieve individual's objectives [5]. Various psychological issues are associated with deficits in ERD, and several studies have developed tools for measuring ERD [6], focusing notably on the DERS [7]. The DERS assesses an individual's level of ERD across six dimensions: (i) clarity about emotions, (ii) acceptance of one's emotions, (iii) ability to act effectively despite emotional distress, (iv) ability to control impulses, (v) awareness of emotional states, and (vi) use of effective emotion regulation strategies. Deficiencies in these dimensions indicate the extent to which an individual struggle with emotional control.

This study explores the use of audio data to examine emotion regulation difficulties in individuals with mental health disorders. Our methodology involves recording audio of participants during interactions with interviewers, alongside their responses to standardized questionnaires. These

 DOI: https://doi.org/10.71058/jodac.v9i1006

recordings provide comprehensive data, capturing both verbal and non-verbal aspects of emotion regulation.

During a Zoom call, participants were asked questions derived from the DERS. Responses were recorded using clip-on microphones and webcams to identify symptoms of MDD and PTSD. Alongside audio data, self-reported information through online forms including the DERS, the Patient Health Questionnaire-8 (PHQ-8), and the Posttraumatic Stress Disorder Checklist-Civilian Version (PCL-C) are collected. Self-reported data served as a reference for validating audio findings. DERS subscale scores were used to estimate ERD, while PHQ-8 and PCL-C total scores assessed depression and PTSD severity respectively.

Gratz and Roemer (2004) provided a comprehensive framework for understanding ERD, emphasizing their prevalence in various mental health conditions such as MDD and PTSD. Aldao et al. (2010) conducted a meta-analytic review highlighting the association between emotion regulation strategies and psychopathology, indicating that maladaptive strategies are commonly observed in individuals with mental health disorders. Gross (2015) reviewed the extended process model of emotion regulation, detailing the sequential nature of emotion regulation strategies and their differential impacts on mental health. John and Gross (2004) discussed the limitations of self-report measures in capturing emotion regulation processes, citing issues such as social desirability bias and lack of introspective accuracy.

Cowie et al. (2001) pioneered the use of audio analysis in emotion recognition, demonstrating the potential of vocal features in detecting emotional states. Scherer (2003) provided a detailed review of vocal indicators of emotion, highlighting the significance of prosody and speech patterns in understanding emotional experiences. Busso et al. (2004) developed a multimodal database for emotion analysis, combining audio and video data to enhance the accuracy of emotion detection systems. Zeng et al. (2009) reviewed the advances in multimodal emotion recognition, emphasizing the integration of audio, video, and physiological data for more robust assessments.

De Silva et al. (2003) demonstrated the efficacy of multimodal approaches in emotion recognition, showing significant improvements over unimodal systems. Koolagudi and Rao (2012) explored the role of speech emotion recognition in mental health applications, discussing the implications for automated diagnosis and intervention. This integrated approach holds promise for enhancing our understanding and treatment of mental health disorders, leveraging advanced technology to provide more precise assessments and interventions.

The paper is structured as follows: the introduction section outlines the importance of emotion regulation and the challenges faced by individuals with mental health disorders. It highlights the limitations of traditional methods and introduces the innovative approach of using audio analysis for a more nuanced understanding of emotion regulation difficulties. Data collection and its justification have been discussed in data collection section. The methodology section details the data collection process, including the use of audio recordings and self-reported questionnaires. The results section presents findings, emphasizing the ROC curve and AUC values for each DERS subscale in predicting PTSD and MDD. Finally, the Conclusion summarizes the work contributions and potential implications for future research and clinical practice.

2. Data collection

The collection of data includes recording of the participants' audio through online interaction with the interviewer as well as their self-reported responses. The audio data is collected through zoom calls. The participants respond to the set of questions (self-reported), which are taken from the DERS questionnaire based on one to one interaction.

2.1 Data collection protocol

Firstly, all participants were informed about the data collection process. Then the participants were requested to be seated comfortably while attending the zoom call on their computer or mobile phone. Initially, they were asked several questions to make them feel comfortable. Subsequently, the questions present in the Difficulties in Emotion Regulation Scale questionnaire [25] were asked in successive order. These zoom calls are recorded and retained in M4A format. As soon as the recordings are completed, the participants' responses to an online form consisting of DERS, PHQ-

 DOI: https://doi.org/10.71058/jodac.v9i1006

8[26] and PCL-C questionnaires collectively [27] are recorded which are further used as self-reported data (SRD).

2.2 Participants

Regarding this work the data of 30 under-graduate students comprising of equal number of male-female participants aged between 20 to 25 years is collected.

2.3 Audio data

The participants' audio responses during zoom call are recorded in M4A audio format. Their face is the primary focus of the webcam. In this work, free and open-source zoom application is utilized for recording audio data. This work involves 2.57 hours of audio recordings.

2.4 Self-reported data

The authenticity for the participant's audio data is ensured by their answers to an online form, which is created by merging the DERS, PHQ-8, and PCL-C questionnaires. Difficulties in Emotion Regulation Scale subscale scores are utilized as an authenticating factor for assessment of emotion regulation difficulty. These scores are calculated by aggregating the weights corresponding to the selected choices for a particular subscale. The 5-point Likert scale is used for rating the choices associated with each question [28].

Table 1 shows the number of items belonging to each subscale, score range as well as their $\mu \pm \sigma$ values. The authenticity for Major depressive disorder detection is recorded by totalizing the participants' PHQ-8 scores. The total score is calculated by aggregating the weights corresponding to the selected choices which denote the degree of depression of a participant. The 4-point Likert scale is used for rating the choices associated with each question [28]. The whole score lies between 0 to 24. A threshold of score 10 is used to detect whether a person suffers from Major depressive disorder or not. The authenticity for Post-traumatic stress disorder detection is recorded by totalizing the participant's PCL-C scores. The total score is calculated by aggregating the weights corresponding to the selected choices which denote the degree of post-traumatic stress disorder detection by which a participant is suffering. The 5-point Likert scale is used for rating the choices associated with each question [28]. The whole score lies between 17 to 85. A threshold of score 30 is used to detect whether a person suffers from PTSD or not.

Therefore, in this dataset 13 participants are found to be suffering from major depressive disorder and 22 participants are suffering from post-traumatic stress disorder. Thus, the proportion of people with MDD and without MDD is 13:17. Also, the proportion of people with post-traumatic stress disorder and without it is 22:8. Therefore, this dataset is disproportionate for Post-traumatic stress disorder which has been addressed in this work.

Table 1. Statistical Summary of DERS Subscale Scores (Mean μ and SD σ)

Subscale	Items	Score Boundaries	μ ± σ(Mean ± SD)
Clarity	5	5-25	10.23 ± 4.59
Non-acceptance	6	6-30	13.20 ± 6.49
Goals	5	5-25	14.6 ± 4.85
Impulse	6	6-30	13.9 ± 4.60
Awareness	6	6-30	12.56 ± 4.82
Strategies	8	8-40	6.73 ± 6.07

 DOI: https://doi.org/10.71058/jodac.v9i1006

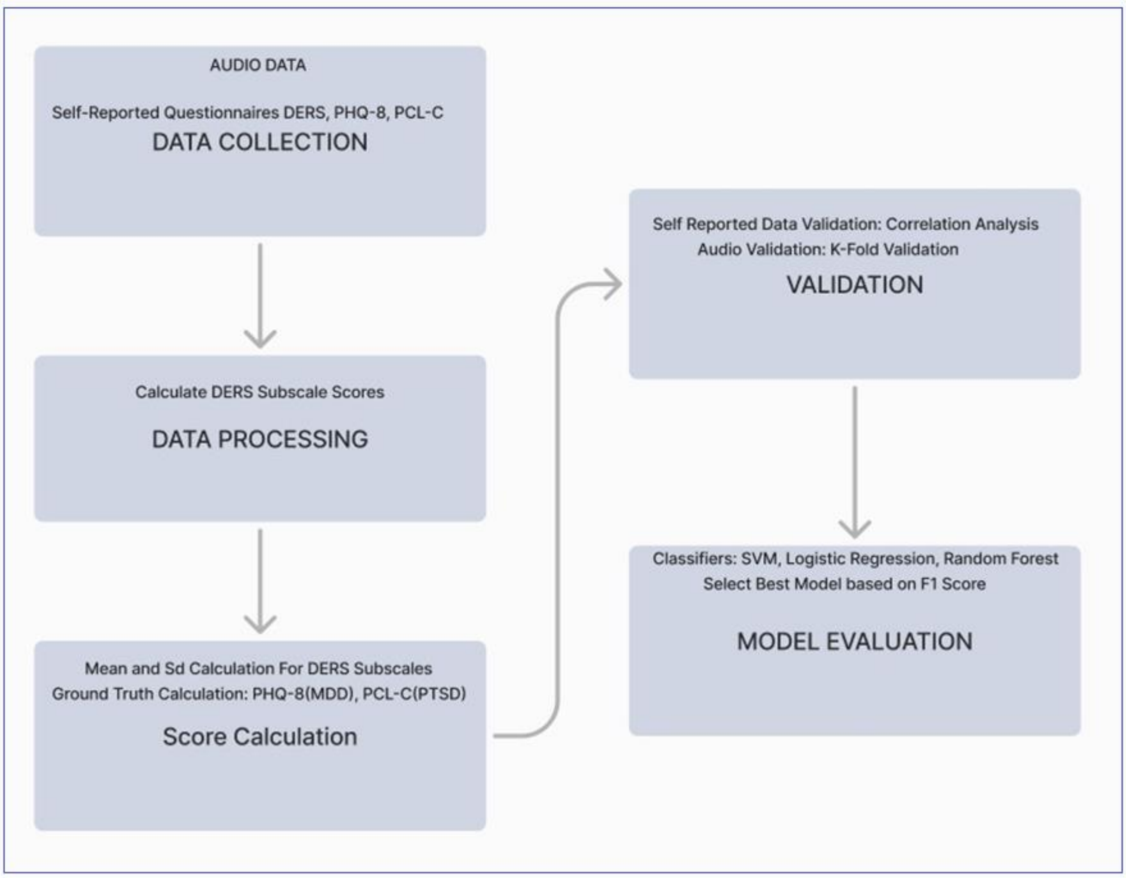
2.5 Rationale for data collection protocol

Interviewing the participants is part of the data gathering procedure for the automated identification of mental diseases. Usually, during the interview, different questions are asked from the participants that are not included in the questionnaire used for identifying a mental disorder [29], [30], [31], [32].The interview between the interviewer and respective participant, stored in audio format is marked by the combined results of PHQ-8 and PCL-C for that participant. The purpose of this work is to identify Post Traumatic Stress Disorder and Major Depressive Disorder using ERD as an intermediary form of audio data. The Emotion Regulation Difficulties level of each participant is measured by the DERS subscales as featured in Section I.

Therefore, in order to determine a DERS subscale's score based on audio data, the interviewer must provide questions about that subscale. For ease of implementation, a question sheet linked to the respective subscale of DERS was not created. Rather, our data collecting technique involves asking individuals items straight out of the Difficulties in Emotion Regulation Scale question sheet. A comparable method has been applied in [33] to estimate quality of life. In that work, the conversation between participants and a computer agent are recorded in audio form. These conversations involve questions from the short form-36 version 2 (SF-36v2) questionnaires [34]. The authenticity of a participant's audio recording is derived from their SRD on the SF-36v2.

3. Methodology

The primary aim is to investigate the necessity of Emotion Regulation Difficulties, inferred from audio data, as an intermediary form for evaluating the degree of Major Depressive Disorder and Post-Traumatic Stress Disorder. The suggested methodology (shown in figure. 1) involves estimating the DERS subscale score and leveraging these scores to gauge the severity of MDD/PTSD.



DOI: https://doi.org/10.71058/jodac.v9i1006

Figure 1: Proposed methodology

In this order, spreadsheets are organized with the computed subscale scores for each participant, derived from both self-reported questionnaires and audio data. The self-reported questionnaire includes DERS, PHQ-8, and PCL-C questionnaires. Further, mean and standard deviation are calculated for each DERS subscale as indicated in Table 1, including data from all 30 participants.

To detect MDD, the ground truth is the combined score from the Patient Health Questionnaire-8, representing the participant's depression severity level. It is computed by aggregating the weights of the selected choices. For detection of Post-Traumatic Stress Disorder, the authenticity is the combined score from the Post-traumatic Stress Disorder Checklist-Civilian Version questionnaire, indicating the PTSD severity level, calculated similarly by summing the associated weights of the responses.

The collected data is validated separately for both self-reported questionnaires and audio sources. Data validation on online form is carried out through correlation analysis, calculating Pearson correlation coefficients among DERS subscales, MDD severity levels, as well as acuteness of PTSD extent. Audio data validation is conducted using the k-fold cross-validation method to ensure robust evaluation of machine learning models. The three classifiers are utilized in this work: a support vector machine, logistic regression classifier and random forest.

4. Results

The investigatory assessment of the preferred approach for evaluating Emotion Regulation Difficulties utilizing the SRD and auditory modality is presented in this section. Further, using the calculated ERD, the evaluation of two selected mental illnesses (PTSD and MDD) are performed. The tests are conducted to verify the generated dataset. These tests are implemented using Python 3.7 and IDLE (v3.7.0) on an Intel Core i3 7th Gen CPU.

The following sub-sections discusses and analyzes the authentication of SRD and audio data. A correlation analysis is performed for self-reported data. Audio data validation involves detailing the use of classifiers and the evaluation of performance metrics for detecting Major Depressive Disorder and Post-Traumatic Stress Disorder.

4.1 Analyzing the validation of self-reported data (SRD)

Table 2 displays the Pearson correlation coefficients derived between the Difficulties Emotion Regulation subscale scores and the seriousness of Major Depressive Disorder and Post-traumatic stress disorder. It is worth noting how the graveness of these both disorders has a logarithmically remarkable connection with all DERS subscales excluding the Strategies subscale. The connection between MDD and PTSD gravity levels is significant (0.83, $p < 0.01$). Thus, strong relationships across all DERS subscales excluding the strategies subscale are found. Our findings show comparable trends, which corroborate the SRD in our sample.

Table 2: Correlations among DERS subscale scores, severeness of MDD and PTSD derived from SRD

SUBSCALE	1	2	3	4	5	6	7	8
CLARITY	1	0.41^{+}	0.32^{+}	0.34^{+}	0.74*	0.4^{+}	-0.11	0.04
NON-ACCEPT		1	0.21	0.5*	0.43^{+}	0.58*	-0.09	0.08
GOALS			1	0.57*	0.02	0.53*	-0.08	-0.04
IMPULSE				1	0.22	0.75*	-0.01	0.06
AWARENESS					1	0.32^{+}	-0.04	0.14

DOI: https://doi.org/10.71058/jodac.v9i1006

STRATEGIES	1	0.08	0.18
MDD		1	0.83*
PTSD			1

*p <0.01; +p <0.05

4.2 Analysing the validation of audio data

In this investigation, the classifiers for determining Major Depressive Disorder and Post Traumatic Stress Disorder are instructed as well as examined on the basis of auditory feedback of each subscale in our dataset individually. For both Major Depressive Disorder and Post-Traumatic Stress Disorder detection, the training set consists of 18 participants and the testing set consists of 12 participants, respectively. In our experiment 13 out of 30 participants suffer from MDD and 22 out of 30 participants suffer from PTSD. Thus, there is an asymmetry in the dataset. To deal with the asymmetry in data, a threshold shifting approach is employed in the composition [35] [36]. Macro-average F1-score [37] is used for choosing optimal decision boundary and k-fold cross validation method is used to identify the optimal threshold boundary. It is applied on the training partition. Table 3 shows the calculated macro F1 score for unbalanced data using SVM with RBF kernel.

TABLE 3: Macro-F1 score for unbalanced audio data

SUBSCALE	MDD	PTSD
Awareness	0.43	0.40
Goals	0.48	0.87
Non-accept	0.73	0.69
Strategies	0.74	0.62
Clarity	0.36	0.45
Impulse	0.65	0.69

For balanced data, experiments are performed using different Machine Learning models and then compared. The models used for classification are Support Vector Machine (SVM), Logistic Regression and Random Forest classifier. Precision, recall, f1 score and accuracy are being calculated and the model which produces the most optimal value of all 4 parameters is chosen. Among the 3 models, the Logistic Regression Classifier produces the most optimal result. For our experiment, balanced data is used instead of unbalanced data as it has greater F1 score value for each DERS subscale.

 DOI: https://doi.org/10.71058/jodac.v9i1006

TABLE 4: Evaluation of performance metrics for each subscale using SVM

SUBSCALE	MDD				PTSD			
	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score	Accuracy
Awareness	0.56	0.54	0.49	0.5	0.61	0.65	0.41	0.42
Goals	0.55	0.53	0.50	0.58	0.83	0.95	0.87	0.92
Non-accept	0.75	0.76	0.75	0.75	0.70	0.85	0.70	0.75
Strategies	0.55	0.53	0.50	0.58	0.67	0.80	0.62	0.67
Clarity	0.73	0.57	0.41	0.50	0.56	0.60	0.56	0.67
Impulse	0.89	0.80	0.81	0.83	0.75	0.90	0.78	0.83

TABLE 5: Evaluation of performance metrics for each subscale using Logistic Regression

SUBSCALE	MDD				PTSD			
	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score	Accuracy
Awareness	0.56	0.54	0.49	0.50	0.61	0.65	0.41	0.42
Goals	0.89	0.80	0.81	0.83	0.83	0.95	0.87	0.92
Non-accept	0.75	0.76	0.75	0.75	0.70	0.85	0.70	0.75
Strategies	0.75	0.76	0.75	0.75	0.67	0.80	0.62	0.67
Clarity	0.78	0.71	0.66	0.67	0.70	0.70	0.70	0.83
Impulse	0.89	0.80	0.81	0.83	0.75	0.90	0.78	0.83

TABLE 6: Evaluation of performance metrics for each subscale using Random Forest

SUBSCALE	MDD				PTSD			
	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score	Accuracy
Awareness	0.73	0.57	0.44	0.50	0.61	0.65	0.41	0.42
Goals	0.66	0.66	0.66	0.67	0.83	0.95	0.87	0.92
Non-accept	0.75	0.73	0.73	0.75	0.70	0.70	0.70	0.83
Strategies	0.81	0.79	0.75	0.75	0.70	0.85	0.70	0.75
Clarity	0.75	0.73	0.75	0.75	0.95	0.75	0.81	0.92
Impulse	0.89	0.80	0.81	0.83	0.75	0.90	0.78	0.83

For audio modality, the key observations that are indistinguishable to those derived from online form are given in Table 7.

 DOI: https://doi.org/10.71058/jodac.v9i1006

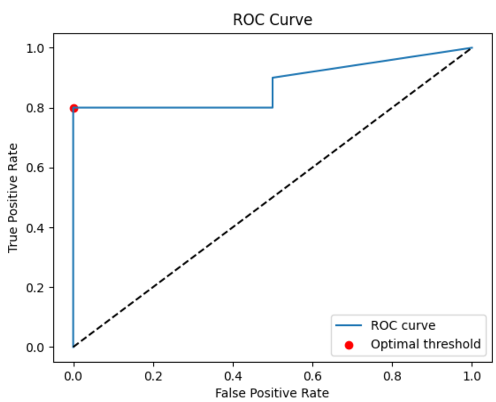
i. The relationship between the graveness of major depressive and post-traumatic stress disorders is computationally remarkable.
ii. The seriousness of major depressive and post-traumatic stress disorders is strongly linked with the majority of DERS subscales.
iii. Except for a few situations, links among all Difficulties in Emotion Regulation subscales, eliminating the Strategies subscale, are statistically important. The biggest dissimilarity with respect to the data collected through online form is over Strategies subscale, which is substantially linked with the acuteness of Post-Traumatic Stress Disorder and Major Depressive Disorder in auditory modality.

TABLE 7: Relational analysis of MDD severity level, DERS subscale scores, and PTSD severity levels based on audio data

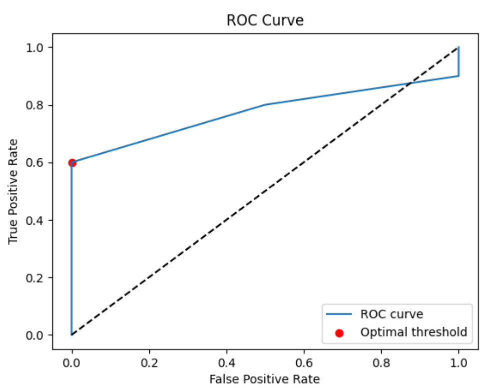
SUBSCALE	1	2	3	4	5	6	7	8
CLARITY	1	0.58*	0.24	0.54*	0.57*	0.73*	0.25	0.29
NON-ACCEPT		1	0.40^{+}	0.76*	0.51*	0.67*	0.58*	0.64*
GOALS			1	0.39^{+}	0	0.3	0.38^{+}	0.37^{+}
IMPULSE				1	0.46^{+}	0.65*	0.49^{+}	0.57*
AWARENESS					1	0.46^{+}	0.2	0.26
STRATEGIES						1	0.45^{+}	0.59*
MDD							1	0.83*
PTSD								**1**

*p <0.01; +p <0.05

Logistic Regression is used as a suitable classifier and for its validation, the **Receiver Operating Characteristic (**ROC) curves of MDD and PTSD against each DERS subscale are calculated.

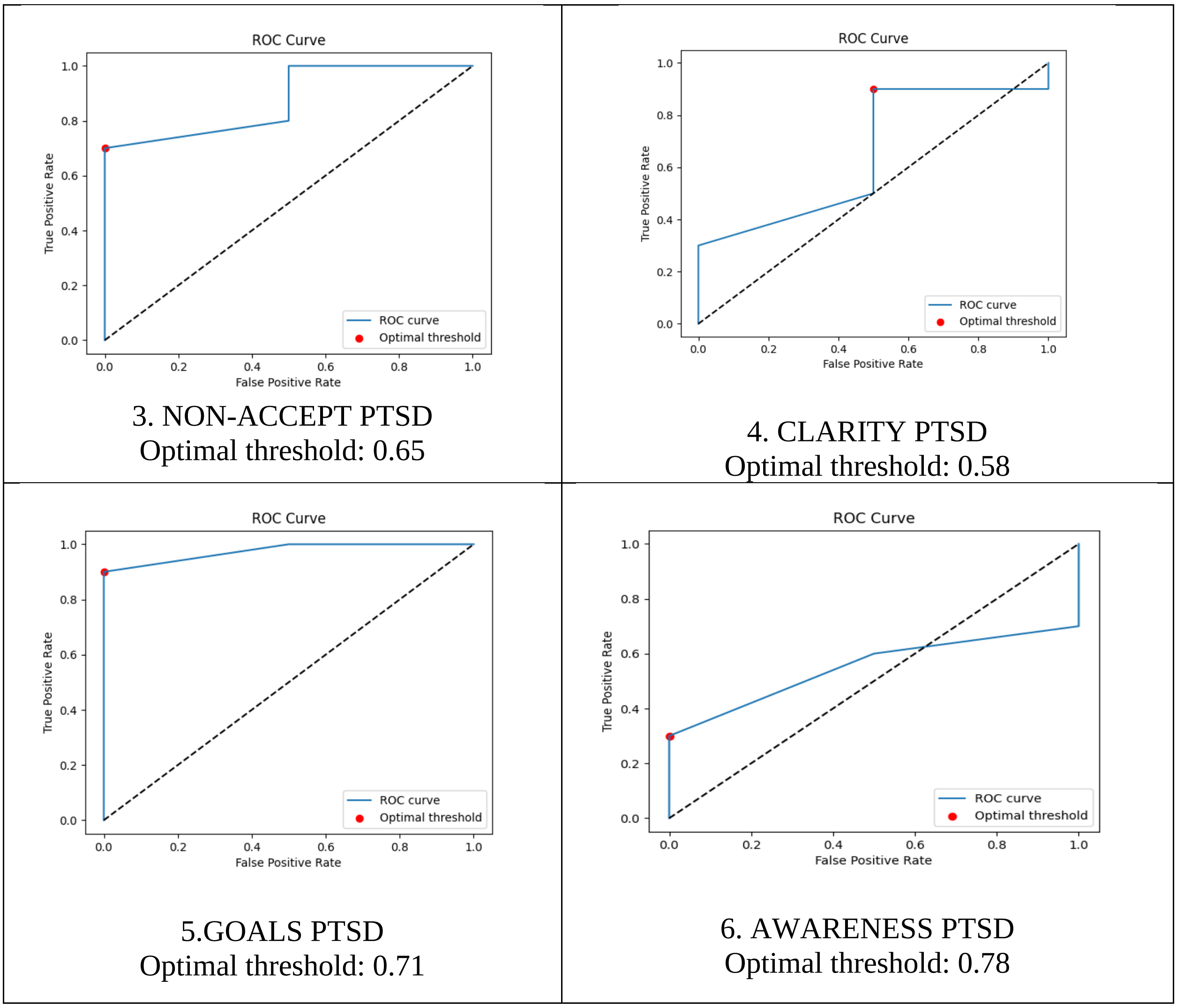


1. IMPULSE PTSD
Optimal threshold: 0.68



2. STRATEGIES PTSD
Optimal threshold: 0.81

DOI: https://doi.org/10.71058/jodac.v9i1006



3. NON-ACCEPT PTSD
Optimal threshold: 0.65

4. CLARITY PTSD
Optimal threshold: 0.58

5.GOALS PTSD
Optimal threshold: 0.71

6. AWARENESS PTSD
Optimal threshold: 0.78

FIGURE 5: ROC CURVES of each DERS subscale for PTSD

 DOI: https://doi.org/10.71058/jodac.v9i1006

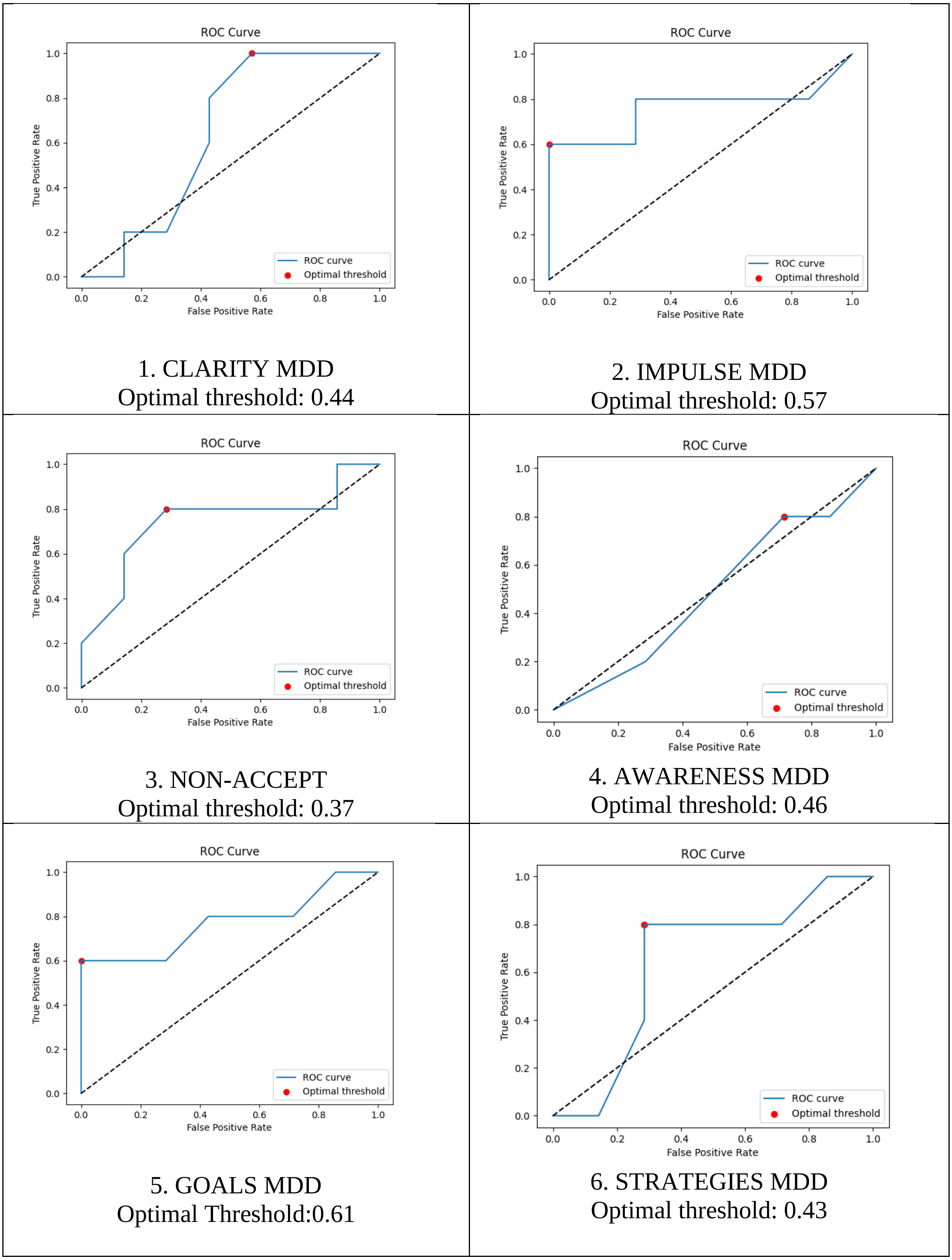


FIGURE 6: ROC CURVES of each DERS subscale for MDD

 DOI: https://doi.org/10.71058/jodac.v9i1006

Figure 5 and 6 represents ROC curve which is used to assess the performance of classification model. The x-axis of this curve represents FPR and y-axis represents TPR. The diagonal line from (0,0) to (1,1) represents a classifier that makes random guesses. A model whose ROC curve is close to this diagonal is no better than random guessing.

Each ROC curve helps in understanding how well each specific subscale of the DERS can be used as a diagnostic tool for MDD and PTSD. The closer the value of Area Under Curve (AUC) is to 1, the better the subscale is at predicting MDD and PTSD [38].

TABLE 8: AUC values of ROC curves of each DERS subscale for PTSD and MDD

CURVE	PTSD	MDD
1.	0.86	0.66
2.	0.76	0.75
3.	0.86	0.71
4.	0.65	0.46
5.	0.95	0.67
6.	0.55	0.65

This table consists of AUC values of all the curves in Figure 5 and 6 which conclude that Goals and Impulse subscale is most efficient at predicting PTSD and MDD respectively.

5. Discussion

This research work further demonstrates that advanced audio analysis techniques can effectively identify emotion regulation difficulties in individuals with MDD and PTSD. By utilizing logistic regression classifiers to analyze audio data from participant-interviewer interactions and self-reported questionnaires, the study found significant correlations between specific DERS subscales and the severity of these disorders. These findings further support the potential of audio-based assessments as diagnostic tools for mental health conditions.

The strengths of this work include its innovative use of audio analysis to provide a more objective and nuanced assessment of emotion regulation difficulties, addressing the limitations of traditional self-reports and behavioral observations. However, the study also has limitations, such as a relatively small sample size and potential for audio quality variations affecting the analysis.

In comparison to similar research, this work aligns with previous studies that highlight the importance of various approaches in understanding emotion regulation. However, this research further advances in the field, specifically focusing on emotion regulation difficulties in clinical populations, such as individuals with MDD and PTSD.

The significant correlations between DERS subscales and disorder severity underscore the utility of audio-based assessments in capturing the complexities of emotion regulation. These findings suggest that audio analysis can provide insights that are difficult to obtain through traditional methods. The implication is that integrating audio-based assessments into clinical practice could further enhance diagnostic accuracy and inform more tailored therapeutic interventions.

DOI: https://doi.org/10.71058/jodac.v9i1006

6. Conclusion

This work delves into the intricate role of emotion regulation within the realm of mental health, with a particular emphasis on its dysregulation across various psychiatric conditions such as anxiety disorders, mood disorders, and post-traumatic stress disorder (PTSD). Leveraging advanced techniques like audio analysis, the research endeavors to comprehensively capture a spectrum of verbal cues to discern patterns of emotion dysregulation. By employing audio recordings of participants engaging with the Difficulties in Emotion Regulation Scale (DERS) questionnaire alongside self-reported data, the work aims to gauge the severity of major depressive disorder (MDD) and PTSD. Through sophisticated correlation analysis and the application of machine learning classifiers, including Support Vector Machines (SVM), logistic regression, and random forest algorithms, the research seeks to validate the self-reported data while identifying significant patterns of emotion dysregulation that align with the severity of mental disorders. Logistic Regression is used as a suitable classifier and for its validation, the Area Under the Curve (AUC) values of the ROC curves were computed to quantify the diagnostic ability of each DERS subscale in predicting MDD and PTSD. The closer the AUC value is to 1, the better the subscale is at prediction. The results, summarized in Table 8, indicates that the Goals subscale is the most efficient at predicting PTSD, with the highest AUC value.

In contrast, the Impulse subscale demonstrates the greatest efficiency in predicting MDD. Moreover, the work underscores the paramount importance of ethical considerations in managing sensitive data throughout the research process. Ultimately, the research strives to contribute to the development of targeted clinical interventions by demonstrating how the integration of audio analysis techniques can offer invaluable insights into the challenges surrounding emotion regulation in the context of mental health. These findings suggest that specific subscales of the DERS can serve as effective diagnostic tools for MDD and PTSD when using logistic regression classifiers. The identification of these subscales provides a targeted approach for diagnosing these mental disorders, enhancing the overall accuracy and reliability of predictions based on audio data.

Future research should aim to address the limitations of this study by including larger and more diverse samples, improving audio data collection techniques, and integrating additional modalities such as physiological data. This would further validate the findings and enhance the robustness of audio-based diagnostic tools. Clinicians and researchers should consider adopting these innovative techniques to improve the assessment and treatment of emotion regulation difficulties in mental health disorders.

Acknowledgement

We would like to thank Dr. R.M.L. Avadh University, Ayodhya for encouraging us in doing research pertaining to real time scenario which created an interest in this topic.

Author Contributions

All authors have equally contributed.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the study of this article.

Funding

Not applicable.

References

[1] L. S. Hallion, S. A. Steinman, D. F. Tolin, and G. J. Diefenbach, "Psychometric properties of the difficulties in emotion regulation scale (DERS) and its short forms in adults with emotional disorders," *Front. Psychol.*, vol. 9, **(2018)**, pp. 1-12.

[2] N. H. Weiss, M. Tull, L. T. Davis, E. E. Dehon, J. J. Fulton, and K. L. Gratz, "Examining the association between emotion regulation difficulties and probable posttraumatic stress disorder

 DOI: https://doi.org/10.71058/jodac.v9i1006

within a sample of African Americans," *Cogn. Behav. Ther.*, vol. 41, no. 1, **(2012)**, pp. 5-14.

[3] K. Yang, K. Qinami, L. Fei-Fei, J. Deng, and O. Russakovsky, "Towards fairer datasets: Filtering and balancing the distribution of the people subtree in the ImageNet hierarchy," in *Proc. Conf. Fairness, Accountability, Transparency*, **(2020)**, pp. 547-558.

[4] A. Bailey and M. D. Plumbley, "Gender bias in depression detection using audio features," *arXiv preprint*, arXiv:2010.15120, **(2020)**.

[5] R. A. Thompson, "Emotion regulation: A theme in search of definition," *Monogr. Soc. Res. Child Dev.*, vol. 59, no. 2-3, **(1994)**, pp. 25-52.

[6] E. L. Ferrell, T. S. Watford, and A. Braden, "Emotion regulation difficulties and impaired working memory interact to predict boredom emotional eating," *Appetite*, vol. 144, **(2020)**, Art. no. 104450.

[7] K. L. Gratz and L. Roemer, "Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale," *J. Psychopathol. Behav. Assess.*, vol. 26, no. 1, **(2004)**, pp. 41-54.

[8] A. Aldao, S. Nolen-Hoeksema, and S. Schweizer, "Emotion-regulation strategies across psychopathology: A meta-analytic review," *Clin. Psychol. Rev.*, vol. 30, no. 2, **(2010)**, pp. 217-237.

[9] J. J. Gross, "Emotion regulation: Current status and future prospects," *Psychol. Inquiry*, vol. 26, no. 1, **(2015)**, pp. 1-26.

[10] O. P. John and J. J. Gross, "Healthy and unhealthy emotion regulation: Personality processes, individual differences, and life span development," *J. Pers.*, vol. 72, no. 6, **(2004)**, pp. 1301-1334.

[11] S. L. Koole, "The psychology of emotion regulation: An integrative review," *Cogn. Emot.*, vol. 23, no. 1, **(2009),** pp. 4-41.

[12] G. A. Bonanno and C. L. Burton, "Regulatory flexibility: An individual differences perspective on coping and emotion regulation," *Perspect. Psychol. Sci.*, vol. 8, no. 6, **(2013)**, pp. 591-612.

[13] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Process. Mag.*, vol. 18, no. 1, **(2001)**, pp. 32-80.

[14] K. R. Scherer, "Vocal communication of emotion: A review of research paradigms," *Speech Commun.*, vol. 40, no. 1-2, **(2003)**, pp. 227-256.

[15] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, et al., "Analysis of emotion recognition using facial expressions, speech and multimodal information," in *Proc. 6th Int. Conf. Multimodal Interfaces*, ACM, **(2004)**, pp. 205-211.

[16] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, **(2009),** pp. 39-58.

[17] L. C. De Silva, B. S. Mendis, and S. Sumathipala, "Prosody-based automatic emotion recognition in speech signals: An overview," *J. Natl. Sci. Found. Sri Lanka*, vol. 31, no. 3, **(2003)**, pp. 139-152.

[18] S. G. Koolagudi and K. S. Rao, "Emotion recognition from speech: A review," *Int. J. Speech Technol.*, vol. 15, no. 2, **(2012)**, pp. 99-117.

[19] F. Nasoz, K. Alvarez, C. L. Lisetti, N. Finkelstein, and A. Shapiro, "Developing empathetic agents: Personalized emotion recognition and empathetic response," *Appl. Artif. Intell.*, vol. 17, no. 3-4, **(2003)**, pp. 245-266.

[20] C. H. Lu and C. H. Wu, "Emotion detection for affective computing based on physiological signals," *J. Sci. Eng. Res.*, vol. 3, no. 1, **(2012)**, pp. 51-56.

[21] N. Cummins, J. Epps, M. Breakspear, and R. Goecke, "An investigation of depressed speech detection: Features and normalization," *IEEE Trans. Affect. Comput.*, vol. 6, no. 2, **(2015)**, pp. 148-161.

[22] B. Schuller, S. Steidl, A. Batliner, A. Vinciarelli, K. Scherer, F. Ringeval, et al., "The INTERSPEECH 2018 Computational Paralinguistics Challenge: Atypical and self-assessed affect, crying and heart beats," in *Proc. Interspeech 2018*, **(2018)**, pp. 171-175.

[23] H. Gunes and B. Schuller, "Categorical and dimensional affect analysis in continuous input: Current trends and future directions," *Image Vis. Comput.*, vol. 31, no. 2, **(2013),** pp. 120-136.

DOI: https://doi.org/10.71058/jodac.v9i1006

[24] E. Kaufman, M. Xia, G. Fosco, M. Yaptangco, C. Skidmore, S. Crowell, "DERS-SF scoring and measure," **(2015)**.

[25] K. Kroenke, T. W. Strine, R. L. Spitzer, J. B. W. Williams, J. T. Berry, A. H. Mokdad, "The PHQ-8 as a measure of current depression in the general population," *J. Affect. Disord.*, vol. 114, no. 1-3, **(2009),** pp. 163-173.

[26] M. A. Andrykowski, M. J. Cordova, J. L. Studts, and T. W. Miller, "Posttraumatic stress disorder after treatment for breast cancer: Prevalence of diagnosis and use of the PTSD checklist—Civilian version (PCL—C) as a screening instrument," *J. Consult. Clin. Psychol.*, vol. 66, no. 3, **(1998),** pp. 586-590.

[27] R. Likert, "A technique for the measurement of attitudes," *Arch. Psychol.*, vol. 22, no. 140, **(2014),** pp. 5-55.

[28] S. Scherer, G. Stratou, M. Mahmoud, J. Boberg, J. Gratch, A. Rizzo, L. P. Morency, "Automatic behavior descriptors for psychological disorder analysis," in *Proc. 10th IEEE Int. Conf. Workshops Autom. Face Gesture Recognit. (FG)*, **(2013)**, pp. 1-8.

[29] K. Schulte Braucks, V. Yadav, A. Y. Shalev, G. A. Bonanno, I. R. Galatzer-Levy, "Deep learning-based classification of posttraumatic stress disorder and depression following trauma utilizing visual and auditory markers of arousal and mood," *Psychol. Med.*, vol. 52, no. 5, **(2022),** pp. 957-967.

[30] E. Victor, Z. M. Aghajan, A. R. Sewart, R. Christian, "Detecting depression using a framework combining deep multimodal neural networks with a purpose-built automated evaluation," *Psychol. Assess.*, vol. 31, no. 8, pp. 1019, 2019.

[31] J. Gratch, R. Arnstein, G. Lucas, G. Stratou, S. Scherer, A. Nazarian, et al., "The distress analysis interview corpus of human and computer interviews," in *Proc. Int. Conf. Lang. Resour. Eval.*, **(2014)**, pp. 3123-3128.

[32] S. Nakagawa, D. Enomoto, S. Yonekura, H. Kanazawa, Y. Kuniyoshi, "New telecare approach based on 3D convolutional neural network for estimating quality of life," *Neurocomputing*, vol. 397, **(2020),** pp. 464-476.

[33] G. Hawthorne, R. Osborne, A. Taylor, J. Sansoni, "The SF36 version 2: Critical analyses of population weights, scoring algorithms and population norms," *Qual. Life Res.*, vol. 16, no. 4, **(2007),** pp. 661-673.

[34] F. J. Provost and T. Fawcett, "Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions," in *Proc. 3rd Int. Conf. Knowl. Discov. Data Min.*, **(1997)**, pp. 43-48.

[35] R. K. Gupta and R. Sinha, "An investigation on the audio-video data-based estimation of emotion regulation difficulties and their association with mental disorders," *IEEE Access*, vol. 11, pp. 74324-74336, **(2023)**, doi: 10.1109/ACCESS.2023.3297091.

[36] Y. Yang and X. Liu, "A re-examination of text categorization methods," in *Proc. 22nd Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, **(1999)**, pp. 42-49