Environment and Social Psychology (2024) Volume 9 Issue 4
doi: 10.54517/esp.v9i4.2362

ORIGINAL RESEARCH ARTICLE

Relationship between mental health and job competence of college teachers: Application of sentiment analysis algorithm

Xiaobin Wu*, **Lizheng Zhuo**

International College, Krirk University, Bangkok 10220, Thailand

*** Corresponding author:** Xiaobin Wu, wuxiaobinhope@163.com

ABSTRACT

In systematically collecting and investigating data related to the teaching field, the association between teachers' Mental Health (MH) and Job Competence (JC) is an important field that has not been systematically studied. Understanding this association is vital due to its direct impact on the Quality of Education (QoE) and its standard educational system. Due to problems in evaluating all qualitative data, such as Open-Ended Survey Feedback (OESF), traditional models frequently find it challenging to epitomize intricate relations accurately. This article reports on these tasks by introducing a Mixed-Method Approach (MMA). The research study was directed through online learning across numerous higher education institutions in three provinces of China, and a combination analysis of quantitative study data using qualitative Sentiment Analysis (SA) was recommended. The new thing about the method is that it suggests an algorithm based on Latent Dirichlet Allocation (LDA) to SA that lets all of the qualitative data from the OESF questions be studied. This algorithm suggests a more philosophical knowledge of teachers' MH and its association with their specialized skills overall OESF. The study's results represent a significant insight into the dynamics between MH and JC related to College Teachers (CT). It highlights how every feature impacts others and the prerequisites for educational strategies supporting teacher well-being. By overcoming boundaries in existing models, the proposed work contributes to a broader and better knowledge of teacher well-being, its impact on educational quality, and the potential for SA in educational research. We compared how well three classifiers—Naïve Bayes (NB), Support Vector Machine (SVM), and Linear Regression (LR)—performed on six topics that were chosen for this SA research. The performance analysis for evaluation is accuracy, precision, recall, and F1-score. The reliability of the OESF measures was confirmed with Cronbach's alpha values signifying high internal consistency: 0.85 for JC, 0.88 for MH status, and 0.82 for innovative teaching ability.

Keywords: teaching and learning environment; latent Dirichlet allocation; sentiment analysis; mental health; education; machine learning

1. Introduction

The dynamic relationship between Mental Health (MH) and Job Competence (JC) is increasingly recognized as critical for numerous professions, particularly education. MH significantly influences the individual's cognitive abilities, emotional resilience, and overall performance. In education, where JC encompasses past subject knowledge, including personal skills, intelligent teaching methods, and Emotional Intelligence (EI), the impact on MH becomes even more noticeable[1]. This involved relationship highlights

ARTICLE INFO
Received: 3 January 2024 | Accepted: 5 February 2024 | Available online: 28 February 2024

CITATION
Wu X, Zhuo L. Relationship between mental health and job competence of college teachers: Application of sentiment analysis algorithm. *Environment and Social Psychology* 2024; 9(4): 2362. doi: 10.54517/esp.v9i4.2362

COPYRIGHT
Copyright © 2024 by author(s). *Environment and Social Psychology* is published by Asia Pacific Academy of Science Pte. Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), permitting distribution and reproduction in any medium, provided the original work is cited.

the need for in-depth knowledge of how MH and skilled value affect each other, particularly with College Teachers (CT). In education, recognizing and understanding the relationship between these two features of CTs is critical so that it can lead to better support systems, knowledgeable policy-making, and targeted involvement to improve teacher MH and educational results. This supports the CTs and the comprehensive academic community, including students and higher education institutions.

Qualitative and quantitative research practices are often active in discovering multipart MH and JC relationships. Quantitative methods provide quantifiable, objective data to help create patterns and relationships[2]. Conversely, qualitative methods offer a more philosophical and context-rich insight into distinct skills and observations. This hybrid method enables a more general understanding of the statistical strength of quantitative analysis and the in-depth details of qualitative research. In qualitative research, sentiment analysis (SA) is employed mainly for understanding OESF[3]. Based on natural language processing (NLP), this system is proficient at extracting independent sentiments from textual data. Numerous approaches, such as Latent Dirichlet Allocation (LDA), Support Vector Machines (SVM), and Naïve Bayes (NB) classifiers with unique strengths, are employed in SA. In the context of teaching research, all these models help recognize novel sentiments related to complex subjects like MH and JC performance[4].

This work embarks on an experimental setup across several higher education institutions in Nanjing province to unravel the relationship between MH and JC among CTs. This work implements a Mixed-Method Approach (MMA) that combines quantitative analysis of OESF data with qualitative performances. The primary role of the proposed method is a state-of-the-art application of SA that mainly focuses on Open-Ended Survey Feedback (OESF). We quantify the relationship between MH and JC using an LDA-based SA algorithm and qualitatively dissecting educators' sentiments and ideas. This dual method of contributing to statistical study with the abstraction of thematic visions over an SA algorithm is the novelty of this work. The work offers a multilayered understanding of how CT'-MH correlates to their skilled value and its more significant outcome. The outcomes of this study are predictable to contribute value to the teaching and learning policy-making system, CT support courses, and overall address while enhancing the quality of education and the well-being of CT.

These studies emphasize the complex interplay of MH, EI, SC, and JC in the educational sector. The proposed research aims to build on these findings by employing an MMA and integrating quantitative and qualitative analyses, including advanced SA techniques, to explore these dynamics among CTs. By experimenting with the proposed context of three provincial universities in China, this research study seeks to contribute novel insights into how MH and JC interact and their broader suggestions for educational policy and CT support systems.

The paper's structure is as follows: Section 2 delves into the literature survey; Section 3 covers the background study; Section 4 introduces the proposed methodology and data collection model; Section 5 details the study analysis; and Section 6 provides the concluding remarks.

2. Literature review

This literature review examines recent research exploring the intricate relationship between MH, EI, Social Competencies (SC), and JC among educators focused on understanding how these elements interplay within the educational domain.

A study[5] on the MH of CTs by examining its connection to professional competence and social support. Their findings reveal a significant negative correlation between MH and JC. Notably, competence significantly mediates CTs' MH and social support. Their study shows the importance of considering holistic well-being in educators in the academic environment. The research[6] delved into the impact of MH activities on CTs' well-being and work-related behaviors. Utilizing theories like Broaden and Build (B&B) and Conservation of Resources (COR), the study found that MH activities are positively associated with well-being and healthy work-related behavior while reducing unhealthy work experiences. The study with a sample of German CTs highlights the potential benefits of MH activities to promote MH among educators[7].

This study focused on[8] Emotional Intelligence Competencies (EIC) among higher education educators. Using Structural Equation Modeling (SEM), researchers established that EIC significantly impacts educator behavior and student success. The study emphasizes the need in higher education institutes to recruit and train educators with high EI skills to ensure effective teaching and remarkable performance. This investigate[9] based on the relationship between EI, SC, and challenges faced by CTs educating students with special needs; their study found significant differences between EI and SC, especially when working among students with intellectual disabilities and mental illnesses[10]. The findings highlight the importance of equipping CTs with advanced soft skills to effectively navigate such complexities of diverse educational needs. The study[11] of language CT struggle, Wang and colleagues explored the connection between psychological well-being, work engagement, and CT struggle among English as a Foreign Language (EFL) teachers in Asia. Their findings indicate that psychological well-being and work engagement positively impact CT struggle, with psychological well-being being a more robust predictor. The study underscores the need for a psychologically healthy work environment for short-term CT commitment and resilience[12–15].

3. Background

A recommendation is presented for a methodology that depends on LDA and can be employed to apply the SA to OESF. The LDA is the generative probabilistic model predominantly implemented to identify hidden conceptual features from large text corpora. It determines subjects, each of which is the probability distribution over phrases[16–18]. At this point, all of these subjects are calculated by applying a hierarchical Bayesian process, in which every sentence in the document can be represented as a combined collection of a number comprising various subjects[19,20]. As a result of its ability to successfully analyze a significant quantity of unlabeled textual data, this model is a highly beneficial tool for ML[21–25]. This research project uses LDA to find sentiment-laden topics from qualitative data for evaluating CTs' MH and JC feedback[26–30]. Thus, providing an understanding of OESF and experiences. The above technique clarifies how MH and JC relate in the teachers' words.

In LDA, data is illustrated mathematically in the following format:

1) *Word Representation: A word within a document is represented as an element in a vocabulary, indexed by* $\{1,2,\ldots,V\}$.
2) *Document Representation: A document is a sequence of* N *words, denoted as* $w=(w_1,w_2,\ldots,w_N)$, *where* w_n *is the* n^{th} *word in the sequence.*
3) *Corpus Representation: The corpus that comprises the entire collection of documents is represented as* $D=\{W_1,W_2,\ldots,W_M\}$, *where* M *is the total number of documents.*

The generative process for each document w *for the corpus* D *is modelled as follows:*

1) Select the length for document N from a Poisson distribution with parameter ξ, represented as $N \sim \text{Poisson}\ (\xi)$.
2) Choose the distribution upon topics θ in a Dirichlet distribution with parameter α, denoted as $\phi \sim \text{Dir}\ (\lambda)$.

3) For each of the N words w_n in the document:

- Select the topic z_n from a Multinomial distribution based on θ, expressed as $z_n \sim \text{Multinomial}\ (\phi)$.
- Choose the word w_n using topic z_n and a topic-word distribution parameter β, formulated as $w_n \sim p(w_n \mid z_n, \beta)$.

The distribution of topics in a document is characterized by a multidimensional Dirichlet random variable ϕ. This variable ϕ assumes a value of within a (k-1)-dimensional simplex, which means each component of the k-dimensional vector ϕ (denoted as ϕ_i) are non-negative, and their sum is 1, as represented by the condition $\phi_i \geq 0$ and $\sum_{i=1}^{k} \phi_i = 1$. The probability density of ϕ over simplex and given the parameter vector λ (an alternative to α in standard LDA notation) are expressed by:

$$P(\phi \mid \lambda) = \frac{\Gamma\left(\sum_{i=1}^{k} \lambda_i\right)}{\prod_{i=1}^{k} \Gamma(\lambda_i)} \times \phi_1^{\lambda_1 - 1} \phi_2^{\lambda_2 - 1} \ldots \phi_k^{\lambda_k - 1} \tag{1}$$

Here, λ is a k-dimensional vector with positive components $\lambda_i > 0$, and $\Gamma(x)$ represents the gamma function. This formulation defines the probabilistic foundation for the distribution of topics within each document as analyzed by our adapted LDA model for SA. The joint distribution of the topic mixture ϕ, a set of N sentiment categories z, and a corresponding set of N words w, given parameters λ and ω, is expressed as:

$$P(\phi, z, w \mid \lambda, \omega) = P(\phi \mid \lambda) \times \prod_{n=1}^{N} \left(P(z_n \mid \phi) \times P(w_n \mid z_n, \omega)\right) \tag{2}$$

Here, $P(\phi \mid \lambda)$ denotes the probability of a particular topic mixture ϕ under the Dirichlet distribution parameterized by λ. $P(z_n \mid \phi)$ is the probability of assigning a sentiment category z_n to a word, based on the topic mixture $\phi.P(w_n \mid z_n, \omega)$ represents the likelihood of observing the word w_n from a sentiment-specific word distribution for a given sentiment class z_n and word distribution parameters ω. Integrating over the topic mixture ϕ and summing over the sentiment classes z, the marginal distribution of a document in our adapted model is derived as follows:

$$P(w \mid \lambda, \omega) = \int P(\phi \mid \lambda) \left(\prod_{n=1}^{N} \sum_{z_n} P(z_n \mid \phi) P(w_n \mid z_n, \omega) \right) d\phi \tag{3}$$

This expression represents the probability of observing a specific sequence of words w in a document, accounting for the variability in topic distributions and the corresponding word distributions within each sentiment class. This study takes the product of these marginal probabilities for each document to determine the overall probability of the corpus, a collection of such documents. This gives us the probability of the corpus D as:

$$P(D \mid \lambda, \omega) = \prod_{d=1}^{M} \int P(\phi_d \mid \lambda) \left(\prod_{n=1}^{N} \sum_{z_{dn}} P(z_{dn} \mid \phi_d) P(w_{dn} \mid z_{dn}, \omega) \right) d\phi_d \tag{4}$$

In this equation: $P(\phi_d \mid \lambda)$ is the probability of the topic mixture for document d, under the Dirichlet parameters λ. The internal sum and product calculate the likelihood of each word in document d based on its assigned sentiment class and the word distribution for that group, parameterized by ω.

4. Proposed work

4.1. Objectives

The primary objective of this study is to explore the intricate relationship between MH and JC among CTs. This paper aims to understand how MH sentiments expressed by CTs correlate with their perceived JC, including aspects such as smart teaching abilities[31–34]. By analyzing these relationships, the study seeks to contribute to the broader understanding of teacher well-being and its impact on educational outcomes.

Based on the objective, the study proposes the following hypotheses:

1) **CTs with positive MH sentiments express higher JC**: This hypothesis posits that CTs who exhibit or express positive MH sentiments are likely to demonstrate higher levels of JC. The underlying assumption is that good MH contributes to effective teaching and job performance.

2) **CT's JC is related to positive MH**: Here, we explore the possibility of a bidirectional relationship, where not only does positive MH led to better JC, but the experience of being competent in one's job could also enhance a CT's MH. This hypothesis acknowledges the cyclical nature of MH and JC performance.

3) **CT's innovative teaching ability is related to the positive relationship between MH and JC: This hypothesis extends the investigation to a specific aspect of JC's** innovative teaching ability. It suggests that CTs' capacity to be innovative and creative in their teaching methods is influenced by the interplay between their MH and overall JC.

Through these hypotheses, the study intends to shed light on the multifaceted connections between a CT's MH and their professional efficacy, particularly in higher education.

4.2. Data collection

To test the above hypotheses of this study, an OESF was conducted for the CTs. Before participating in the OESF, the participants were provided with detailed information about the study and were required to give their informed consent. This ensures ethical compliance and respect for the participant's autonomy. The questionnaire is designed to ensure confidentiality and anonymity, allowing participants to withdraw from the OESF at any time without any consequences. This approach ensured that no data was collected without explicit consent from the participants.

The questionnaire was crafted through a multi-step process:

1) **Initial interviews**: The initial phase involves conducting interviews among educators across several areas (science, engineering, arts, management, and social science) at a university located in Nanjing province. The interviews are for acquiring a comprehension of the fundamental principles of the study, namely concerning JC and the MH of tutors.

2) **Drafting the questionnaire**: A draft questionnaire is created by incorporating insights from interviews and data from prior investigations. The objective of this draft was to accurately depict the particulars of CTs' MH and their perceived level of JC.

3) **Pilot study**: A qualitative study was undertaken to evaluate the initial questionnaire, and the feedback obtained from this limited-scale OESF was used to improve the questionnaire, guaranteeing its pertinence and comprehensibility.

4) **Finalization**: The questionnaire was finalized based on the pilot study's findings, guaranteeing its correctness and usefulness in assessing the targeted components.

The finalized questionnaire is emailed to CTs across different positions (lecturers, assistant professors, associate professors, and professors) in China. The OESF will be conducted from March to April 2022, with participants from various universities in provinces such as Nanjing, Hubei, and Shandong. Approximately 250 questionnaires were sent to multiple disciplines, including medicine, agriculture, fine arts, science and technology, law and economics, and music. OESF was selected to have a high completion rate, and the samples with a completion rate below 70% were excluded from the study. In the end, 116 valid samples were collected and included in the analysis. The samples represent the broad spectrum of academic disciplines and provide a comprehensive view of the relationship between MH sentiments and JC among CTs. **Table 1** presents the characteristics of the samples.

Table 1. Characteristics of OESF participants.

Characteristic	**Category**	**Percentage of participants (%)**
Experience	0–5 years	20%
	6–10 years	30%
	11–15 years	25%
	16–20 years	15%
	20+ years	10%
Age	25–30 years	15%
	31–35 years	25%
	36–40 years	20%
	41–45 years	15%
	46+ years	25%
Highest degree	Bachelor's	10%
	Master's	40%
	PhD	50%
Academic stream	Science and technology	25%
	Medicine and health	20%
	Arts and humanities	15%
	Social sciences	20%
	Business and management	10%
	Other	10%
Province	Nanjing	35%
	Hubei	30%
	Shandong	20%
	Other	15%

4.3. Measurement

In this study, a questionnaire consisting of 36 questions was designed to assess numerous aspects related to the JC of CTs and their MH status. The variables listed below have been addressed in the OESF: JC (the dependent variable), MH status (the independent factor), novel teaching skills (the mediating variable), and several types of control factors (demographic information and professional history). In the questionnaire, participants are advised to rate many different variables using 5-point scales based on Likert ratings. The structure of the paper specified a conventional scale from "strongly disagree" to "strongly agree", "never", to "always", and so on, allowing listeners to choose their degree of agreement or frequency.

These numerical problems were matched by open-ended textual ones in the questionnaire in order to elicit qualitative data. The OESF to these questionnaires provide precious details about how the respondents dealt with mental health and their level of JC. Examples of such questions include "Please describe your experience regarding the impact of MH on teaching abilities" and "How do you feel your emotional state impacts students' performance in innovative teaching practices?" The section on control variables consisted of categorical or numerical questions, such as age, years of teaching experience, and highest degree obtained. This combination of quantitative and qualitative data enriches the dataset, allowing for a more comprehensive analysis using statistical methods and SA. **Table 2** provides a detailed overview of each variable in the study, illustrating examples of the questionnaire objects used, their corresponding Cronbach's alpha values for Likert-scale items, and a brief description of the open-ended questions.

Table 2. Measurement of variables.

Variable type (No. of questions)	Variable description	Example questionnaire items	Cronbach's α
Dependent variable (10)	JC of CTs	1. I am confident in my teaching abilities. 2. I can effectively engage and motivate students. 3. I am proficient in developing innovative teaching methods. 4. Describe a situation where you felt exceptionally competent in your teaching role	0.85
Independent variable (10)	MH status	1. I feel content with my overall emotional state with this job. 2. I seldom feel stressed or overwhelmed due to work. 3. I am satisfied with my work-life balance. 4. Share your experiences or challenges related to maintaining MH in your teaching profession	0.88
Mediating variable (10)	Innovative teaching ability	1. I frequently try new teaching methods or tools. 2. I am enthusiastic about experimenting with new educational technologies. 3. I can adapt well to different teaching scenarios. 4. Can you provide an example of your innovation in your teaching approach?	0.82
Control variables (6)	Demographic and professional background	1. Age. 2. Highest academic degree. 3. Years of teaching experience.	N/A

4.4. Data preprocessing

The preprocessing commenced with a thorough cleaning process, identifying approximately 5% of missing values, predominantly in the sections assessing MH and JC. To address this, a mean imputation method was employed for Likert-scale items, replacing missing values with the corresponding item's average OESF, minimizing the impact on the overall data patterns. Preprocessing Likert-Scale Responses (LSR): LSR, which form a significant part of the OESF, were normalized by converting them into numerical values ranging from 1 to 5. For categorical data like the academic discipline and type of institution, the one-hot encoding was applied to convert these nominal values into a binary format suitable for analysis.

A. Preprocessing OESF: The OESF are preprocessed using:

- *Text Cleaning:* Initial steps involve the removal of noise like the unique characters and the numbers and standardizing the text.
- *Tokenization and Lemmatization:* The text is split into individual words or phrases, and the words are reduced to their base and root form to ensure consistency.
- *Stop Words Removal:* Common words with little analytical value are removed to focus on the more meaningful content in the OESF.

B. Data Reduction and Transformation: Principal Component Analysis (PCA) is utilized to reduce the dimensionality of the dataset, focusing on the most informative variables. The statistics collected from the LSR were vital for this process stage.

In order to maintain the confidentiality of participants and follow research ethics standards, every personally identifiable detail has been eliminated or anonymized. Verifying that the LSR for MH and JC are in line with internal standards demonstrates the dataset's trustworthiness. Approaches to preliminary processing are being enhanced with the help of feedback collected through the initial investigation. A high value is given to accuracy, reliability, and compliance with ethical principles when preparing the final dataset, which includes both normalized OESF on a Likert scale and textual data that has been automated.

4.5. Sentiment-oriented topic analysis: An LDA-based approach

The main objective of this section of the paper is to suggest an LDA-based algorithm for the goal of SA of OESF from college instructors. The methodology uses SEM to find and evaluate key topics reflecting several MH and JC ideas. The steps start with the preliminary processing of the open text of the OESF in order to extract word unigrams and then move on to the actual use of LDA for topic discovery. It follows with an investigation to find out the related sentiment polarity of the subjects, such as once they have been identified. After that, a method of classification such as SVM or Naive Bayes is applied in order to label the sentiments. A summary of the research results can be seen in **Table 3**, which includes a graphic representation of the key subjects and emotional reactions that could be determined from the collected feedback.

Table 3. Summary of topics and sentiments identified from OESF.

Topic number	Identified topic	Sentiment orientation	Example phrases from OESFs
1	Work-life balance	Mostly positive	"Satisfactory work-life harmony", "Adequate time for personal life"
2	Teaching resources availability	Mixed	"Need for better resources", "Satisfactory classroom tools"
3	Institutional support	Mostly negative	"Lack of administrative support", "Feeling undervalued by the institution"
4	Student engagement challenges	Mixed	"Challenges in online engagement", "Rewarding classroom interactions"
5	Personal MH	Mostly negative	"Stress and anxiety", "Concerns about burnout"
6	Professional development	Mostly positive	"Opportunities for growth", "Valuable training sessions"

Algorithm 1 Topic Modeling & Sentiment Topic Mapping for OESF

This algorithm takes the corpus of OESF as input. Initially, the corpus is pre-processed to remove stop words and symbols, enhancing the data quality for analysis. LDA is then performed on this pre-processed corpus to conduct topic modelling, extracting numerous topics. Each topic, representing a key theme or sentiment within the OESF, is mapped to specific sentiment-oriented aspects. The algorithm scans the corpus to identify sentences that align with the words describing each topic, extracting these as topic-specific sentences. These sentences are then grouped and stored, considered by their corresponding issues.

Algorithm: Sentiment-Oriented Topic Analysis (SOTA) for OESF

Input:

A set of OESF, $S = \{s_1, s_2, \dots, s_M\}$

Output:

A set of sentiment-oriented topics $T = \{t_1, t_2, \dots, t_K\}$
A mapping of each OESF to a sentiment-oriented topic

Algorithm Steps:

1: Preprocessing:
For each OESF, s_i in S :
Remove stop-words and symbols to obtain a clear OESF s_i'.

2: Topic Modeling with LDA:
Apply LDA on the preprocessed OESFs $\{s_1', s_2', \dots, s_M'\}$ to extract topics.
Let $LDA(s_i') \rightarrow \{l_1, l_2, \dots, l_K\}$ be the set of topics for OESF, s_i' , where K is the number of topics.

3: Sentiment Analysis:
For each topic l_k :
Analyze the word distribution within l_k to determine the dominant sentiment d_k.
Assign a sentiment label to each topic, resulting in sentiment-oriented topics t_k.

4: OESF to Topic Mapping:
Initialize an empty map M.
For each OESF s_i' :
Determine the most relevant topic t_k based on $LDA(s_i')$.
Update the map $M(s_i') = t_k$.

5: Output Generation:
Return the set of sentiment-oriented topics T and the mapping M.

End of Algorithm

Algorithm 2 Topic-Specific Sentiment Analysis

Algorithm 2 processes the topic-specific sentences obtained from **Algorithm 1**. It evaluates the sentiment polarity score of each sentence. The polarity score is the determinant of the sentence's sentiment; a positive score indicates a "Positive" sentiment, a negative score indicates a "Negative" sentiment and a zero score corresponds to a "Neutral" sentiment. This step is crucial for classifying the sentiments of OESF related to each topic, allowing for a nuanced understanding of the sentiment distribution within the OESF data.

Algorithm: Topic-Specific Sentiment Analysis (TSSA) for OESF

Input:
A file containing sentences mapped to specific topics: File = ["Topic – SpecificSentences"]
Output:
Sentiment polarity scores and sentiment classes for each sentence
Algorithm Steps:
1: Initialize SA:
Read the file containing topic-specific sentences.
2: Iterate Over Sentences:
For each sentence line in the file:
Calculate the sentiment polarity score PolarityScore(line).
3: Determine Sentiment Category:
For each line:
If PolarityScore(line) > 0 :
Assign Sentiment(line) = "Positive".
Else if PolarityScore(line) < 0 :
Assign Sentiment(line) = "Negative".
Else if PolarityScore(line) $= 0$:
Assign Sentiment(line) = "Neutral".
4: Output Sentiment Results:
Return the sentiment polarity scores and corresponding sentiment classes for each sentence in the file.
End of Algorithm

Algorithm 3 Overall Sentiment Assessment for Topics

The final algorithm operates on the aggregate polarity counts of sentences associated with each topic. It assesses the overall sentiment for each subject based on the polarity scores. If the count of positive polarity sentences outweighs the negative and neutral ones, the general sentiment for that topic is labelled as "Positive". Conversely, if negative polarity sentences dominate, the sentiment is "Negative". If neutral counts are higher, the sentiment is deemed "Neutral". This algorithm provides an overarching sentiment perspective for each topic, summarizing the predominant emotional tone reflected in the OESF.

Algorithm: Overall Sentiment Assessment for Topics (OSAT)

Input:

Polarity counts of sentences corresponding to each topic: "Polarity Count per Topic"

Output:

Aggregate sentiment for each topic

Algorithm Steps:

1: Initialize Sentiment Counters for Each Topic:
For each topic T_i from the set of topics T :
Initialize PositiveCount $T_i = 0$
Initialize NegativeCount $_{T_i} = 0$
Initialize NeutralCount $T_{T_i} = 0$

2: Aggregate Polarity Counts:
For each topic T_i, accumulate the polarity counts:
PositiveCount $T_{T_i} \leftarrow$ Total count of positive polarity sentences
NegativeCount $T_i \leftarrow$ Total count of negative polarity sentences
NeutralCount $T_{T_i} \leftarrow$ Total count of neutral polarity sentences

3: Determine Overall Sentiment for Each Topic:
For each topic T_i :
If PositiveCount $T_{T_i} >$ (NegativeCount $_{T_i} +$ NeutralCount T_i):
OverallSentiment $T_i =$ "Positive"
Else if NegativeCount $T_{T_i} >$ (PositiveCount $_{T_i} +$ NeutralCount T_i):
OverallSentiment $T_i =$ "Negative"
Else if NeutralCount $T_i >$ (PositiveCount $_{T_i} +$ NegativeCount $_{T_i}$):
OverallSentiment $T_i =$ "Neutral"
Else:
OverallSentiment $T_i =$ "Indeterminate"

4: Output Aggregate SA:
Return the overall sentiment for each topic based on the aggregated polarity counts.

End of Algorithm

5. Analysis of the hypothesis

5.1. Analysis of variables

The study used SPSS for statistical analysis to investigate the interplay between MH and JC in CTs. **Table 4**'s correlation matrix revealed moderate relationships among the primary variables, confirming that they maintain distinct conceptual identities while interconnected. The research concentrated on three principal variables: the dependent variable of JC, the independent variable of MH Status, and the mediating variable of Innovative Teaching Ability. Demographic and professional background characteristics, including age and teaching experience, were control variables. Descriptive statistics showed a positive skew in JC perceptions, with an average score of 3.5 and a standard deviation of 0.8 on a 5-point Likert scale. Multicollinearity was assessed with VIFs for each independent variable, all falling below the threshold of 10, indicating no multicollinearity issues. Hierarchical linear modelling (**Table 5**) yielded insightful results; Model 2 substantiated a significant positive correlation between MH Status and JC ($\beta = 0.256$, $p < 0.01$), while Model 3 demonstrated a strong positive influence of innovative teaching ability on JC.

Table 4. Correlation coefficients among study variables.

Variables	1	2	3	4	5	6	7	Mean	S.D.
Age	1	-	-	-	-	-	-	42.7	9.4
Teaching experience (years)	0.21*	1	-	-	-	-	-	15.3	7.5
MH status	–0.03	0.14	1	-	-	-	-	3.8	0.5
Innovative teaching ability	0.04	0.19*	0.28*	1	-	-	-	3.7	0.6
JC	0.01	0.17*	0.36*	0.47*	1	-	-	3.9	0.7

Note: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 5. Hierarchical linear modeling results for JC.

Variables	Model 1	Model 2	Model 3
Age	0.02	−0.01	0.03
Teaching experience	−0.05	0.04	0.06*
MH status		0.25**	
Innovative teaching ability			0.31***
R^2	0.05	0.10	0.15
Adjusted R^2	0.04	0.09	0.14
F-value	2.53	4.67*	6.81**

Note: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Mediation analyses, presented in **Table 6**, elucidated the pivotal role of innovative teaching ability in bridging MH Status and JC, with a notable mediating effect. This suggests that innovation in teaching significantly leverages the positive impact of MH on JC. The reliability of the OESF scales was confirmed with Cronbach's alpha values signifying high internal consistency: 0.85 for job competence, 0.88 for MH status, and 0.82 for innovative teaching ability. Control variables were meticulously chosen to address endogeneity, with subsequent analyses underscoring the absence of significant endogeneity, thus bolstering the study's findings. The hierarchical linear modeling, supported by **Tables 5** and **6**, validates the hypothesized positive associations within the study's context and the mediating effect of innovative teaching ability, reinforcing the importance of MH in educational JC.

Table 6. Hierarchical linear modeling results for mediation analysis.

Variables	Model 4	Model 5	Model 6	Model 7
JC (dependent)				
Innovative teaching ability (mediator)	0.44***	0.42***		
MH status (independent)	0.13	0.11		
Work engagement (dependent)			0.28***	
Innovative behavior (dependent)				0.33***
R^2	0.27	0.26	0.30	0.35
Adjusted R^2	0.26	0.25	0.29	0.34
F-value	7.91***	7.44***	9.28***	10.52***

Note: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

5.2. Analysis of the proposed SA algorithm for the OESF

In this section, the results obtained from the LDA algorithm are presented. **Table 7** showcases the primary topics extracted by the algorithm, limited to a few due to space constraints. These topics, derived as word unigrams, reflect key themes from the OESF related to MH and JC among CTs. Each identified topic is carefully examined to discern its predominant sentiment orientation. The results of this SA, detailing the sentiment polarity for each topic, are displayed in **Table 8**. This table helps to understand the general sentiment—whether primarily positive, negative, or mixed—associated with each subject based on the language used in the OESF.

Table 7. Result of topic-sentiment mapping.

Topic number	Topic words	Sentiment orientation
1	Workload, grading, curriculum, lesson planning	Work-life balance
2	Interactive whiteboards, projectors, textbooks	Teaching resources
3	Administrative support, policy, recognition	Institutional support
4	Student participation, online platforms, engagement	Student engagement
5	Stress management, self-care, workload	Personal mental health
6	Workshops, conferences, professional growth	Professional development

Table 8. Polarity score and sentiment of OESF for the topic "work-life balance".

#	OESF excerpt from CTs	Polarity score	Sentiment	Accuracy
1	Balancing grading and personal life are challenging.	−0.15	Negative	Correct
2	I manage to maintain an excellent work-life balance.	0.25	Positive	Correct
3	Work often spills into my weekends.	−0.20	Negative	Correct
4	I've found a rhythm between teaching and family time.	0.30	Positive	Correct

Table 8 presents the results of analyzing textual data from OPSF. It maps the polarity scores and sentiment to the CTs' comments regarding work-life balance. The scores correspond to the SA algorithm's interpretation in each OESF, and the accuracy column indicates whether the sentiment was accurately classified based on manual validation. **Table 9** offers a breakdown of polarity counts across the identified topics. This quantifies the sentiment expression and calculates how frequently positive, negative, or neutral sentiments are expressed concerning each topic.

Table 9. Polarity count for identified topics from OESF.

Topic	Total	Positive	Negative	Neutral
Work-life balance	1667	750	583	333
Teaching resources availability	1333	500	467	367
Institutional support	2000	333	1167	500
Student engagement challenges	1500	667	583	250
Personal MH	1833	417	1000	417
Professional development	2167	1500	333	333

Table 10 summarizes the overall findings of the topic-oriented SA. It presents the insight conceived from the previous tables that show a comprehensive view of the collective sentiments of CTs as related to numerous elements of their professional experiences and MH.

Table 10. Result of topic-oriented SA.

No	Identified topic	Overall sentiment
1	Work-life balance	Mixed
2	Teaching resources availability	Positive
3	Institutional support	Negative
4	Student engagement challenges	Mixed
5	Personal MH	Negative
6	Professional development	Positive

To conduct a thorough assessment of the model's effectiveness, we employed Receiver Operating Characteristics (ROC) plots to examine the True Positive Rate (TPR) and False Positive Rate (FPR) for different topics found from the OPSF with a NB classifier. ROC plots for the topics "work-life balance", "Teaching Resources Availability", "Institutional Support", "Student Engagement Challenges", "Personal MH", and "Professional Development" are illustrated in **Figure 1(a–f)**, respectively.

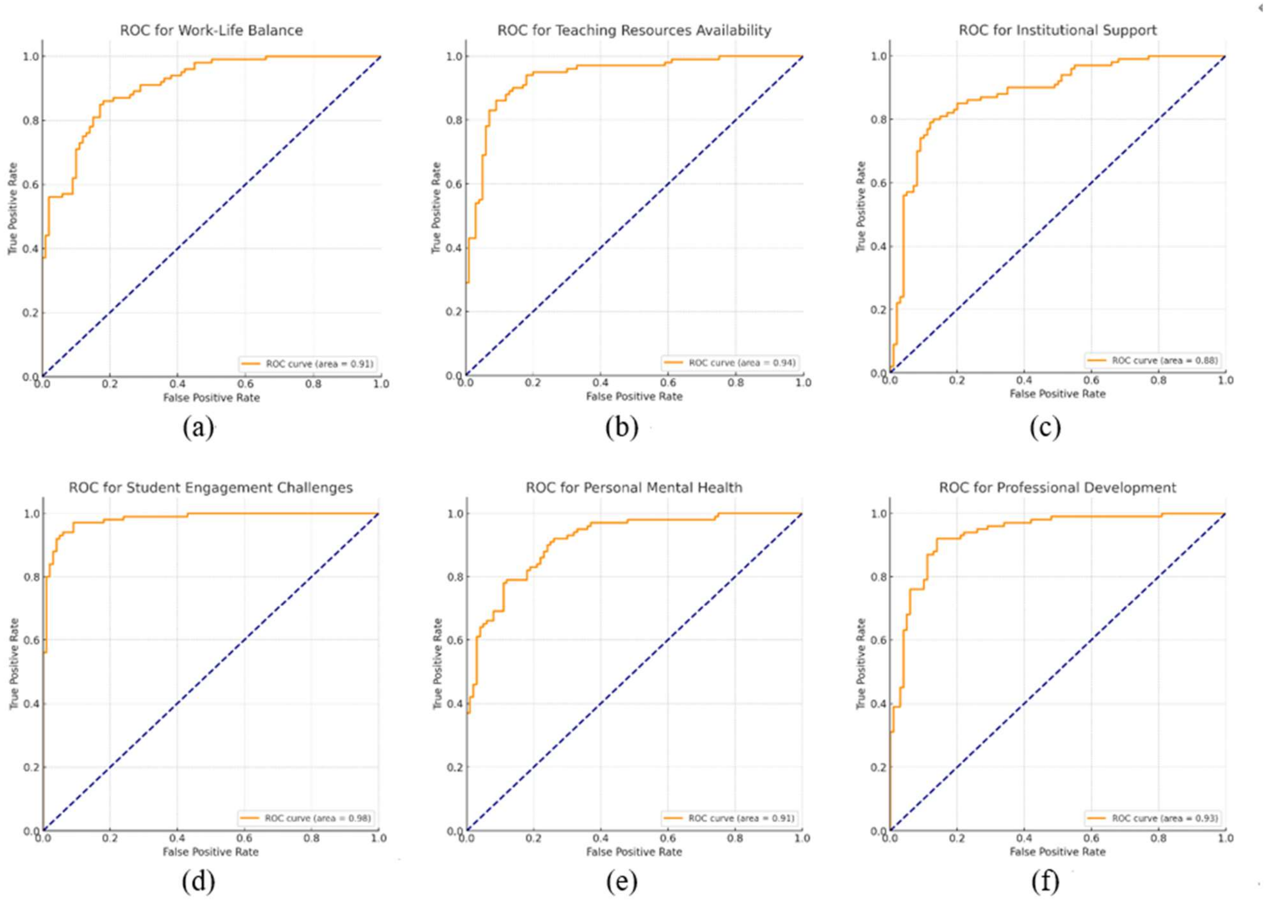

Figure 1. ROC for **(a)** work-life balance; **(b)** teaching resources availability; **(c)** institutional support; **(d)** student engagement challenges; **(e)** personal MH; **(f)** professional development.

5.3. Comparative analysis of classifier performance

This section conducted a comparative analysis to evaluate the performance of three classifiers, NB, SVM, and LR, for six topics identified in our SA research. The performance metrics for comparison are accuracy, precision, recall, and F1-score, and the results are in **Table 11**.

Table 11. Comparison between the classifiers.

Topic	Naïve Bayes				SVM				LR			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score
Work-life balance	96.30	87.32	90.75	92.09	79.87	74.26	89.26	82.70	88.09	80.79	90.01	87.39
Teaching resources availability	93.59	88.46	87.21	87.32	86.94	89.47	84.84	74.74	90.26	88.97	86.02	81.03
Institutional support	93.00	90.35	87.03	87.46	86.14	83.50	76.91	79.51	89.57	86.92	81.97	83.48
Student engagement challenges	97.43	95.46	91.35	91.62	89.94	87.22	74.18	80.77	93.69	91.34	82.76	86.20
Personal mental health	95.37	87.41	83.16	78.28	84.49	85.14	81.55	76.38	89.93	86.27	82.35	77.33
Professional development	97.67	97.95	87.16	88.88	84.46	45.10	104.61	63.02	91.07	71.52	95.89	75.95

For work-life balance, the NB demonstrates better accuracy and good balance among precision, recall, and F1-score. The SVM classifier exhibits a lower accuracy and precision and shows a high recall, which indicates effectiveness for identifying relevant cases. LR presented a balanced performance, making it an alternative with consistent results in all metrics. For Teaching resource availability, the NB again shows high accuracy and better performance in all metrics. SVM decreased recall and F1-score while the precision was high. LR surpassed SVM in accuracy and displayed an impressive balance of performance metrics. Upon reviewing the performance for the institutional support topic, the Naïve Bayes maintained a high level of performance, particularly in precision and recall. The SVM classifier's performance declined notably for Recall, F1-score, and LR, offering a balanced set of results. In the case of Student Engagement Challenges, the NB excelled, achieving the highest accuracy and F1-score among classifiers. SVM showed a weaker performance, especially in recall, and LR delivered robust results, especially for precision and recall. Regarding Personal MH, NB demonstrated a high level of accuracy, where it had difficulties in precision and recall. SVM showed slightly lower performance with precision, whereas LR displayed a balanced performance.

Similarly, in professional development, NB stood out with high accuracy, excellent precision, and recall. SVM showed a lower F1-score, reflecting a significant imbalance of precision and recall. LR displayed high accuracy and a more balanced metric distribution than SVM. Across all topics, NB consistently demonstrated high accuracy and a better balance of precision, recall, and F1 score, making it the most effective classifier for SA.

6. Conclusion

The comprehensive OESF conducted across numerous universities in Nanjing Province has yielded important insights into the relationship between MH and JC among College Teachers (CT). The findings indicated a significant correlation between educators' mental well-being and their perceived professional efficacy. CTs who reported higher levels of MH positivity were likelier to demonstrate enhanced JC, including innovative teaching methods and effective student engagement. Furthermore, using a Latent Dirichlet Allocation (LDA)-based algorithm, the Sentiment Analysis (SA) found important topics in CTs'-OESF. The findings show that a prevalent sentiment was the need for excellent institutional support in addressing MH concerns, which many educators felt directly impacted their JC and satisfaction.

Additionally, the OESF underscored the importance of work-life balance with CT well-being and the subsequent effect on teaching efficacy. The study's Mixed-Method Approach (MMA) combines quantitative data analysis with qualitative SA, providing an understanding of the intricate dynamics between MH and JC in the educational sector. These findings show the importance of developing a sympathetic environment and policies within educational institutions that must address the MH requirements of CTs. By doing this, there is potential to enhance educators' well-being and elevate the overall quality of education.

Author contributions

Conceptualization, XW and LZ; methodology, XW; software, LZ; validation, LZ, and XW; formal analysis, XW; investigation, XZ; resources, LZ; data curation, LZ; writing—original draft preparation, XW; writing—review and editing, LZ; visualization, LZ; supervision, LZ; project administration, XW. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Reference

1. Mirbahaeddin E, Chreim S. A Narrative Review of Factors Influencing Peer Support Role Implementation in Mental Health Systems: Implications for Research, Policy and Practice. Administration and Policy in Mental Health and Mental Health Services Research. 2022, 49(4): 596-612. doi: 10.1007/s10488-021-01186-8
2. Greenier V, Fathi J, Behzadpoor SF. Teaching for creativity in an EFL context: The predictive roles of school climate, teaching enthusiasm, and metacognition. Thinking Skills and Creativity. 2023, 50: 101419. doi: 10.1016/j.tsc.2023.101419
3. Derakhshan A, Wang Y, Wang Y, et al. Towards Innovative Research Approaches to Investigating the Role of Emotional Variables in Promoting Language Teachers' and Learners' Mental Health. International Journal of Mental Health Promotion. 2023, 25(7): 823-832. doi: 10.32604/ijmhp.2023.029877
4. Alkhnbashi OS, Nassr RM. Topic Modelling and Sentimental Analysis of Students' Reviews. Computers, Materials & Continua. 2023, 74(3): 6835-6848. doi: 10.32604/cmc.2023.034987
5. Peng R, Photchanachan S. Influence of mental health status of university teachers on professional competence. The Journal of Sirindhornparithat. 2021, 22(2): 241-255.
6. Hofmann H, Groß D, Kohlmann CW. On the Role of Mental Health Activities for Teachers' Work and Life. Applied Research in Quality of Life. 2020, 17(1): 205-227. doi: 10.1007/s11482-020-09885-4
7. Khassawneh O, Mohammad T, Ben-Abdallah R, et al. The Relationship between Emotional Intelligence and Educators' Performance in Higher Education Sector. Behavioral Sciences. 2022, 12(12): 511. doi: 10.3390/bs12120511
8. Skura M, Świderska J. The role of teachers' emotional intelligence and social competences with special educational needs students. European Journal of Special Needs Education. 2021, 37(3): 401-416. doi: 10.1080/08856257.2021.1885177
9. Wang Y, Derakhshan A, Azari Noughabi M. The interplay of EFL teachers' immunity, work engagement, and psychological well-being: Evidence from four Asian countries. Journal of Multilingual and Multicultural Development. 2022, 1-17. doi: 10.1080/01434632.2022.2092625
10. Kumar S, Rani S, Jain A, et al. Face Spoofing, Age, Gender and Facial Expression Recognition Using Advance Neural Network Architecture-Based Biometric System. Sensors. 2022, 22(14): 5160. doi: 10.3390/s22145160
11. Borkar P, Wankhede VA, Mane DT, et al. Deep learning and image processing-based early detection of Alzheimer disease in cognitively normal individuals. Soft Computing. 2023. doi: 10.1007/s00500-023-08615-w
12. Alnuaim AA, Zakariah M, Alhadlaq A, et al. Human-Computer Interaction with Detection of Speaker Emotions Using Convolution Neural Networks. Computational Intelligence and Neuroscience. 2022, 2022: 1-16. doi: 10.1155/2022/7463091
13. Kimmatkar NV, Babu BV. Novel Approach for Emotion Detection and Stabilizing Mental State by Using Machine Learning Techniques. Computers. 2021, 10(3): 37. doi: 10.3390/computers10030037
14. Balajee RM, Mohapatra H, Deepak V, et al. Requirements identification on automated medical care with appropriate machine learning techniques. In: Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT); 20–22 February 2021; Coimbatore, India. pp. 836-840. doi: 10.1109/icict50816.2021.9358683

15. Mannepalli K, Sastry PN, Suman M. Emotion recognition in speech signals using optimization based multi-SVNN classifier. Journal of King Saud University - Computer and Information Sciences. 2022, 34(2): 384-397. doi: 10.1016/j.jksuci.2018.11.012
16. Depuru S, Nandam A, Ramesh PA, et al. Human emotion recognition system using deep learning technique. Journal of Pharmaceutical Negative Results. 2022, 13(4): 1031-1035.doi: 10.47750/pnr.2022.13.04.141
17. Sekhar C, Rao MS, Nayani ASK, Bhattacharyya D. Emotion recognition through human conversation using machine learning techniques. In: Bhattacharyya D, Thirupathi Rao N (editors). Machine Intelligence and Soft Computing: ICMISC 2020. Springer; 2021. pp. 113-122. doi:10.1007/978-981-15-9516-5_10
18. Durga BK, Rajesh V. A ResNet deep learning based facial recognition design for future multimedia applications. Computers and Electrical Engineering. 2022, 104: 108384. doi: 10.1016/j.compeleceng.2022.108384
19. Ashok Kumar PM, Maddala JB, Martin Sagayam K. Enhanced Facial Emotion Recognition by Optimal Descriptor Selection with Neural Network. IETE Journal of Research. 2021, 69(5): 2595-2614. doi: 10.1080/03772063.2021.1902868
20. Bharti SK, Varadhaganapathy S, Gupta RK, et al. Text-Based Emotion Recognition Using Deep Learning Approach. Computational Intelligence and Neuroscience. 2022, 2022: 1-8. doi: 10.1155/2022/2645381
21. Thirumuru R, Gurugubelli K, Vuppala AK. Novel feature representation using single frequency filtering and nonlinear energy operator for speech emotion recognition. Digital Signal Processing. 2022, 120: 103293. doi: 10.1016/j.dsp.2021.103293
22. Kumar S, Haq M, Jain A, et al. Multilayer neural network based speech emotion recognition for smart assistance. Computers, Materials & Continua. 2023, 74(1): 1523-1540. doi: 10.32604/cmc.2023.028631
23. Srinivas PVVS, Mishra P. A novel framework for facial emotion recognition with noisy and de noisy techniques applied in data pre-processing. International Journal of System Assurance Engineering and Management. 2022. doi: 10.1007/s13198-022-01737-8
24. Mishra P, Srinivas PVVS. Facial emotion recognition using deep convolutional neural network and smoothing, mixture filters applied during preprocessing stage. IAES International Journal of Artificial Intelligence (IJ-AI). 2021, 10(4): 889. doi: 10.11591/ijai.v10.i4.pp889-900
25. Setiawan R, Devadass MMV, Rajan R, et al. IoT Based Virtual E-Learning System for Sustainable Development of Smart Cities. Journal of Grid Computing. 2022, 20(3). doi: 10.1007/s10723-022-09616-z
26. Irrinki MK. Learning through ICT role of Indian higher education platforms during pandemic. Library Philosophy and Practice. 2021, 1-19.
27. Lakshmi AJ, Kumar A, Kumar MS, et al. Artificial intelligence in steering the digital transformation of collaborative technical education. The Journal of High Technology Management Research. 2023, 34(2): 100467. doi: 10.1016/j.hitech.2023.100467
28. Madhavi E, Lavanya Sivapurapu, Vijayakumar Koppula, et al. B. Esther Rani. Developing Learners' English-Speaking Skills using ICT and AI Tools. Journal of Advanced Research in Applied Sciences and Engineering Technology. 2023, 32(2): 142-153. doi: 10.37934/araset.32.2.142153
29. Nanduri VNPSS, Sagiri C, Manasa SSS, et al. A Review of multi-modal speech emotion recognition and various techniques used to solve emotion recognition on speech data. In: Proceedings of the 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA); 3-5 August 2023; Coimbatore, India. pp. 577-582. doi: 10.1109/icirca57980.2023.10220691
30. Srinivas P, Khamar SN, Borusu N, et al. Identification of facial emotions in Hitech modern era. In: Proceedings of the 2023 2nd International Conference on Edge Computing and Applications (ICECAA); 19-21 July 2023; Namakkal, India. pp. 1202-1208. doi: 10.1109/icecaa58104.2023.10212285
31. Sugumar R, Sharma S, Kiran PBN, et al. Novel Method for Detection of Stress in Employees using Hybrid Deep learning Models. In: Proceedings of the 2023 8th International Conference on Communication and Electronics Systems (ICCES); 1-3 June 2023; Coimbatore, India. pp. 984-989. doi: 10.1109/icces57224.2023.10192609
32. Kuchibhotla S, Dogga SS, Vinay Thota NVSLG, et al. Depression Detection from Speech Emotions using MFCC based Recurrent Neural Network. In: Proceedings of the 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN); 5-6 May 2023; Vellore, India. pp. 1-5. doi: 10.1109/vitecon58111.2023.10157779
33. Chintalapudi KS, Patan IAK, Sontineni HV, et al. Speech Emotion Recognition Using Deep Learning. In: Proceedings of the 2023 International Conference on Computer Communication and Informatics (ICCCI); 23-25 January 2023; Coimbatore, India. pp. 1-5. doi: 10.1109/iccci56745.2023.10128612
34. Dimlo UMF, Bhanarkar P, Jayalakshmi V, et al. Innovative Method for Face Emotion Recognition using Hybrid Deep Neural Networks. In: Proceedings of the 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI); 11-13 April 2023; Tirunelveli, India. pp. 876-881. doi: 10.1109/icoei56765.2023.10126007