Received: 7 April 2025 | Revised: 19 August 2025 | Accepted: 13 November 2025 | Published online: 10 February 2026

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2026, Vol. 4(3) 365–377
DOI: 10.47852/bonviewJDSIS62025867

BON VIEW PUBLISHING

Predicting Mental Health Outcomes for Native and Non-native Language Students Through ICT-Based Learning Analytics

Kunal Ghosh[1] and Saiyed Umer[1,*]

[1] *Department of Computer Science and Engineering, Aliah University, India*

Abstract: This paper presents the impact and assessment of Information and Communication Technology (ICT)-based learning techniques employed by native and non-native language medium schools and predicts the factors that affect the mental health outcomes of students in these medium schools. The implementation of this study has been executed into five stages: (i) implication of reliable & aligned questionnaire associated with this study, (ii) targeting a group of students for data collection, (ii) performing the data reliability and data analysis tasks, (iv) deriving the key factors associated with the objective of this study, and (v) finally, building the prediction models as prototypes based on the textual analysis of the employed questionnaire for the mental health predictions among students. This study is being carried out using questionnaire-based data collected from 400 students in the Indian education system. Here, the medium of schools is Bengali (native) and English (non-native) languages. After data collection, rigorous statistical and machine learning-based data analysis techniques are employed. In the findings, Statistical tests (e.g., Mann–Whitney U, Kaiser-Meyer-Olkin (KMO), and Bartlett) revealed that English-medium students, particularly girls, reported higher levels of perceived stress compared to their Bengali-medium counterparts. Factor analysis confirmed more complex stress structures in English-medium groups, and predictive models achieved over 81% accuracy in classifying stress profiles. These findings suggest that language background is a significant factor in shaping students' psychological well-being in ICT-based education systems. The outcome of this study will be used to derive the key factors for mental health problems among students and build the prediction model that will be derivable from this study.

Keywords: ICT-based learning, mental health prediction, native and non-native medium schools, data analysis, predictive modeling, student well-being

1. Introduction

Education systems are different in every country, but some challenges are the same everywhere [1]. E-learning is a way of learning using digital technology. It is used in schools and at home. E-learning depends on Information and Communication Technology (ICT), which helps to store, share, and process information. ICT includes things like computers, the internet, and mobile phones [2]. These tools make online learning possible. ICT helps teachers and students by providing the technology needed for digital education. ICT connects students with people around the world and makes learning flexible anytime, anywhere. It also supports mental health by offering online help, relaxation apps, and ways to stay connected with friends and teachers. Personalized learning tools help students feel confident and engaged. A mental health prediction system in education helps track student well-being using ICT [3]. ICT has both benefits and risks for mental health. It provides access to information, social connections, and remote mental health services. Social media and online platforms help reduce loneliness and connect people with mental health professionals. However, ICT can also lead to addiction, stress, low self-esteem, and cyberbullying, especially among young users. Gamification and Virtual Reality can help with mental health treatment, but they also have risks, like unrealistic expectations or worsening symptoms. Using ICT wisely and promoting digital literacy can help balance its benefits and challenges.

The study examines how ICT is incorporated into teaching and learning, stressing its value in education for the twenty-first century while also pointing out that teachers have limited ICT proficiency and use [4]. Through an analysis of instructor feedback, the study investigates the effects of ICT use in Greek higher education institutions during the COVID-19 lockdown [5]. Inadequate technical support, insufficient infrastructure, and issues with remote instruction and testing are major obstacles. The limitations of manual clinical diagnosis are addressed in the study by proposing an automated system for detecting stress levels using speech samples [6]. Based on answers to questions approved by psychiatrists, the study examines the acoustic properties of 800 voice recordings of 100 participants at LNIPE, Guwahati, India. The study examines how technostress affects higher education, highlighting the difficulties even though integrating technology has many advantages [7]. A systematic review of 83 studies from 1,861 publications and visualization analysis are used in this study to determine the causes and consequences of student technostress. People with dyslexia can benefit from ICT tools by using them to overcome learning obstacles and enhance their performance, especially in the classroom, as has been discussed by Vouglanis and Driga [8]. They are becoming more and more effective, according to research. Further research is required to improve the effectiveness of the dyslexia detection software that has been developed, including for Greek students. The stress, burnout, resilience, and technostress of 168 teachers in Murcia, Spain, are all examined by Pagán Garbín et al. [9]. It looks at how these variables relate to one another as well as how gender and age affect things. The study discovered that while resilience has a positive correlation with personal achievement, it has a negative correlation with stress and technostress. Resilience levels were predicted by an artificial neural network (ANN)

*Corresponding author: Saiyed Umer, Department of Computer Science and Engineering, Aliah University, India. Email: saiyed.umer@aliah.ac.in

with an accuracy of 86.7%, with personal accomplishment being the most important predictor.

The relationship between language and mental health can be contextualized using the Cultural Adaptation Stress Theory, which highlights the psychological strain faced by individuals adapting to new linguistic or cultural environments [10]. Similarly, the Cognitive Load Theory suggests that students in non-native language settings may face additional processing burdens, affecting their stress levels and academic engagement. This study is driven by the hypothesis that the medium of instruction whether native or non-native can significantly influence a student's psychological well-being. Language may serve as both a cognitive and a cultural barrier, increasing mental load and contributing to stress [11]. Non-native learners often expend additional cognitive effort decoding the language before engaging with academic content, resulting in a dual burden that may lead to elevated stress levels among English-medium students (EMS) compared to their Bengali-medium (BMS) counterparts. According to the Cultural Adaptation Stress Theory, students in non-native language environments often experience disorientation, social withdrawal, and anxiety. When this is combined with high academic expectations and ICT-driven content, the perceived stress may escalate. Therefore, this research focuses on comparing English-medium (non-native) and Bengali-medium (native) students to investigate the relationship between language background and mental health outcomes.

Balsalobre-Lorente et al. [12] examine how tourism, urbanization, and natural resource rents influence environmental sustainability, emphasizing the critical role of Artificial Intelligence (AI) and ICT in driving progress toward the Sustainable Development Goals (SDGs). Their study highlights that digital technologies serve as key enablers for balancing economic growth with ecological preservation in the digital era. In Olatunde-Aiyedun [13] work, the determinants of artificial intelligence literacy were explored, with emphasis placed on the roles of the digital divide, computational thinking, and cognitive absorption. It was highlighted that these factors significantly shape individuals' capacity to develop AI literacy. In Celik [14] study, the integration of artificial intelligence into science education curricula in Nigerian universities was examined. Emphasis was placed on how AI adoption in education was approached to enhance teaching and learning outcomes. The study by Bibri et al. [15] was carried out to provide an extensive literature review on environmentally sustainable smart cities. Within this work, the convergence of AI, Internet of Things (IoT), and big data technologies was emphasized as playing a crucial role in enabling integrated solutions for sustainable urban futures. The rise of artificial intelligence across European regions, focusing on how a strong local ICT base fosters AI development has been investigated by Xiao and Boschma [16]. Their findings reveal that regions with established ICT capabilities are better positioned to drive innovation and capture the economic benefits of AI.

The rapid integration of ICT into education has transformed teaching and learning practices worldwide. While ICT offers benefits such as increased access to learning resources and remote collaboration, it can also contribute to stress, anxiety, and reduced social interaction among students. Current research on the link between ICT-based education and mental health is limited in scope, often lacking predictive models that can identify at-risk students early. This study addresses this gap by developing an AI-driven framework that predicts students' mental health outcomes, specifically stress, anxiety, and depression based on psychometric data. The aim of this research work is to provide schools and policymakers with a tool that supports timely interventions and promotes students' well-being in ICT-enabled learning environments.

The objective of this work is to predict the mental health of students in terms of whether they are stressed or unstressed by asking the questions of the questionnaire. Hence, based on these studies, this research makes the following contributions:

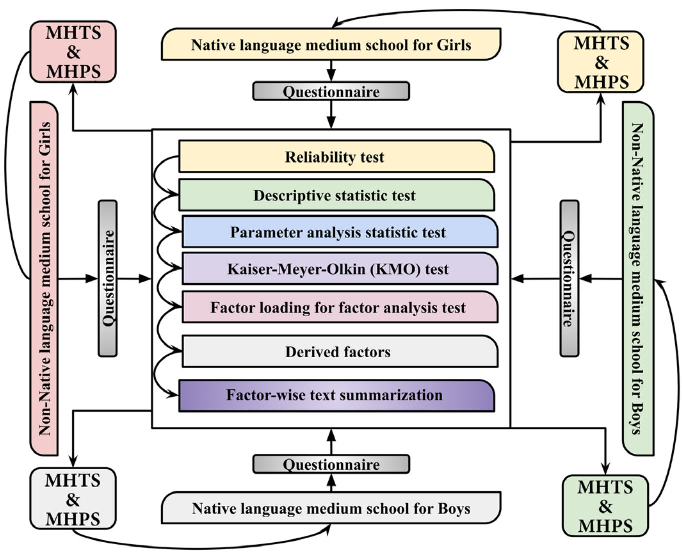
1) This research thoroughly investigates the integration of ICT-based education systems to predict children's mental well-being through learning analytics.
2) This research analyzes students from different language mediums—native (regional) and non-native (international)—as well as gender (boys and girls) to explore differences in mental health status across these groups.
3) This research develops prototypes for categorizing overall mental health conditions by utilizing statistical data analysis techniques and textual summarization of student responses to the questionnaire.
4) This research implements a gender classification system for real-time test analytics, effectively handling questionnaire responses with unknown label data.

The organization of this work is as follows: Section 2 describes the materials and methods used for this research. The results and discussions have been performed in Section 3. The work is concluded in Section 4.

2. Materials and Methods

This section provides the detailed explanation of materials in terms of data collection procedure and the sequence of methods employed for performing the ensemble of statistical and machine learning based data analysis techniques to build the prediction models for deriving the mental health status in the ICT-Based School Children's Education System [17]. The analysis comprises three stages: (i) Developing an architectural framework for implementing the prototypes for categorizing overall mental health conditions by utilizing statistical data analysis techniques and textual summarization of student responses to the questionnaire (demonstrated in Figure 1); (ii) Analyzing students from different language mediums native (regional) and non-native (international) as well as gender (boys and girls) to explore differences in mental health status across these groups (demonstrated in Figure 2); (iii) Implementing a gender classification system using machine learning-based data analysis techniques for the real-time test analytics, effectively handling questionnaire responses with unknown label

Figure 1
An architectural framework for developing prototypes to categorize overall mental health conditions in native and non-native language medium schools



data (demonstrated in Figure 3); lastly (iv) Validating the considered hypotheses by providing appropriate justifications and explaining the significance of the findings in the context of mental health prediction in ICT-based children's education [18].

Figure 2
A workflow diagram for developing the Mental Health Text Summarization (MHTS) model for children with known gender information in the test samples

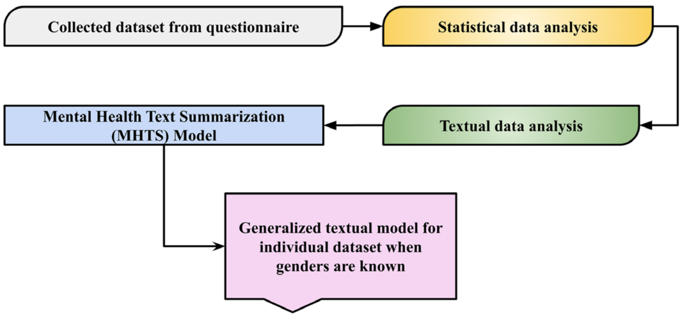
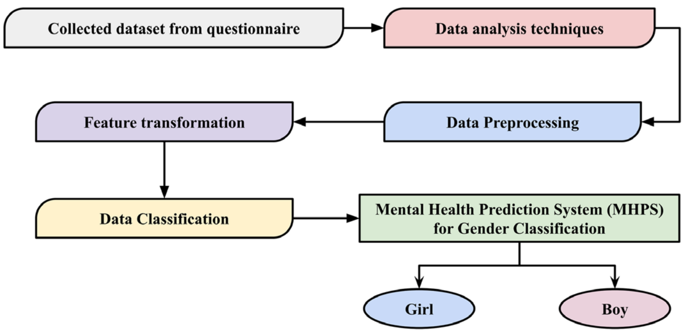


Figure 3
A workflow diagram for developing a prediction model for gender classification by analyzing questionnaire responses from test samples without gender information



2.1. Data collection

In this work, the first step of the proposed system is the data collection. A questionnaire, which is the 'Perceived Stress Questionnaire (PSQ)', has been considered here for support by this work [19]. This questionnaire is a widely used psychological instrument for measuring the perception of stress. It comprises thirty questions (features) prompting individuals to reflect on their feelings and thoughts over the past month. The children will rate each question of this questionnaire on a scale from 0 to 4, with zero being "never" and four being "very often." After gathering the responses to all questions, an overall score is generated, which is the indicator for measuring the mental health score for the corresponding children. This generated score is the PSQ index, which is calculated as $r = \frac{\sum_{i=0}^{30}(r_i - 30)}{90}$. The value of r lies in the interval [0, 1]. This data collection is conducted among the 12$^{\text{th}}$-grade students (typically around 17–18 years old) of English (Non-native) and Bengali (Native) language schools in West Bengal, India. To minimize potential biases such as social desirability or expectation effects, data are collected anonymously, and no school staff is present in front of the students during the survey. Students are also informed that their responses would have no academic consequences. However, we acknowledge that residual bias may exist, especially related to self-perception and peer conformity. These limitations are inherent in self-reported mental health assessments.

A total of 100 girls and 100 boys from each school medium participated in the data collection process. Hence, data from 400 students have been collected for this study, preserving ethical guidelines to protect their rights, privacy, and well-being. In addition to parental consent, permission from the students has also been sought to ensure they understand the process and agree to participate voluntarily. In addition, we obtained permission from the head of the institution for data collection, which is conducted in a supervised environment. The list of questions of the PSQ is demonstrated in Table 1.

Table 1
List of items in the PSQ

S. no.	Statement	S. no.	Statement	S. no.	Statement
S1	You feel rested	S11	You have too many decisions to make	S21	You enjoy yourself
S2	You feel that too many demands are being made on you	S12	You feel frustrated	S22	You are afraid for the future
S3	You are irritable or grouchy	S13	You are full of energy	S23	You feel you're doing things because you have to, not because you want to
S4	You have too many things to do	S14	You feel tense	S24	You feel criticized or judged
S5	You feel lonely or isolated	S15	Your problems seem to be piling up	S25	You are lighthearted
S6	You find yourself in situations of conflict	S16	You feel you're in a hurry	S26	You feelmentally exhausted
S7	You feel you're doing things you really like	S17	You feel safe and protected	S27	You have trouble relaxing
S8	You feel tired	S18	You have many worries	S28	You feel loaded down with responsibility
S9	You fear you may not manage to attain your goals	S19	You are under pressure from other people	S29	You have enough time for yourself
S10	You feel calm	S20	You feel discouraged	S30	You feel under pressure from deadlines

2.2. Data analysis

After data collection, the next step is data analysis. Here, the first step of data analysis is a data reliability test.

2.2.1. Data reliability test

A Cronbach's alpha is performed for this purpose [20]. This measure ranges from 0 to 1 and can be interpreted as follows: Strong correlations between the test features are indicated by values nearer 1, which indicate improved internal consistency.

2.2.2. Descriptive statistics test

In this analysis, the mean represents the average value of a dataset and is calculated by summing up all the values and dividing by the total number of observations, and the variance measures the spread or dispersion of the data points around the mean [21]. It is calculated by taking the average squared differences between each data point and the mean. Mathematically, it is described as

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad \sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}$$

where, μ is the mean, σ^2 is the variance, x_i represents each individual's value in the dataset, and n denotes the total number of observations. Higher variance signifies increased diversity among the dataset's data points, whereas lower variance implies data points are more closely clustered around the mean.

2.2.3. Parameter statistics test

This statistical test assesses sampling adequacy when performing parameter (factor) analysis on a dataset [22]. It establishes whether the variables in the dataset are suitable for factor analysis, which is a technique for identifying underlying dimensions or factors that account for correlations between variables. Here, the Mann-Whitney U test has been performed for the parameter statistics test [23]. This test starts with two sub-tests, which are the Kaiser-Meyer-Olkin (KMO) and Bartlett's tests. The KMO test yields a value ranging from 0 to 1, with closer to 1 suggesting a high correlation among variables, indicating suitability for factor analysis. Bartlett's test supports and leads to factor analysis tests of structural equation modeling (SEM) techniques [24].

2.2.4. Factor analysis test

This statistical method is employed to determine the underlying factors or dimensions that account for the correlations between observed variables [23]. The factor loadings, which show the strength and direction of relationships between the underlying factors and the observed variables, are examined here. It is considered that variables with high loadings on a particular factor are closely related to that factor.

2.2.5. Derived factors

A theoretical model of factor analysis is specified in advance, and its fit to the observed data is tested using confirmatory factor analysis (CFA) [24]. The CFA and exploratory factor analysis (EFA) are integrated into SEM. All things considered, factor analysis is a reliable technique for revealing obscure patterns in data, clarifying underlying concepts, and intricate datasets for additional examination.

2.2.6. Exploratory factor analysis (EFA)

It is employed to uncover latent stress dimensions across the EMS/BMS groups. Here, EFA is done using CFA and SEM techniques. The CFA is used here to validate the consistency of these factors within the sample. SEM is used to test the theoretical structure of stress constructs derived from PSQ. SEM is justified here as it allows simultaneous assessment of observed variables and latent constructs, supporting the multidimensional modeling of student stress.

2.2.7. Machine learning-based prediction model test

A machine learning technique-based test analysis has been performed for factor-wise text summarization purposes [25]. The derived factors lead to a hierarchical structure with the association of each questionnaire item to its derived factor. So, the items represented by the sentence question of the employed questionnaire undergo text summarization techniques. This text summarization technique aims to obtain the particular subject, verb, and object of the question for which the students got stressed. This work aims to predict students' mental health in terms of whether they are stressed or non-stressed by asking the questionnaire questions. So, the text summarization is essential to make a ground model for them by justifying their mental health status among girls and boys of native and non-native language medium schools [26].Text summarization was applied to identify the dominant stress themes within each factor derived from factor analysis. Using NLP preprocessing (stop-word removal, tokenization), each question was distilled to its core semantics (Subject-Verb-Object triads). These summaries served as prototype feature labels for the prediction model, enabling pattern mapping between input responses and stress profiles. Then, each student's questionnaire data is labeled using the factor structure and summarization outputs. These outputs are then used to train machine learning classifiers to predict mental health status (stressed vs. non-stressed) and gender, with evaluation metrics computed on a test split. The algorithmic steps of text summarization of the PSQ sentences have been presented in Algorithm 1.

Algorithm 1: For Text Summarization

Step-1: Initialization of Stop Words: Define a stop-word set W = {W_1, W_2, …, W_n} containing common, non-informative words to be filtered from all statements.

Step-2: Preprocessing Function φ(s): For a given statement s ϵ S_i, define a transformation function φ: s⟶s` such that: φ(s) = Join ({w ϵ Tokenize(s)\W}), Where: (i) Tokenize(s) splits s into words, (ii) \ removes elements in W, (iii) Join () reconstructs the text without stop words.

Step-3: Summarization for Each Factor: For every factor F_i, apply φ to each s ϵ S_i, producing a cleaned set $S_{\cdot i}$. Then compute the combined summary C_i as: C_i = φ($s1_i$); φ($s2_i$); ... ; φ(sn_i).

Step-4: Tabular Representation: Construct a two-column matrix (or DataFrame): D = [(F_1,C_1), (F_2,C_2), …, (F_n,C_n)].

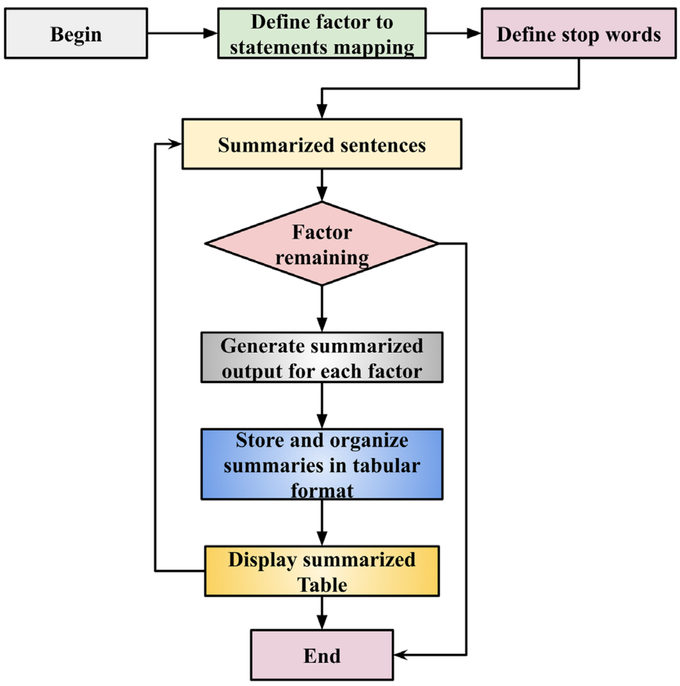
Step-5: Sorting by Factor Index: Extract numerical indices from Fi using: Index(F_i) = i, where F_i = "Factor-" + i. Sort D in ascending order based on Index(F_i).

Step-6: Final Output: Present the matrix D as the final, ordered summary of factors and their associated processed statements.

From this algorithm, it has been observed that factor-wise text summarization involves removing stop words, extracting key tokens (subject, verb, object), and combining statements into concise labels to generate summarized features for modeling. Figure 4 shows the work flow diagram of Algorithm 1. Here, for the predictive modeling task, the dataset is randomly split into training (50%) and testing (50%) sets. The features are then extracted from the PSQ responses and factor summarizations (Algorithm 1). A well-known Support Vector Machine (SVM) classifier has been employed to build the prediction model. Since questionnaire responses are often collected using Likert scales, resulting in a large number of numerical features across multiple items, SVMs are particularly effective because they handle high-dimensional feature spaces well—even when the number of features exceeds the number of samples—making them suitable for modeling complex response patterns and achieving reliable classification or prediction outcomes. The performance of the prediction model has been shown in terms of accuracy, precision, recall, and F1-score metric. The best-performing model achieved will be used as the prediction models of the proposed study.

2.3. Hypothesis testing

Figure 4
A workflow diagram for Algorithm 1 for building the prediction model for mental health frameworks



This study examines two hypotheses: the Alternative Hypothesis (H1), which contradicts the Null Hypothesis (H_0), which represents the assumption under examination [27]. The null hypothesis is rejected when a statistical test's p-value is less than a selected significance level (such as 0.05). Based on the collected dataset, the data analysis has been verified using these hypotheses to predict children's mental health outcomes. Since in this study the students from the Naive (Bengali) Medium Language (BML) and Non-native (English) Medium Language (EML) schools are considered, the hypotheses under this study are explained in Table 2.

3. Experiments

This section explains the experimentation carried out in the study by identifying the findings and providing a rationale based on the proposed experimentation. For this study, 400 students took part in the data collection process. After that, the gathered dataset is put through the data reliability test, which is shown in Table 3. From this table it has been observed that the value of Cronbach's Alpha is consistently greater than 0.8, which means there exists a good association of the questionnaire with the collected data samples.

During data analysis, the collected dataset is categorized into four subsets based on school medium and gender: (i) EMS_Girls (English-medium school girls), (ii) EMS_Boys (English-medium school boys), (iii) BMS_Girls (Bengali-medium school girls), and (iv) BMS_Boys (Bengali-medium school boys). The next data analysis test is the descriptive statistics test, where the mean and variance of the collected data correspond to each dataset, have been discussed. The PSQ has been validated across various cultures and was adapted in this study for Indian students with minor linguistic adjustments for clarity. A pilot test with 30 students is conducted to confirm the cultural and linguistic suitability of the instrument. Cronbach's alpha values exceed 0.82 across all subgroups (see Table 3), indicating high reliability. Moreover, Bartlett's test of sphericity ($p < 0.001$) and acceptable KMO values (0.68–0.74) support both construct validity and sampling adequacy, confirming that the dataset is appropriate for factor analysis and useful for this study.

Table 3
Results of the data reliability test

No. of samples	Cronbach's alpha	Interpretation
30	0.8215	Good
100	0.8221	Good
200	0.8315	Good
300	0.8562	Good
400	0.8762	Good

3.1. Descriptive statistics test

This test (discussed at Section 2.2 (b)) has been performed with the collected dataset and the results are reported in Figure 5. This figure demonstrates that stress levels of girls are almost identical in both EMS and BMS conditions, indicating that the two conditions have little effect on them. The fact that stress levels of boys are marginally higher in BMS than in EMS suggests that they may feel a little more stressed in BMS. Overall, compared to boys, girls experience slightly higher levels of stress. Figure 6 shows that under both EMS and BMS conditions,

Table 2
List of hypotheses considered in this study

Hypothesis	Null hypothesis (H_0)	Alternative hypothesis (H_1)
Hypothesis 1	Girls experience similar perceived stress levels in EMS and BMS.	Girls experience dissimilar perceived stress levels in EMS and BMS.
Hypothesis 2	Boys experience similar perceived stress levels in EMS and BMS.	Boys experience dissimilar perceived stress levels in EMS and BMS.
Hypothesis 3	Girls experience similar perceived stress levels compared to boys.	Girls experience dissimilar perceived stress levels compared to boys.
Hypothesis 4	Students from EMS experience similar perceived stress levels compared to students from BMS.	Students from EMS experience higher perceived stress levels compared to students from BMS.

girls exhibit nearly no change in stress variability. Compared to BMS_Boys, EMS_Boys variance is higher, indicating that their stress levels fluctuate more under EMS. Boys exhibit a slightly higher overall variance in stress than girls, which means that their stress levels vary more.

Figure 5
Comparison of the mean PSQ scores for boys and girls in EMS and BMS

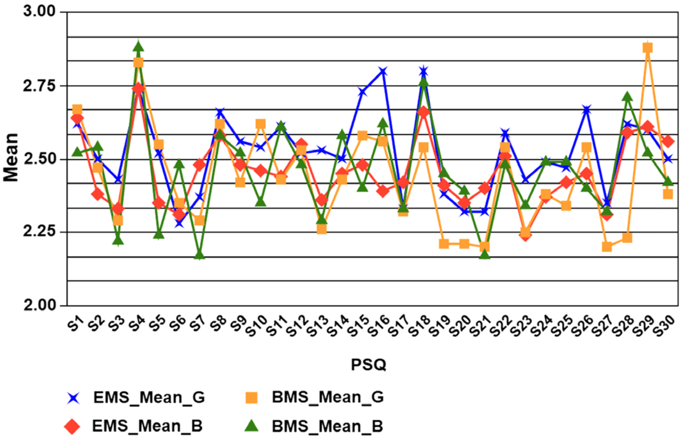
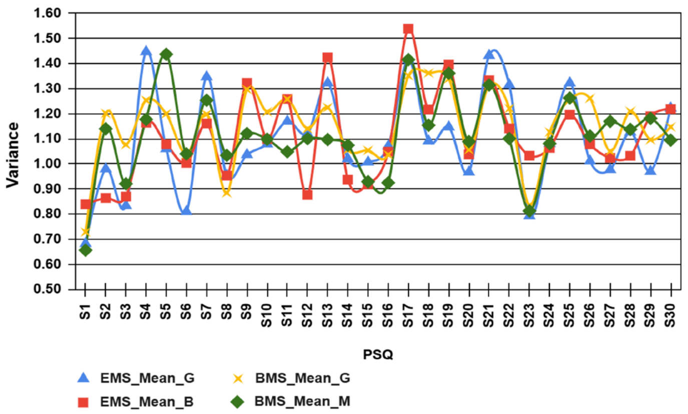


Figure 6
Comparison of PSQ score variance for boys and girls in EMS and BMS



3.2. Parameter statistics test

The results of the Mann-Whitney U test (discussed in Section 2.2 (c)) show in Table 4 that the stress levels in both EMS and BMS schools varied significantly between boys and girls. According to the impact size (r), the degree of these differences varies depending on the survey item. A significant effect size is seen for a number of categories, including S3, S5, S15, and S26, suggesting a significant difference in the stress levels of boys and girls. A moderate difference is shown by the medium effect size of other items, such as S6, S17, and S19. Furthermore, differences in effect sizes between EMS and BMS imply that stress levels and gender-specific variances may be more noticeable in one school than the other. This suggests that environmental or institutional variables may affect stress levels of students, which calls for more research into the root causes. Hence, based on impact sizes and U-statistics, the stress level ranking with BMS Boys < BMS Girls < EMS Boys < EMS Girls with is therefore largely accurate. Accordingly, with Higher effect sizes and lower U-values in the Mann-Whitney U test indicate that students in English Medium Schools (EMS) report higher levels of stress than those in Bengali Medium Schools (BMS with). This score results from girls experiencing more significant stress than boys inside each school. Mann-Whitney U results do not directly justify factor analysis, so KMO/Bartlett tests are required, as reported in Table 5.

To assess whether sampling is adequate for factor analysis or

Table 4
Non-parametric test analysis using Mann-Whitney U test for ordinal or non-normally distributed dataset

Question	EMS_U_Statistic	EMS_p_Value	EMS_Effect_Size_r	EMS_Effect_Interpretation	BMS_U_Statis-tic	BMS_p_Value	BMS_Effect_Size_r	BMS_Effect_Interpretation
S1	4874.00	0.74	0.49	Medium	5405.00	0.28	0.54	Large
S2	5372.00	0.34	0.54	Large	4816.00	0.64	0.48	Medium
S3	5275.50	0.48	0.53	Large	5169.50	0.67	0.52	Large
S4	5034.00	0.93	0.50	Large	4897.50	0.79	0.49	Medium
S5	5456.00	0.24	0.55	Large	5793.00	0.05	0.58	Large
S6	4937.00	0.87	0.49	Medium	4650.50	0.37	0.47	Medium
S7	4717.50	0.48	0.47	Medium	5320.00	0.42	0.53	Large
S8	5236.00	0.55	0.52	Large	5120.00	0.76	0.51	Large
S9	5226.00	0.57	0.52	Large	4728.00	0.49	0.47	Medium
S10	5210.00	0.59	0.52	Large	5712.00	0.07	0.57	Large
S11	5433.00	0.27	0.54	Large	4545.00	0.25	0.45	Medium
S12	4903.50	0.81	0.49	Medium	5123.00	0.76	0.51	Large
S13	5402.50	0.31	0.54	Large	4893.50	0.79	0.49	Medium
S14	5135.00	0.73	0.51	Large	4616.50	0.33	0.46	Medium
S15	5699.50	0.08	0.57	Large	5480.00	0.22	0.55	Large
S16	6072.50	0.01	0.61	Large	4847.00	0.70	0.48	Medium
S17	4804.00	0.62	0.48	Medium	4987.00	0.97	0.50	Medium
S18	5356.00	0.37	0.54	Large	4476.00	0.19	0.45	Medium
S19	4955.50	0.91	0.50	Medium	4414.50	0.14	0.44	Medium
S20	4944.50	0.89	0.49	Medium	4531.00	0.23	0.45	Medium
S21	4770.00	0.56	0.48	Medium	5057.50	0.88	0.51	Large
S22	5192.50	0.63	0.52	Large	5160.00	0.69	0.52	Large
S23	5597.00	0.13	0.56	Large	4704.50	0.45	0.47	Medium
S24	5345.00	0.38	0.53	Large	4719.00	0.48	0.47	Medium
S25	5117.50	0.77	0.51	Large	4623.50	0.34	0.46	Medium
S26	5623.00	0.11	0.56	Large	5350.00	0.38	0.54	Large
S27	5121.50	0.76	0.51	Large	4703.00	0.45	0.47	Medium
S28	5086.00	0.83	0.51	Large	3777.50	0.00	0.38	Medium
S29	4942.00	0.88	0.49	Medium	5933.00	0.02	0.59	Large
S30	4848.00	0.70	0.48	Medium	4901.00	0.80	0.49	Medium

Table 5
Kaiser-Meyer-Olkin (KMO) & Bartlett chi-square test analysis

Dataset	KMO overall	Bartlett chi-square	Bartlett p-value	Remark
EMS_Girls	0.72	910.54	0.00	Good
EMS_Boys	0.69	936.22	0.00	Moderate
BMS_Girls	0.68	921.69	0.00	Moderate
BMS_Boys	0.74	1083.88	0.00	Good

principal component analysis (PCA), the KMO Test is used, and results are reported in Table 5. Analyzing the strength of the correlations between the variables establishes whether the dataset is appropriate for factor extraction. This table shows the experiments for KMO and Barlett tests and reveals that the statistical validity of factor analysis is confirmed by Bartlett's Test. The KMO values support factor analysis because they fall within an acceptable range. KMO values (0.68–0.74) indicate sampling adequacy for factor analysis, while Bartlett's test ($p < 0.001$) confirms that the correlation matrix is not an identity matrix. Thus, factor analysis is statistically justified for all subgroups.

3.3. Factor analysis test

The demonstration of the factor analysis (discussed at Section 2.2 (d)) is shown in Figure 7. This shows that three to five principal components appear relevant for each group according to the Kaiser criterion and the elbow approach. The precise number of retained components could be verified by the SEM techniques applied to each dataset to find the factors (groups) and parameters inside each factor, such as a factor loading investigation. Figure 7 shows that EMS students display a steeper elbow in the scree plot, indicating a denser structure of stress-related factors, suggesting deeper psychological strain in this group. Screen plots for factor analysis of PSQ responses across different student groups. The 'elbow point' indicates the number of meaningful latent factors for each group, with EMS cohorts showing more components suggesting more complex stress dimensions compared to BMS groups.

3.4. Factor loading test

The obtained factors with association of items of the questionnaire corresponding to each employed dataset have been demonstrated in Figure 8. Using EFA (discussed in Section 2.2 (e)), 11 unique factors are identified for the EMS_Girls dataset (Figures 8 and 9 (a)). These factors group items that show high factor loadings, indicating strong conceptual and statistical coherence. Emotional overload and responsibility are represented by statements S2, S23, S27, and S28 in Factor 1, whereas positive affect and well-being are represented by statements S10, S13, S17, and S21 in Factor 2. Factor 3 uses items S3, S12, and S19 to capture emotional agitation, and Factor 4 uses items S15 and S30 to reflect task-related stress. Even though each of the factors (S4, S26, S22, S5, S6, S1, and S7) only has one statement, they each represent distinct psychological aspects like task overload, mental exhaustion, future anxiety, and intrinsic motivation. Their strong loadings and unique thematic contributions support their functions as separate latent constructs in the model.

Through EFA (Figures 8 and 9 (b)), ten unique factors have been identified for the EMS_Boys dataset; each grouping represented conceptually and statistically coherent items. A central theme of emotional strain and stress is indicated by the clustering of statements S7, S13, S17, S21, and S29 in Factor 1. S6, S9, S20, and S22 are among the items in Factors 2 through 4, which represent aspects of task-related stress and outside pressure. Independent psychological constructs like positive affect, fatigue, and social pressure are represented by factors like Factor-5 to Factor-10, which include S14, S16, S24, and S25. Each of these factors has a high factor loading, guaranteeing its statistical

Figure 7
Demonstration of the factor analysis comparison for Girls & Boys of EMS and BMS datasets

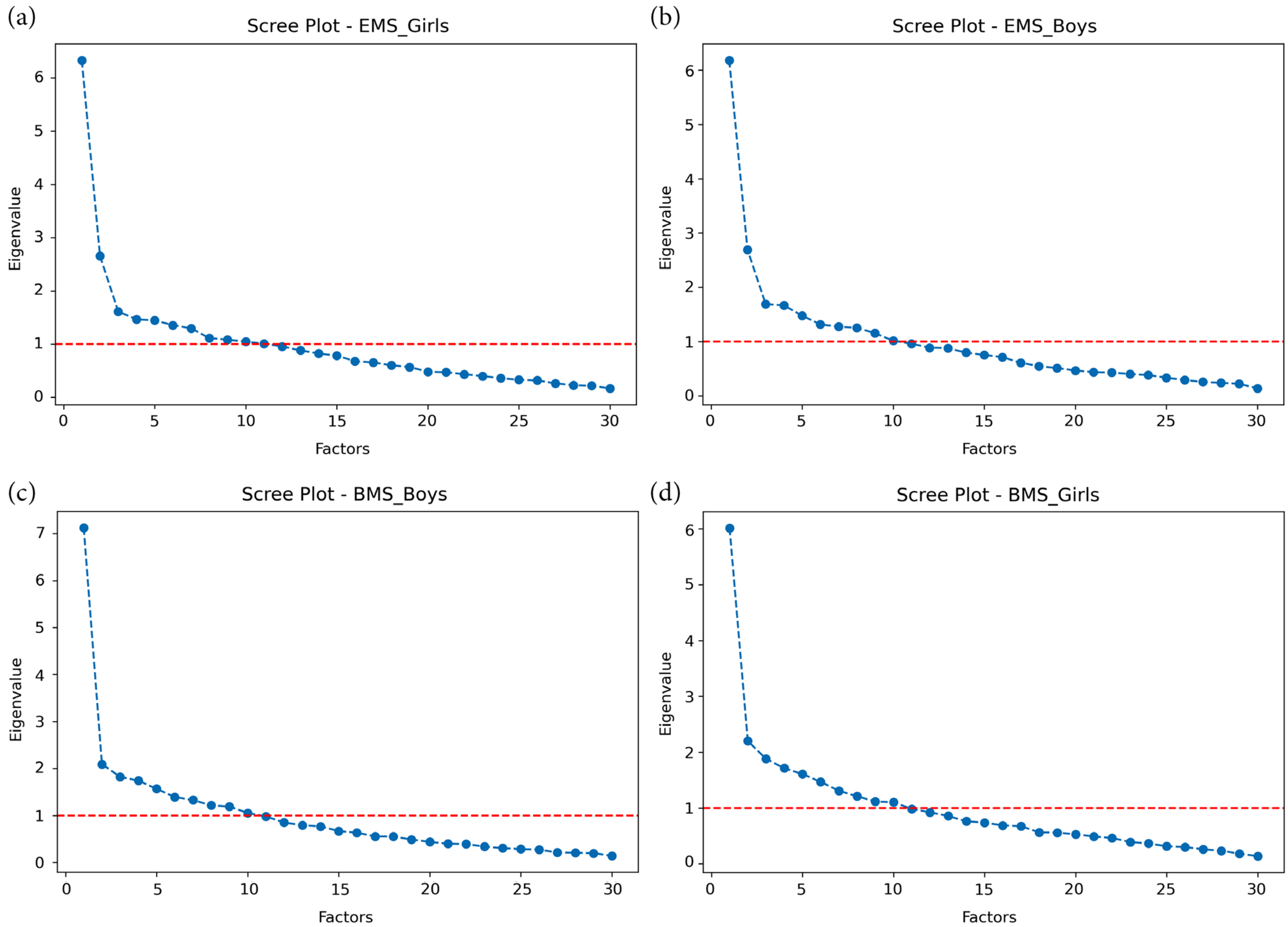


Figure 8
Derived factors containing corresponding statements for Girls & Boys of EMS and BMS datasets: (a) EMS_Girls, (b) EMS_Boys, (c) BMS_Girls, and (d) BMS_Boys

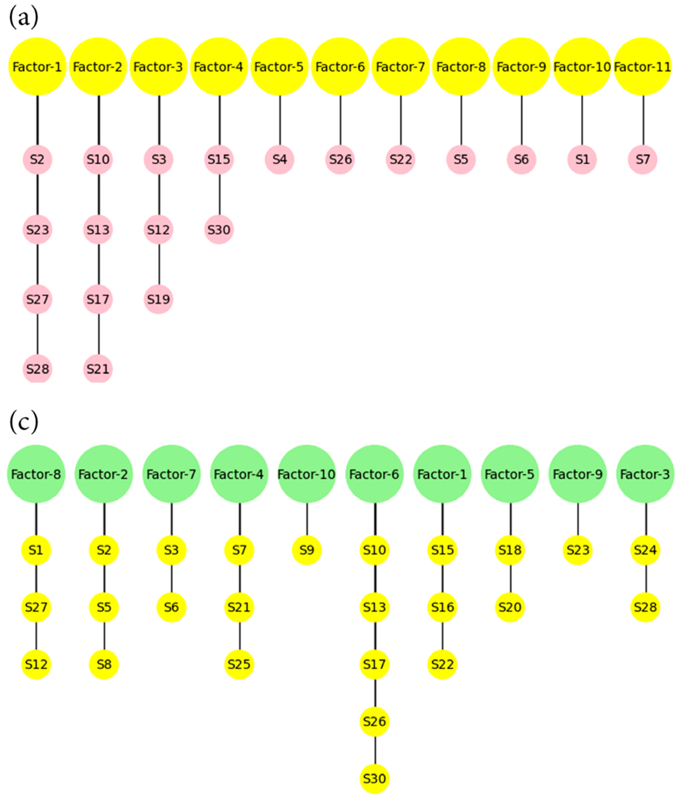
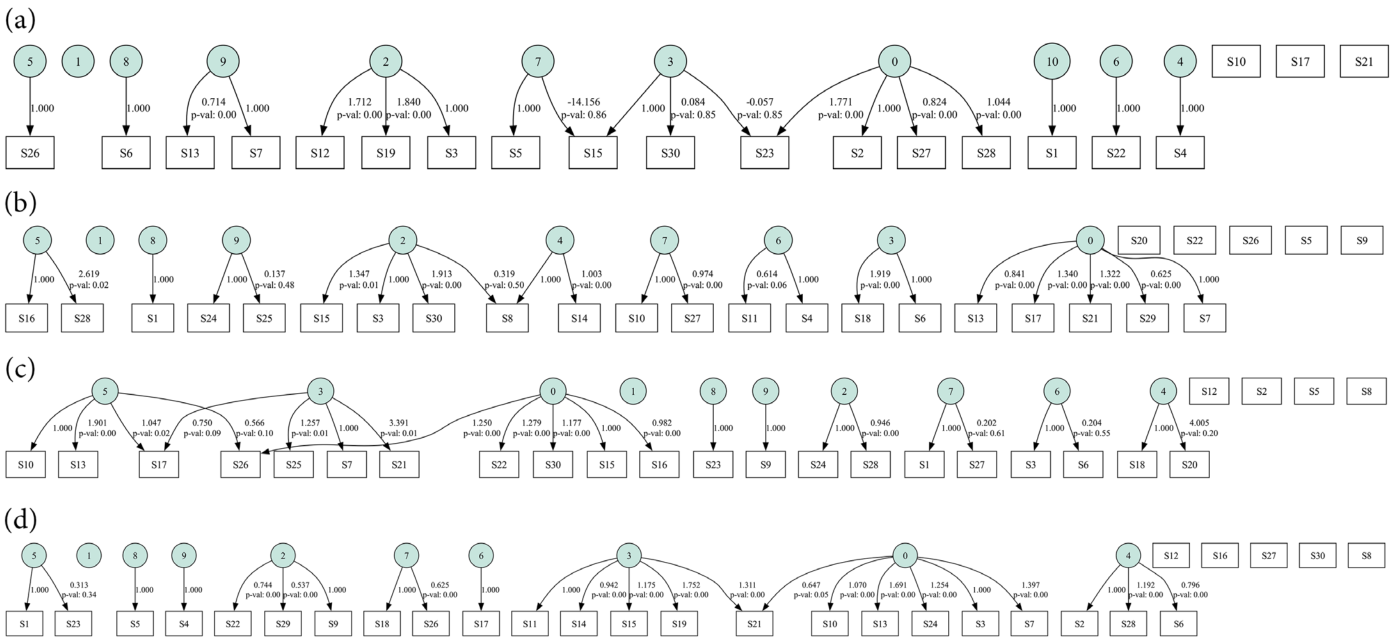


Figure 9
Confirmatory factor analysis (CFA) derived from Figure 7 factor analysis for Girls & Boys of EMS and BMS datasets: (a) EMS_Girls, (b) EMS_Boys, (c) BMS_Girls, and (d) BMS_Boys



significance. These factors are appropriate for additional psychological evaluation and interpretation because of their structure, which supports their thematic distinction and dependability.

For the BMS_Girls dataset, 10 distinct factors are extracted using EFA (Figures 8 and 9 (c)), ensuring that each grouped item exhibits strong conceptual and statistical coherence. Factor-1 consists of S10, S13, S17, S26, and S30, indicating a dimension associated with emotional regulation and stress response. Factor-2 to Factor-6 include items such as S2, S5, S7, S9, and S15, reflecting aspects of task pressure, personal well-being, and external challenges. Meanwhile, Factor-7 to Factor-10, containing statements like S23, S24, and S28, represent unique psychological constructs such as social stress, resilience, and mental fatigue, which demonstrate high factor loadings and contribute meaningfully to the dataset's structure. Since each component successfully captures unique psychological dimensions due to the strong thematic separation among factors, this categorization is appropriate for additional interpretation and analysis. EFA (Figures 8 and 9 (d)) was used to identify 10 unique factors for the BMS_Boys dataset, each of which grouped statements that were conceptually and statistically related. S3, S7, S10, S13, and S24 are included in Factor-1, suggesting a strong correlation with task pressure and emotional control. Stress reactions, cognitive load, and interpersonal difficulties are reflected in Factors 2 through 6, which include statements like S8, S9, S11, S16, and S22. In contrast, different psychological dimensions like mental exhaustion, resilience, and emotional well-being are represented by items like S17, S18, and S26 in Factors 7 through 10. High factor loadings and the distinct separation of these factors guarantee a theoretically sound and statistically valid classification, offering a strong basis for additional research.

3.5. Hypothesis testing

The hypotheses which are considered (see Table 2) in this study have been tested and the results are reported with justification in Table 6.

Table 6
Experiments for the hypothesis testings considered under this study

Hypothesis	Null hypothesis (H_0)	Alternative hypothesis (H_1)
Hypothesis 1	Girls experience similar perceived stress levels in EMS and BMS.	Girls experience dissimilar perceived stress levels in EMS and BMS.
	Findings: Based on the experiments of (i) Descriptive statistics test, it has been observed that girls' stress levels are almost identical in both EMS and BMS conditions, hence, accepting the null-hypothesis, while rejecting the alternative hypothesis.	
Hypothesis 2	Boys experience similar perceived stress levels in EMS and BMS.	Boys experience dissimilar perceived stress levels in EMS and BMS.
	Findings: Based on the experiments of (i) Descriptive statistics test, it has been observed that g boys' stress levels are marginally higher in BMS than in EMS. Hence, rejecting the null-hypothesis, while accepting the alternative-hypothesis.	
Hypothesis 3	Girls experience similar perceived stress levels compared to boys.	Girls experience dissimilar perceived stress levels compared to boys.
	Findings: Based on the experiments of (ii) Parametric test, it has been observed that stress levels and gender-specific variances may be more noticeable in one school than the other. Hence, based on this test, it has been obtained that the stress level ranking can be arranged as "BMS Boys < BMS Girls < EMS Boys < EMS Girls". Hence, rejecting the null-hypothesis, while accepting the alternative hypothesis.	
Hypothesis 4	Students from EMS experience similar perceived stress levels compared to students from BMS.	Students from EMS experience higher perceived stress levels compared to students from BMS.
	Findings: Based on the experiments of (iv) factor analysis and factor loadings, it has been observed that the visual comparison of the factor clusters across EMS and BMS groups reveals that EMS students exhibit a more complex and concentrated factor structure, particularly in the EMS_Girls and EMS_Boys datasets. For instance, the EMS_Girls group demonstrates 11 distinct factors with several high-loading stress-indicative items (e.g., S2, S23, S27, and S28 grouped under emotional overload), whereas BMS_Girls show a relatively distributed clustering with fewer items per factor, suggesting lower cumulative stress impact. Similarly, the EMS_Boys group shows more concentrated clusters like Factor-1 containing deeply stress-related items such as S7, S13, S17, S21, and S29. In contrast, BMS_Boys show a more dispersed structure with items like S3, S10, S13, and S24 spread across more loosely defined factors. This denser clustering of high-stress items within fewer dominant factors in EMS students indicates stronger internal coherence of stress dimensions, supporting the assertion that EMS students face higher perceived stress levels than BMS students. Thus, Hypothesis 4 is supported, as the factor structures reveal both a greater number and depth of stress-related constructs in the EMS cohorts compared to their BMS counterparts. Hence, rejecting the null-hypothesis, while accepting the alternative hypothesis.	

3.6. Prediction of models for Mental Health Status (MHS)

Now, utilizing NLP-based text summarization techniques (discussed in Algorithm 1), each factor is represented by interpreting its corresponding statements from the PSQ, and then the machine learning (ML)-based prediction model has been built. In the four datasets, each derived factor serves as a prototype for mental health situations, prioritizing the identified factors from highest to lowest. Table 7 presents a factor-wise text summarization for the EMS-Girls dataset (see Figure 8(a)), ranking factors from highest to lowest priority. Each factor is derived from grouped statements, capturing key mental health themes such as stress, relaxation, social pressure, exhaustion, and isolation. This summarization provides insights into the psychological state of individuals based on the prioritized factors from the PSQ. The derived prototype will be used as a prediction model shown in Figure 10(a).

Table 7
Factor-wise text-summarization and prediction model for EMS-Girls dataset, where the factor precedence is from top to down

Factor	Statements included	Combined summary
Factor-1	S2, S23, S27, S28	Demands; things; trouble relaxing; responsibility
Factor-2	S10, S13, S17, S21	Calm; energy; safe; enjoy
Factor-3	S3, S12, S19	Grouchy; frustrated; pressure people
Factor-4	S15, S30	Problems; pressure deadlines
Factor-5	S4	Things
Factor-6	S26	Mentally exhausted
Factor-7	S22	Future
Factor-8	S5	Isolated
Factor-9	S6	Situations conflict
Factor-10	S1	Feel rested
Factor-11	S7	Things

Table 8 presents a factor-wise text summarization for the EMS-Boys dataset (see Figure 8(b)), prioritizing mental health themes from highest to lowest. Key concerns include energy levels, safety, time management, isolation, discouragement, stress from deadlines, conflicts, and worries, along with tension, responsibility, and mental exhaustion. This structured summarization offers insight into individuals' psychological states by highlighting the most significant mental health challenges. The derived prototype will be used as a prediction model shown in Figure 10(b).

Table 8
Factor-wise text summarization and prediction model for EMS-Boys dataset, where the factor precedence is from top to down

Factor	Statements	Combined summary
Factor-1	S7, S13, S17, S21, S29	Things; energy; safe; time yourself
Factor-2	S5, S9, S20, S22	Isolated; goals; discouraged; future
Factor-3	S3, S8, S15, S30	Grouchy; tired; problems; pressure deadlines
Factor-4	S6, S18	Situations conflict; worries
Factor-5	S14	Tense
Factor-6	S16, S28	Hurry; responsibility
Factor-7	S4, S11, S26	Things; decisions; mentally exhausted
Factor-8	S10, S27	Calm; trouble relaxing
Factor-9	S1	Rested
Factor-10	S24, S25	Criticized, lighthearted

Table 9 presents a factor-wise text summarization for the BMS-Girls dataset (see Figure 8(c)), ranking mental health themes from highest to lowest priority. Key concerns include problems, hurriedness, future uncertainty, demands, isolation, tiredness, criticism, responsibility, worries, and discouragement. Additionally, factors highlight emotional states such as frustration, grouchiness, conflict, trouble relaxing, mental exhaustion, and goal-setting, providing insight into the psychological well-being of individuals. The derived prototype will be used as a prediction model shown in Figure 10(c).

Table 9
Factor-wise text summarization and prediction model for BMS-Girls dataset, where the factor precedence is from top to down

Factor	Statements	Combined summary
Factor-1	S15, S16, S22	Problems; hurry; future
Factor-2	S2, S5, S8	Temands; isolated; tired
Factor-3	S24, S28	Criticized; responsibility
Factor-4	S7, S21, S25	Things; enjoy; lighthearted
Factor-5	S18, S20	Worries; discouraged
Factor-6	S10, S13, S17, S26, S30	Calm; energy; safe; mentally exhausted; pressure deadlines
Factor-7	S3, S6	Grouchy; situations conflict
Factor-8	S1, S27, S12	Trouble relaxing; frustrated
Factor-9	S23	Things
Factor-10	S9	Goals

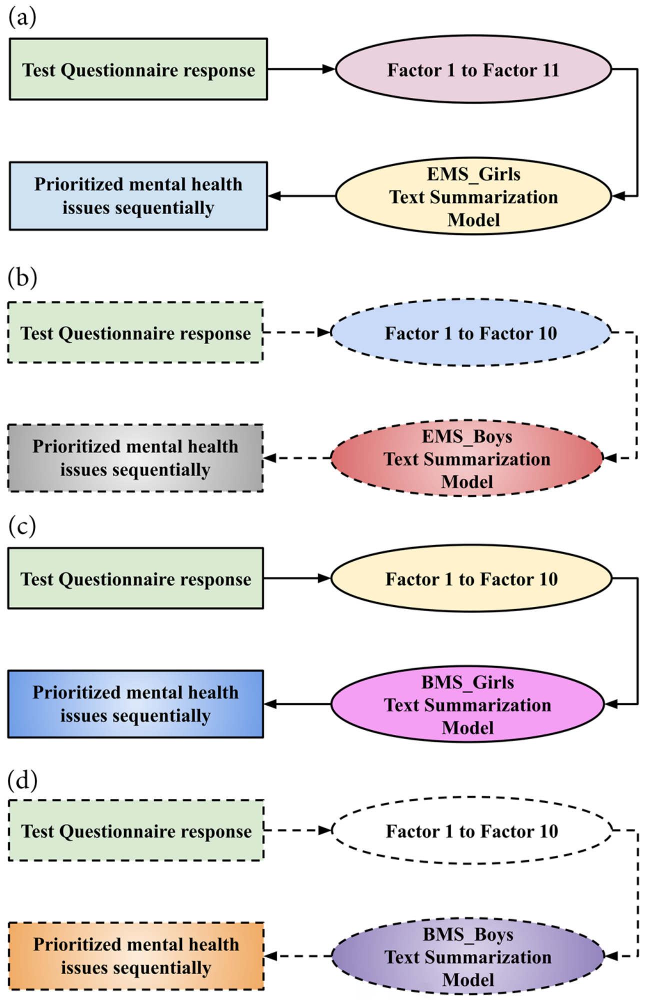
Table 10 presents a factor-wise text summarization for the BMS-Boys dataset (see Figure 8(d)), prioritizing mental health themes from highest to lowest. Key concerns include emotional states such as grouchiness, frustration, tiredness, and tension, along with issues related to pressure, decision-making, and responsibility. Additional factors

Table 10
Factor-wise text summarization and prediction model for BMS-Boys dataset, where the factor precedence is from top to down

Factor	Statements	Combined summary
Factor-1	S3, S7, S10, S13, S24	Grouchy; things; calm; energy; criticized
Factor-2	S8, S12, S16, S27, S30	Tired; frustrated; hurry; trouble Relaxing; pressure deadlines
Factor-3	S9, S22, S29	Goals; future; time yourself
Factor-4	S11, S14, S15, S19, S21	Decisions; tense; problems; pres-sure people; enjoy
Factor-5	S2, S6, S28	Demands; situations conflict; responsibility
Factor-6	S1, S23	Rested, things
Factor-7	S17	Safe
Factor-8	S18, S26	Worries, mentally exhausted
Factor-9	S5	Isolated
Factor-10	S4	Things

highlight aspects of relaxation, future goals, isolation, and mental exhaustion, offering insight into individuals' psychological well-being. The derived prototype will be used as a prediction model shown in Figure 10(d).

Figure 10
Demonstration of the prediction of models for MHS derived from the factor analysis comparison for Girls & Boys of EMS and BMS datasets



The factor analysis reveals distinct stress clusters. For EMS-Girls, high-loading items indicate stress linked to academic overload and emotional fatigue (e.g., S2, S27, and S28). BMS students, in contrast, show more diffuse factor loading, suggesting a broader but less intense pattern of stress. These latent constructs support the hypothesis that EMS students experience more concentrated and intense stress dimensions.

The summarization of this prediction model has been shown in Table 11 that shows that the best performance is achieved with an SVM classifier for EMS-Girls indicating that mental health prediction models based on stress-related factors are effective, particularly in English-medium contexts where stress dimensions are more coherent and predictive.

Hence, the findings of this study show that EMS students, especially girls, experience higher levels of perceived stress. This aligns with cognitive load theory, where non-native language instruction imposes dual processing demands—comprehending content and

Table 11
Performance evaluation for prediction model utilized for MHS among students

Group	Accuracy	Precision	Recall	F1-score
EMS-Girls	84.2%	0.82	0.85	0.83
EMS-Boys	81.7%	0.80	0.79	0.79
BMS-Girls	79.4%	0.77	0.80	0.78
BMS-Boy	76.3%	0.75	0.74	0.74

translating language—thereby increasing mental fatigue. The richer factor structure observed in EMS-Girls indicates more defined and intense stress clusters, likely influenced by high academic expectations and linguistic challenges. These results suggest that instructional language is not a neutral medium but a meaningful variable in student well-being. While the findings are promising, there may be one limitation that is acknowledged, that this study is geographically limited to students in West Bengal state of India, which may affect generalizability. The future work should explore longitudinal data, broader geographic samples, and adaptive interventions based on real-time stress monitoring.

4. Conclusions

This study critically examines the perceived stress levels among students in English Medium Schools (EMS) and Bengali Medium Schools (BMS), utilizing a multi-stage ICT-enabled approach involving questionnaire formulation, data collection, statistical analysis, and predictive modeling. The findings provide substantial insights into how language medium and gender intersect to influence student mental health. Across all hypotheses tested, the results revealed significant patterns: (i) girls' stress levels remain consistent across both EMS and BMS, (ii) boys in BMS show higher stress levels than those in EMS, (iii) EMS students—both girls and boys—exhibit higher perceived stress compared to their BMS counterparts, and (iv) EMS students demonstrate more complex and concentrated stress factor structures, indicating deeper psychological stress. These outcomes highlight the need for tailored mental health support systems, particularly in English medium educational environments.

This study provides empirical evidence that instructional language medium and gender significantly influence students' perceived stress. Using validated psychometric tools and multi-stage machine learning analytics, we found that EMS students, especially girls exhibited both higher stress levels and more complex latent stress dimensions than their BMS counterparts. These results emphasize the need for context-sensitive mental health interventions in ICT-based education environments. Furthermore, the predictive models developed herein demonstrate the feasibility of early detection systems for student stress based on simple questionnaire inputs. Ultimately, the study emphasizes the critical role of educational context and language medium in shaping students' psychological well-being, paving the way for more inclusive and adaptive educational policies.

This research advances the scientific understanding of how ICT-integrated education affects children's mental health by combining validated psychometric tools with AI-based prediction models—an approach rarely applied in this field. By integrating demographic factors and psychological assessment data, we developed a framework capable of predicting stress, anxiety, and depression risk levels in students. This moves beyond previous descriptive studies by offering a practical, predictive tool that schools can use for early detection and targeted interventions. In doing so, the study bridges a key gap in the literature between educational technology research and child mental health analytics, providing a foundation for future cross-disciplinary work and evidence-based policy development.

Acknowledgment

We gratefully acknowledge the support of the school authorities and the participation of the students who contributed their responses during data collection.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are available from the corresponding author upon reasonable request.

Author Contribution Statement

Kunal Ghosh: Conceptualization, Methodology, Software, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Saiyed Umer:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Alenezi, M., Wardat, S., & Akour, M. (2023). The need of integrating digital education in higher education: Challenges and opportunities. *Sustainability*, *15*(6), 4782. https://doi.org/10.3390/su15064782

[2] Ben Youssef, A., Dahmani, M., & Ragni, L. (2022). ICT use, digital skills and students' academic performance: Exploring the digital divide. *Information*, *13*(3), 129. https://doi.org/10.3390/info13030129

[3] Muzumdar, P., Basyal, G.P., & Vyas, P. (2022). An empirical comparison of machine learning models for student's mental health illness assessment. *Asian Journal of Computer and Information Systems*, *10*(1), 1–10. https://doi.org/10.24203/ajcis.v10i1.6882

[4] Ros Faizah binti Mohd, R. F., Zulkifli, H., Hamzah, M. I., & Tamuri, A. H. (2024). Systematic literature review of ICT integration in teaching and learning. *TEM Journal*, *13*(4), 3146–3159. https://doi.org/10.18421/TEM134-49

[5] Kyriakidis, A., Zervoudakis, K., Krassadaki, E., & Tsafarakis, S. (2025). Exploring the impact of ICT on higher education teaching during COVID-19: Identifying barriers and opportunities through advanced text analysis on instructors' experiences. *Journal of the Knowledge Economy*, *16*(3), 11896-11926 https://doi.org/10.1007/s13132-024-02346-5

[6] Narzary, D., Sharma, U., & Khanna, A. (2025). An automated stress detection model based on dual approach of clinical psychologist prediction and machine learning. *International Journal of Information Technology*, *17*(2), 755–765. https://doi.org/10.1007/s41870-024-02213-1

[7] Lin, Y., & Yu, Z. (2025). An integrated bibliometric analysis and systematic review modelling students' technostress in higher education. *Behaviour & Information Technology*, *44*(4), 631–655. https://doi.org/10.1080/0144929X.2024.2332458

[8] Vouglanis, T., & Driga, A. M. (2023). The use of ICT for the early detection of dyslexia in education. *TechHub Journal*, *5*, 54–67.

[9] Pagán Garbín, I., Méndez, I., & Martínez-Ramón, J. P. (2024). Exploration of stress, burnout and technostress levels in teachers: Prediction of their resilience levels using an artificial neural network (ANN). *Teaching and Teacher Education*, *148*, 104717. https://doi.org/10.1016/j.tate.2024.104717

[10] Meca, A., & Schwartz, S. J. (2024). Cultural stress theory: An overview. *Cultural Diversity & Ethnic Minority Psychology*, *30*(4), 603–612. https://doi.org/10.1037/cdp0000704

[11] Chen, Y., Zhang, L., & Yin, H. (2022). A longitudinal study on students' foreign language anxiety and cognitive load in gamified classes of higher education. *Sustainability*, *14*(17), 10905. https://doi.org/10.3390/su141710905

[12] Balsalobre-Lorente, D., Abbas, J., He, C., Pilař, L., & Shah, S. A. R. (2023). Tourism, urbanization and natural resources rents matter for environmental sustainability: The leading role of AI and ICT on sustainable development goals in the digital era. *Resources Policy*, *82*, 103445. https://doi.org/10.1016/j.resourpol.2023.103445

[13] Olatunde-Aiyedun, T. G. (2024). Artificial intelligence (AI) in education: integration of AI into science education curriculum in Nigerian universities. *International Journal of Artificial Intelligence for Digital, 1*(1), 1–14. https://doi.org/10.13140/RG.2.2.31699.76320

[14] Celik, I. (2023). Exploring the determinants of artificial intelligence (AI) literacy: Digital divide, computational thinking, cognitive absorption. *Telematics and Informatics*, *83*, 102026. https://doi.org/10.1016/j.tele.2023.102026

[15] Bibri, S. E., Alexandre, A., Sharifi, A., & Krogstie, J. (2023). Environmentally sustainable smart cities and their converging AI, IoT, and big data technologies and solutions: An integrated approach to an extensive literature review. *Energy informatics*, *6*(1), 9. https://doi.org/10.1186/s42162-023-00259-2

[16] Xiao, J., & Boschma, R. (2023). The emergence of artificial intelligence in European regions: The role of a local ICT base. *The Annals of Regional Science*, *71*(3), 747–773. https://doi.org/10.1007/s00168-022-01181-3

[17] Zhang, Y., Yang, X., Li, X., Yu, S., Luan, Y., Feng, S., Wang, D., & Zhang, Y. (2024). *PsyDraw: A multi-agent multimodal system for mental health screening in left-behind children.* arXiv. https://doi.org/10.48550/arXiv.2412.14769

[18] Agarwal, J., & Sharma, S. (2024). Artificial intelligence enabled cognitive computer-centered digital analysis model for examination of the children's mental health. *Evolutionary Intelligence*, *17*(5), 3675–3685. https://doi.org/10.1007/s12065-024-00951-6

[19] Levenstein, S., Prantera, C., Varvo, V., Scribano, M. L., Berto, E., Luzi, C., & Andreoli, A. (1993). Development of the perceived stress questionnaire: A new tool for psychosomatic research. *Journal of Psychosomatic Research, 37*(1), 19–32. https://doi.org/10.1016/0022-3999(93)90120-5

[20] Zakariya, Y. F. (2022). Cronbach's alpha in mathematics education research: Its appropriateness, overuse, and alternatives in estimating scale reliability. *Frontiers in Psychology*, *13*, 1074430. https://doi.org/10.3389/fpsyg.2022.1074430

[21] Kaur, P., Stoltzfus, J., & Yellapu, V. (2018). Descriptive statistics. *International Journal of Academic Medicine*, *4*(1), 60–63. https://doi.org/10.4103/IJAM.IJAM_7_18

[22] Rahayu, N. I., Muktiarni, M., & Hidayat, Y. (2024). An application of statistical testing: A guide to basic parametric statistics in educational research using SPSS. *ASEAN Journal of Science and Engineering*, *4*(3), 569–582. https://doi.org/10.17509/ajse.v4i3.76092

[23] Wall Emerson, R. (2023). Mann-Whitney *U* test and *t*-test. *Journal of Visual Impairment & Blindness*, *117*(1), 99–100. https://doi.org/10.1177/0145482X221150592

[24] Zhu, J., Yang, Y., Zhao, Y., Qin, G., & Su, X. (2024). Using structural equation modeling to examine correlations between health-related physical fitness and cell health among Chinese college students. *BMC Public Health*, *24*(1), 1557. https://doi.org/10.1186/s12889-024-19067-8

[25] Ni, P., Zhang, S., Zhang, G., Zhang, W., Zhang, H., Zhu, Y., ..., & Diao, M. (2025). Development and validation of machine learning-based prediction model for outcome of cardiac arrest in intensive care units. *Scientific Reports*, *15*(1), 8691. https://doi.org/10.1038/s41598-025-93182-3

[26] Sutar, S., Surve, I., Munawwar, M., Nanaware, V., & Dhumal, P. (2024). Text Summarization Using NLP. *International Journal of Novel Research and Development*, *9*(2), b350–b353.

[27] Thams, N., Saengkyongam, S., Pfister, N., & Peters, J. (2023). Statistical testing under distributional shifts. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *85*(3), 597–663. https://doi.org/10.1093/jrsssb/qkad018

How to Cite: Ghosh, K., & Umer, S. (2026). Predicting Mental Health Outcomes for Native and Non-Native Language Students Through ICT-Based Learning Analytics. *Journal of Data Science and Intelligent Systems, 4*(3), 365–377. https://doi.org/10.47852/bonviewJDSIS62025867