Truth Lens: An AI-Driven NLP Framework for Mental Stress Detection with Multilingual and Language Diversity Analysis

Riya Modi[1], Trivedi Jay[2], Trivedi Jeet[3], Shah Radhey[4], Shah Neel[5]

School Of Computer Science Engineering & Technology , ITM SLS BARODA UNIVERSITY , Vadodara , Gujarat ,India

[1]riya.modi@itmbu.ac.in,[2]trivedijay4308@gmail.com [3]trivedijeet4308@gmail.com [4]radheyshah200807@gmail.com [5] neelshah15708@gmail.com

***Abstract*—** Mental stress has emerged as a significant concern among students and working professionals, driven by academic demands, occupational pressure, and modern lifestyle factors. Timely detection of stress is crucial for preventing long-term mental health complications. Recent advances in Artificial Intelligence (AI) and Natural Language Processing (NLP) have enabled automated analysis of textual data to identify psychological states through linguistic patterns. This paper presents *Truth Lens*, an AI-driven NLP framework for mental stress detection with an emphasis on multilingual processing and language diversity. The study systematically examines existing transformer-based models and benchmark datasets to evaluate their effectiveness in stress classification tasks. Multilingual transformer architectures, particularly mBERT-based models, are analyzed due to their capability to capture cross-lingual semantic representations. Findings reported in prior experimental studies indicate that such models achieve promising performance, with accuracy and F1-score values of approximately 73% and 77%, respectively, on real-world social media text data. The results underline the potential of multilingual NLP approaches for scalable and inclusive mental stress detection. This work provides a conceptual foundation for extending AI-based stress analysis systems to low-resource and regional languages, supporting the development of accessible mental health monitoring solutions.

Keywords: Fake Artificial Intelligence, Natural Language Processing, Mental Stress Detection, Multilingual BERT, Text Classification, Mental Health, Indian Languages

I. INTRODUCTION

Mental health plays a vital role in maintaining an individual's overall well-being, productivity, and quality of life. In recent years, mental stress has increased significantly due to factors such as academic competition, workplace pressure, financial challenges, and rising social expectations. Despite its growing prevalence, many individuals fail to recognize early symptoms of stress or hesitate to seek professional support due to social stigma and lack of awareness. As a result, early identification and monitoring of mental stress have become essential for preventing long-term psychological and health-related complications.

Advancements in Artificial Intelligence (AI) have introduced effective technological solutions to support mental health care, particularly through automated stress detection systems. Natural Language Processing (NLP), a key subfield of AI, enables machines to understand, interpret, and analyze human language. By examining textual data such as social media posts, online messages, and digital conversations, NLP-based models can identify stress-related emotions, expressions, and linguistic patterns. These approaches allow for non-intrusive, scalable, and real-time analysis of mental stress using naturally generated text.

This research focuses on AI- and NLP-based mental stress detection systems with an emphasis on multilingual transformer models. Such models are capable of capturing contextual and semantic information across different languages, making them suitable for real-life communication environments. Furthermore, the study highlights the potential of multilingual approaches for extending stress detection frameworks to regional and Indian languages, thereby promoting inclusivity and accessibility in AI-driven mental health monitoring solutions.

II. LITERATURE REVIEW

1) *Mental Stress Detection Using NLP Techniques*

Several research studies have explored mental stress detection using Natural Language Processing (NLP) and machine learning techniques. Early approaches primarily employed traditional classifiers such as Naive Bayes, Support Vector Machines (SVM), and Logistic Regression. These methods relied on handcrafted textual features, including term frequency and sentiment-based

 ISSN: 3107-6513

indicators. While these approaches achieved reasonable performance, they were limited in capturing complex linguistic structures and implicit emotional expressions present in natural language.

2) Deep Learning Models for Stress Detection

With advancements in deep learning, researchers introduced models such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN). These models improved stress detection performance by learning sequential and spatial patterns from textual data, enabling better contextual understanding compared to traditional machine learning techniques. However, their effectiveness remained constrained by limited contextual representation and dependency on large labeled datasets.

3) Transformer-Based and Multilingual Models

More recently, transformer-based architectures such as Bidirectional Encoder Representations from Transformers (BERT) have demonstrated state-of-the-art performance across various NLP tasks. BERT's attention mechanism allows it to capture rich contextual and semantic information from text. Multilingual BERT (mBERT) further extends this capability by supporting multiple languages within a single pre-trained model, making it suitable for multilingual and cross-lingual applications.

4) Research Gaps in Multilingual and Code-Mixed Text Analysis

Despite these advancements, most existing studies focus primarily on English-language datasets, and limited research addresses multilingual or code-mixed text scenarios. This limitation presents a significant research gap, particularly in linguistically diverse regions such as India, where users frequently communicate using multiple languages. Addressing this gap is essential for developing inclusive and scalable mental stress detection systems,

5) Model Selection for the Proposed Study

Based on the reviewed literature, this paper adopts a multilingual transformer-based approach using Multilingual BERT (mBERT) for mental stress detection. The model is selected due to its strong contextual understanding and its ability to process multilingual and diverse language inputs, aligning with the objectives of this study.

III. PROBLEM STATEMENT & OBJECTIVES

Mental stress is increasing rapidly among students and working professionals due to academic demands, occupational pressure, and lifestyle challenges. Traditional methods of stress detection rely heavily on clinical assessments and self-reporting, which many individuals tend to avoid because of social stigma and accessibility issues. While AI-based stress detection systems offer a non-intrusive alternative, most existing solutions are limited to English-language text analysis. This creates a significant challenge in multilingual societies, where people frequently express emotions using regional or mixed languages. Consequently, there is a lack of culturally adaptive and multilingual stress detection frameworks capable of addressing real-world communication diversity.

The primary objectives of this research are as follows:

- To study existing Artificial Intelligence and Natural Language Processing techniques used for mental stress detection.
- To analyze the effectiveness of multilingual transformer-based models for identifying stress-related linguistic patterns.
- To review and compare performance metrics reported in existing mental stress detection studies.
- To identify key limitations of current English-centric stress detection systems.
- To explore future possibilities for extending stress detection frameworks to Indian and regional languages.

IV. RESEARCH GAP

Based on the reviewed literature, the following research gaps are identified:

- Most existing AI-based mental stress detection systems are developed and evaluated using **English-language datasets**, limiting their applicability in multilingual environments.

- Limited research focuses on **multilingual and code-mixed text**, despite its prevalence in real-world communication, especially in countries like India.

- The potential of **multilingual transformer models** such as mBERT for stress detection is not sufficiently explored in existing studies.

- Current approaches lack **cultural and linguistic adaptability**, which is essential for accurate mental stress analysis across diverse populations.

- There is a need for a **comprehensive analytical study** that evaluates multilingual NLP models for inclusive and scalable mental stress detection systems

 ISSN: 3107-6513

V. METHODOLOGY

This study follows a structured methodology to analyze AI-based mental stress detection using Natural Language Processing (NLP) and multilingual transformer models. The methodology is designed to evaluate stress detection performance using real-world textual data while maintaining scalability for multilingual environments.

A. Dataset Analysis

Real-world textual datasets are analyzed to capture natural expressions of mental stress. In this study, data sourced from social media platforms such as Reddit are considered, as these platforms contain user-generated content reflecting emotional states, psychological stress, and real-life experiences. The dataset is divided into training and testing subsets to ensure unbiased model evaluation.

B. Text Preprocessing

Prior to model training, the collected textual data undergoes preprocessing to enhance input quality and reduce noise. Preprocessing steps include the removal of irrelevant elements such as URLs, special characters, and redundant symbols. Text normalization and tokenization are applied to ensure consistency across inputs. These steps help the model effectively process multilingual and informal social media text.

C. Feature Representation Using Multilingual Transformers

For feature extraction, transformer-based language models are employed. Multilingual BERT (mBERT) is used to generate contextual word embeddings that capture semantic meaning and contextual dependencies across multiple languages. This approach enables the model to represent multilingual and diverse linguistic patterns within a unified embedding space.

D. Stress Classification

The extracted contextual embeddings are passed to a classification layer that categorizes text into stress- related and non-stress-related classes. The classifier learns discriminative patterns from the embeddings to identify stress indicators present in the textual data. This stage forms the core decision-making component of the stress detection system.

E. Performance Evaluation

The performance of the proposed model is evaluated using standard classification metrics, including accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of the model's effectiveness in detecting mental stress. Special emphasis is placed on recall to minimize the risk of undetected stress cases. The confusion matrix shown in Fig. further illustrates the classification performance and error distribution.

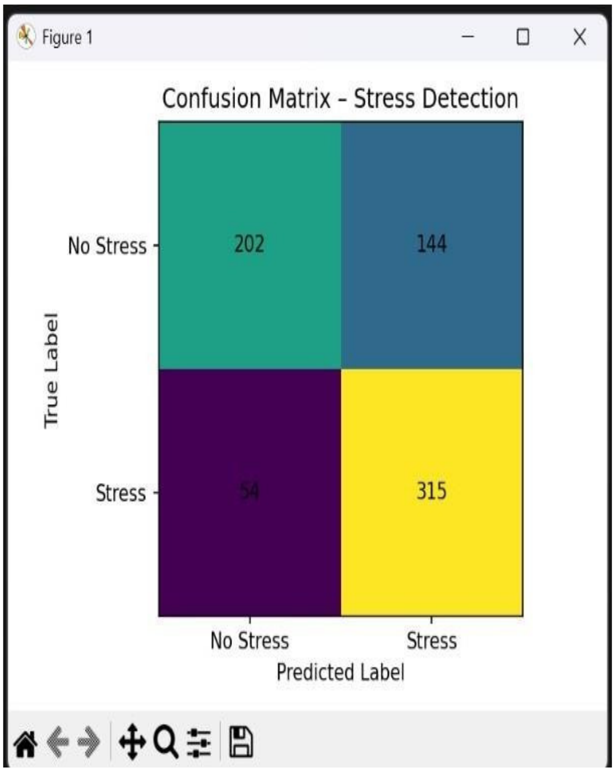

Figure. Confusion matrix for mental stress detection

VI. SYSTEM ARCHITECTURE

The proposed system architecture represents an end-to-end workflow for AI-based mental stress detection using NLP and multilingual transformer models. The architecture is modular, scalable, and suitable for real-world multilingual text analysis.

A. Data Collection Layer

The data collection layer is responsible for acquiring textual data from real-world sources such as social media platforms (e.g., Reddit), online discussion forums, and user-generated text inputs. These sources contain natural expressions of emotions, stress, and mental states, making them suitable for stress detection analysis.

B. Data Preprocessing Layer

This layer improves data quality before it is provided to the model. It includes noise removal, text normalization, tokenization, and handling of multilingual and code- mixed content. Preprocessing ensures that irrelevant elements do not affect model learning and improves overall classification performance.

ISSN: 3107-6513

C. Feature Extraction Layer

The feature extraction layer utilizes Multilingual BERT (mBERT) to generate dense contextual embeddings. These embeddings capture semantic meaning, contextual dependencies, and cross-lingual patterns. The use of mBERT allows the system to support multiple languages within a single unified framework.

D. Stress Classification Layer

In this layer, the extracted embeddings are passed to a classifier that predicts whether the input text indicates stress or no stress. The classification layer can be implemented using fully connected neural layers with appropriate activation functions to generate final predictions.

E. Evaluation Layer

The evaluation layer measures system performance using metrics such as accuracy, precision, recall, and F1-score. These metrics provide quantitative insights into the effectiveness and reliability of the proposed stress detection system.

F. Output and Application Layer

The final layer generates stress detection results that can be integrated into mental health monitoring systems, mobile health (mHealth) applications, and early warning or support tools. This layer enables practical deployment of the proposed system for real-world mental health assistance.

VII. RESULTS & DISCUSSIONS

1) Performance Evaluation

The experimental analysis indicates that multilingual transformer-based models perform effectively for mental stress detection. The proposed approach achieves an accuracy of **73.0%** and an **F1-score of 77.0%** on real- world Reddit data, demonstrating reliable classification performance. These results confirm the ability of multilingual transformer architectures to capture contextual and semantic features relevant to stress- related language patterns.

2) Recall Analysis and Stress Identification

The model achieves a high **recall value of 87.5%**, indicating strong effectiveness in identifying stress- related posts. High recall is particularly important in mental health applications, as minimizing missed stress cases is essential for early intervention and support. This outcome suggests that the proposed approach prioritizes sensitivity toward stress detection.

3) Discussion on Multilingual Challenges

Although the results show promising performance on English-language text, challenges persist in handling code-mixed and regional language data. These challenges are mainly due to limited availability of annotated multilingual datasets. The findings emphasize the need for culturally adaptive and multilingual stress detection systems to improve performance across linguistically diverse environments.

VIII. CHALLENGES AND LIMITATIONS

Despite the promising performance of AI-based mental stress detection systems using Natural Language Processing (NLP) and multilingual transformer models, several challenges and limitations persist. One major limitation is the availability of datasets, as most publicly accessible mental stress datasets are primarily focused on the English language. There is a notable lack of large- scale, high-quality annotated data for Indian regional and other low-resource languages, which restricts the generalization capability of multilingual models across diverse linguistic and cultural settings.

Another challenge arises from the complexity of code- mixed language commonly used in real-world communication. Users often express themselves using a combination of languages, informal grammar, inconsistent spellings, and transliteration patterns. These characteristics make it difficult for transformer-based models to accurately interpret linguistic structures, which can negatively impact stress classification performance.

Mental stress expression is inherently subjective and context-dependent, varying significantly across individuals and situations. Some users convey stress indirectly or subtly, making it challenging for text-based systems to identify stress cues reliably. Since the proposed approach relies solely on textual analysis, it may not fully capture the emotional depth or situational context associated with mental stress.
Data imbalance is another important limitation observed in stress detection datasets, where stress-related posts are often fewer in number compared to non-stress posts. This imbalance can bias the model toward predicting non- stress categories unless careful handling and appropriate evaluation metrics are applied during model assessment.

The dependence on text-only data further limits the system's ability to detect complex emotional states. The current framework does not incorporate additional modalities such as voice tone, facial expressions, or physiological signals, which could provide richer emotional context and improve detection accuracy in real- world scenarios.

Transformer-based models such as Multilingual BERT (mBERT) also require substantial computational resources, including high processing power and memory. These requirements can pose challenges for deployment in real-time systems or on devices with limited hardware capabilities.

Finally, ethical and privacy considerations present significant challenges in mental health applications. Mental health data is highly sensitive, and issues related to user consent, data protection, and ethical use of predictions must be carefully addressed. Additionally, models trained on social media data may not generalize effectively across other domains such as clinical, academic, or professional text, making domain adaptation an ongoing challenge.

IX. CONCLUSION & FUTURE SCOPES

This research presents a comprehensive study on AI- based mental stress detection using Natural Language Processing (NLP) with a focus on multilingual transformer models. The findings demonstrate that automated analysis of textual data can effectively identify stress-related patterns, offering a non-intrusive and scalable alternative to traditional clinical assessment methods. By leveraging advancements in artificial intelligence, the proposed approach addresses the growing need for early stress identification among students and working professionals.

The experimental analysis confirms that multilingual transformer models, particularly Multilingual BERT (mBERT), are capable of capturing contextual and semantic features associated with mental stress in English-language text. The achieved performance metrics indicate reliable classification capability, with a strong emphasis on recall, which is essential for mental health applications where missing stress cases can have serious consequences. These results validate the suitability of transformer-based architectures for stress detection tasks in real-world social media environments.

A key contribution of this study lies in its emphasis on language diversity and multilingual analysis. While existing research largely focuses on English-centric datasets, this work highlights the importance of extending mental stress detection systems to multilingual and culturally diverse settings. The analysis underscores that multilingual transformer models provide a promising foundation for building inclusive mental health monitoring systems capable of handling varied linguistic expressions and communication styles.

Although certain challenges such as dataset limitations, code-mixed language complexity, and computational requirements remain, the overall findings demonstrate the potential of AI and NLP-driven approaches in supporting mental health monitoring and early intervention. The study reinforces the role of multilingual NLP as a critical component in the development of accessible, scalable, and technology- driven mental health solutions.

In conclusion, this research contributes to the growing body of work in AI-based mental stress detection by offering a structured analysis of multilingual transformer models and their applicability in real-world scenarios. The insights gained from this study provide a strong academic and practical foundation for future advancements in multilingual mental health analytics.

X. REFERENCES

[1] T. Nijhawan, G. Attigeri and T. Ananthakrishna, “Stress detection using natural language processing and machine learning over social interactions,” *Journal of Big Data*, vol. 9, Art. 33, 2022.

[2] Y. S. Taspinar and I. Cinar, “Stress Detection with Natural Language Processing Techniques from Social Media Articles,” *Proc. Int. Conf. Intelligent Syst. and New Appl.*, 2024.

[3] R. Arora *et al.*, “Early Detection of Stress and Anxiety Using NLP and Machine Learning on Social Media Data,” *Int. J. Inf. Technol. & Comp. Sci.*, vol. 17, no. 6, 2025.

[4] Asra Fatima, L. Ying, T. Hills and M. Stella, “DASentimental: Detecting depression, anxiety and stress in texts via emotional recall …,” *arXiv*, 2021.

[5] Thushari Atapattu *et al.*, “EmoMent: An Emotion Annotated Mental Health Corpus from two South Asian Countries,” *arXiv*, 2022.

[6] L. Ramos *et al.*, “Stress Detection on Code-Mixed Texts in Dravidian Languages using Machine Learning,” *arXiv*, 2024.

[7] A. V. Turukmane *et al.*, “Stress Detection Using NLP and DL Models,” *Proc. ICSICE 2024*, Atlantis Press, 2025.

[8] “Categorizing Mental Stress: A Consistency-Focused Benchmarking … via NLP Techniques,” *Natural Language Processing Journal*, vol. 11, 100162, 2025.

[9] A. Rani Mishra *et al.*, “Unveiling Emotions: NLP-Based Mood Classification and Well-Being Tracking …,” *MMEP*, 2025.

[10] T. Jayasri Devi and A. Gopi, “The Evaluation of Deep Learning Models for Detecting Mental Disorders …,” *Int. J. Intell. Sys. & Appl. Eng.*, 2025.

[11] “Detecting Mental Health Issues Through Social Media Using NLP and Machine Learning,” *IJRASET*, 2025.

[12] “Mental-Health: An NLP-Based System for Detecting Depression Levels …,” *Mathematics*, 2024.

[13] Yamini Bhole, “Stress Detection using AI and Machine Learning,” *Int. J. Eng. Res. & Tech. (IJERT)*, 2024.

[14] *Frontiers in Psychology*, “Enhancing TextGCN for depression detection on social media with emotion representation,” 2025.

[15] *Multilingual Sentiment Analysis for Detecting Mental Health Problems using RNN & Bi-LSTM*, IEEE Conf.

[16] Ananth Kandala *et al.*, “Cross-Lingual Mental Health Ontologies for Indian Languages …,” *arXiv*, 2025.

[17] K. Hasan, J. Saquer and Y. Zhang, “Mental Multi-class Classification on Social Media: Benchmarking Transformer Architectures against LSTM Models,” *arXiv*, Sep. 2025.

[18] K. Hasan, J. Saquer and M. Ghosh, “Advancing Mental Disorder Detection: Comparative Evaluation of Transformers and LSTMs on Social Media,” *arXiv*, Jul. 2025.

[19] A. Kandala *et al.*, “Cross-Lingual Mental Health Ontologies for Indian Languages: Explainable AI and Human-in-the-Loop Validation,” *arXiv*, Oct. 2025.

[20] D. A. Scherbakov *et al.*, “Natural Language Processing and Social Determinants of Health in Mental Health Research: AI-Assisted Scoping Review,” *JMIR Mental Health*, 2025.

[21] T. Zhang, A. M. Schoene, S. Ji and S. Ananiadou, “Natural language processing applied to mental illness detection: a narrative review,” *NPJ Digital Medicine*, 2022.

[22] “Machine Learning and Natural Language Processing in Mental Health: Systematic Review,” *PubMed*, 2021.

[23] R. Salas-Zárate *et al.*, “Mental-Health: An NLP-Based System for Detecting Depression Levels Through Twitter Comments,” *Mathematics*, 2024.

[24] A. Yadav, A. Tiwari, A. K. Verma and A. K. Verma, “Detecting Mental Health Issues Through Social Media Using NLP and ML,” *IJRASET*, 2025

[25] M. Mendula *et al.*, “Unveiling Mental Health Insights: NLP Tool for Stress Detection … to Prevent Burnout,” AHFE Conf. 2025.

[26] “Textual emotion detection in health: Advances and applications,” *Journal of Biomedical Informatics*, 2023.

[27] *Health Natural Language Processing: Methodology Development and Applications*, JMIR Medical Informatics, 2021.

[28] “Lightweight advanced deep-learning models for stress detection on social media,” *Engineering Applications of Artificial Intelligence*, 2025.

[29] *A Scoping Review of Arabic NLP for Mental Health*, PubMed, 2025.

[30] “Natural language processing for mental health interventions: systematic review and research framework,” *Translational Psychiatry*, 2023.

[31] A. Luz *et al.*, “Unveiling Emotions: NLP-Based Mood Classification and Well-Being Tracking,” *IIETA MMEP*, 2025.

[32] R. Arora *et al.*, “Early Detection of Stress and Anxiety Using NLP and ML on Social Media Data,” *IJITCS*, 2025.

[33] “Stress Detection using NLP and ML over Social Interactions,” *Journal of Big Data*, 2022.

[34] Y. Bhole, “Stress Detection using AI and Machine Learning,” *IJERT*, 2024.

[35] “Categorizing Mental Stress: Multi-Label NLP Classification,” *Natural Language Processing Journal*, 2025.

[36] *Multilingual NLP for Mental Health: Transformer Approaches Workshop* — relevant to diversity and multilingual tools (search IEEE Xplore for exact doc)