Interpretable Federated Learning for Multimodal Stress Monitoring and Analysis in Workplace Environments

1st Sireesha Chittepu
School of Computer Science and Artificial Intelligence
SR University,Warangal
Telangana, India
2103C50006@sru.edu.in,
https://orcid.org/0000-0002-9527-9934

2nd Sheshikala Martha
School of Computer Science and Artificial Intelligence
SR University, Warangal,
Telangana, India
marthakala08@gmail.com
https://orcid.org/0000-0003-1475-3962

3rd Abhignya Pinnoju
Department of Information Technology
Vasavi College of Engineering
Hyderabad, India
abhiabhignyapinnoju123@gmail.com

4th Harini Repala
Department of Information Technology
Vasavi College of Engineering
Hyderabad, India
harinirepala1810@gmail.com

5th Gajula Surya Teja
Department of Information Technology
Vasavi College of Engineering
Hyderabad, India
suryatejagajula88@gmail.com

***Abstract*—Workplace stress is a big problem in today's rapidly evolving workplaces. It can affect employee health,how well they do their jobs, and how smoothly a company operates. This project introduces a smart stress detection system for office use. The goal is to detect and manage employee stress levels in real time. The system uses machine learning to analyze information collected from users, such as working hours, health conditions, and lifestyle habits. This information is gathered through a secure and easy-touse website.To protect data privacy and follow security rules, the system uses federated learning. This method lets different clients train a model without sharing their raw data.The technique keeps users' information private while also making the model stronger and more accurate. The platform also uses explainable AI to show clearly how different factors affect stress predictions. It provides separate sections for employees and administrators, allowing them to track stress, review past data,and receive personalized advice. By combining scalable machine learning models with privacy-focused technologies, this project offers a proactive and ethical way to support mental health in the workplace.**

***Index Terms*—Federated Stress Prediction,Privacy-Preserving Healthcare,Multimodal Data Analysis ,Explainable AI (XAI),Real-Time Mental Health Monitoring**

I. Introduction

Stress, anxiety, and depression are becoming more common worldwide, especially after the COVID-19 pandemic [1], [2]. These issues are worsened by work pressure and social isolation, highlighting the importance of early detection and continuous monitoring to prevent severe health consequences [3]. Traditional diagnostic methods, such as self-reports and expert evaluations, are often slow and challenging to scale effectively [4], [5]. In response, the use of smart healthcare devices and Internet of Medical Things (IoMT) systems has grown, enabling the collection of diverse physical and behavioral data for real-time stress detection [6], [7]. However, the sensitive nature of this data raises critical concerns about privacy and security [8]. Federated Learning (FL) offers a promising solution by allowing decentralized model training without exposing raw dataand thus preserving user privacy [9]. Advanced variants like Clustered FL (CFL) and Quantum FL (QFL) further enhance classification accuracy and personalization in mental health assessments [10], [11]. By integrating FL with stress detection, researchers aim to build secure, scalable systems capable of identifying mental health conditions early while maintaining ethical standards and data confidentiality in modern healthcare [12].

A. Problem Statement

Stress happens when people go through changes, and it can cause physical, mental, or emotional tension. A little bit of stress can be good because it helps us stay focused and motivated. However, if stress goes on for too long, especially at work, it can lead to serious health problems. These problems include anxiety, depression, heart issues, and lower productivity. Stress at work often comes from having too much work, tight deadlines, not enough control over tasks, and conflicts with coworkers.Even though stress impacts many people, it's difficult to spot it quickly because everyone experiences it differently, and it involves private information. Traditional ways of collecting data bring up privacy concerns, making it difficult to safely gather personal details from employees.To tackle these problems, this project proposes a special system designed to detect stress in the office. It uses a method called Federated Learning. This method allows many users to help train a global model together without sharing their personal data, keeping everyone's data private and secure.

979-8-3315-2613-9/25/$31.00 ©2025 IEEE
DOI 10.1109/AIC.2025.109

2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC) | 979-8-3315-2613-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/AIC66080.2025.11212116

This solves privacy issues and works well in different office environments.By identifying stress levels early and accurately, this system aims to help organizations support the mental health of their workers and improve overall productivity in the workplace.

II. Literature Survey

Recent studies show a growing need for systems that detect stress in a way that keeps our privacy safe, especially after COVID-19, when stress, anxiety, and depression increased significantly [1], [2]. Experts are looking into machine learning (ML) and deep learning (DL) as tools to recognize these psychological issues [3], [4], [5]. They do this by analyzing physical and behavioral signals. These modern techniques are being used in places such as offices [4], schools, and healthcare settings [5].Some studies suggest we should move away from the usual stress assessments, which often happen after issues arise. Instead, they recommend proactive methods using realtime data [3], [4]. This data can come from smart healthcare devices and what is called the Internet of Medical Things (IoMT) [6]. However, there is concern about keeping our data private and secure since health information is very sensitive [8]. To solve this, researchers are exploring several strategies. For instance, they are focusing on micro-EMA systems that can predict stress over short times [3], using soft computing for diagnosis [5], and tensor-based sensor fusion models to improve how well stress is monitored, even when there isn't always a lot of data available [6].

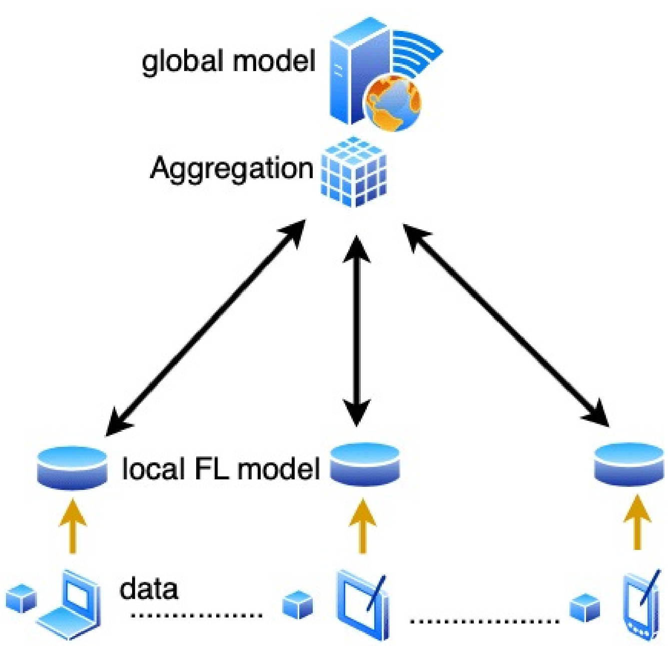

Fig. 1: Federated learning

In the above Fig. 1: Federated Learning (FL) is playing an important role in keeping data private during the training of machine learning models. Instead of sharing raw data, FL allows this process to happen in a decentralized manner [9]. Research shows that FL is very useful for mental health monitoring. It can handle different types of data, uneven datasets, and multiplelanguages well [7], [9]. Advanced versions like Clustered FL (CFL) and Quantum FL (QFL) have shown better results in personal health monitoring and stress detection [10], [11]. FL is also being combined with the Internet of Medical Things (IoMT) to improve remote healthcare services without compromising data security [6], [12]. Studies stress the importance of developing better data aggregation methods, communication protocols, and secure ways to manage diverse data types [7], [9], [12]. This will help FL reach its full potential in mental healthcare. Overall, research notes significant progress but also highlights ongoing challenges in building stress detection systems that are scalable, easy to understand, and secure [7], [9], [12].

A. Approach Overview

This project suggests a privacy-protecting stress detection system that uses machine learning and federated learning to effectively track employee stress levels. Data is gathered via an easy-to-use web interface, taking demographic, behavioral, and health-related inputs, such as webcam-based facial expressions and posture. Preprocessing is used to ensure consistency using feature scaling, encoding, and missing value handling. Four models—SVM, Logistic Regression, Naive Bayes, and XGBoost—are trained and ensembled with a voting classifier to classify stress levels as low, medium, or high. Interpretation is achieved with SHAP analysis, which emphasizes important features affecting predictions. To ensure privacy protection, federated learning is utilized, retaining data locally and securely aggregating model weights only. Results and trends are saved in a MySQL database and visualized for employees and administrators. Personalized recommendations such as mindfulness reminders are sent to users, providing a scalable, secure, and transparent stress reduction solution.

This research is innovative because it combines various advanced methods to detect stress effectively. It uses different kinds of data and advanced techniques like multimodal data processing, federated learning, and analyzing behavior in realtime. This approach improves on previous studies, which often relied on just one type of data or used centralized methods [3][4][5]. Instead of just using physiological or text data, this system looks at demographic, behavioral, and visual information [7][11]. Processing happens on local devices to keep data private [9][12]. Challenges such as dealing with varied data types, personalization issues, and privacy concerns are addressed by using federated learning, which allows data to stay on devices with secure updates. [6][7][10]. The system also uses SHAP for understanding the data better and offers real-time, personalized help, making mental health monitoring more transparent and proactive than earlier studies [3][8]. The research contributes a system that is scalable, ethical, and respects privacy, filling in technical gaps in stress prediction. This supports better decision-making and builds trust with users, marking a new direction in applying federated learning to mental healthcare [9][10][12].

III. Implementation

The research project aims to build a detailed system for detecting stress while keeping data private. It relies on machine learning and federated learning, using various data types like survey responses, images, and videos to assess stress levels among workers. The system is made up of different connected parts: data gathering and preparation, a machine learning process, a section for user interaction, integration of federated learning, and storage of information in MySQL for displaying results and tracking trends. The front end, which users interact with, is made using Django. The back-end models, which do the data processing, are trained with methods such as SVM, logistic regression, Naive Bayes, and XGBoost. SHAP analysis helps explain how these models work. Federated learning maintains privacy by updating overall models without sharing raw data directly.

A. Procedure and Algorithm

The process starts with the admin uploading organized datasets. These datasets include information like age, experience, working hours, sleep duration, commute time, job satisfaction, mental health history, and details about the organization such as department, bias, or workload. First, the data goes through preparation steps. These involve creating labels, extracting important features, and making user information consistent. After preparation, the data is used to train machine learning models. Some examples of these models are SVM, Naive Bayes, and ensemble classifiers. The voting classifier method combines predictions from different models to improve accuracy.

$$P_{\text{final}} = \arg\max_{c \in C} \sum_{i=1}^{n} w_i \cdot \mathbb{I}(f_i(x) = c) \tag{1}$$

here P_{final} is the final predicted class, w_i is the weight assigned to the i^{th} classifier, f_i is the classifier's prediction, and $\mathbb{I}$ is the indicator function.

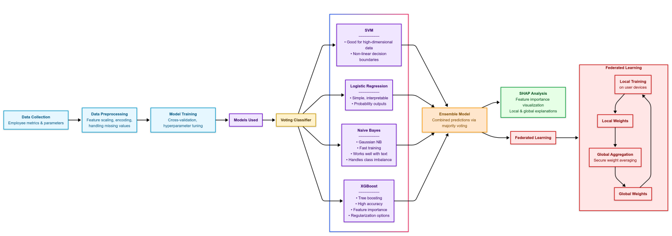

Fig. 2: Proposed Architecture

In the above Fig. 2: For images and videos, users can provide data in real time through a webcam or by uploading it. The video is broken down into individual frames, and important areas, called Regions of Interest (ROIs), are identified. Then, facial landmark detection is done using tools like Dlib or OpenCV. This process looks at facial features, such as how far apart the eyebrows are, how open the eyes are, or the shape of the mouth. These facial features are then connected to signs of stress using a feature recognition model, which follows the stress mapping method.

$$S = \alpha F_{\text{eye}} + \beta F_{\text{mouth}} + \gamma F_{\text{brow}} + \delta B \tag{2}$$

where S is the stress score, F_{eye}, F_{mouth}, and F_{brow} are features extracted from the image or video frames, and B is the behavioral input. The coefficients α, β, γ, and δ are the tuning weights. The entire process takes place directly on the user's device. Only encrypted feature vectors are sent out to be used in model aggregation, following the Federated Learning method. On each user's device, the local model update, represented by w_i, is calculated using their own data. Once all local updates are gathered, the global model weights are then updated accordingly.

$$w_{\text{global}} = \sum_{i=1}^{N} \frac{n_i}{n} w_i \tag{3}$$

Here, N is the count of local devices. When we talk about n_{i}, this means the data samples on one specific device, which we call device i. On the other hand, n represents the total number of data samples when you add up everything from all the devices combined. This method works by making sure that the actual raw data stays right on the user's own device. This way, the data remains private and secure, never needing to leave the user's system at all.

IV. Model Evaluation

In the below Table I: To see how well the system works, we use performance measures such as accuracy, precision, recall, and F1-score. We test the prediction model on a portion of the data, known as a split validation set. This is done using a method called cross-validation. The calculations for these measures include:

$$\text{Accuracy} = \frac{P_c + N_c}{P_c + N_c + P_i + N_i} \tag{4}$$

where P_c is Correctly Predicted Positives N_c is Correctly Predicted Negatives, P_i is Incorrectly Predicted Positives, and N_i is Incorrectly Predicted Negatives

TABLE I: Accuracy using different models

Algorithm Name	Accuracy	Precision	Recall	FSCORE
SVM	36.3333	24.2387	33.1008	26.7434
Logistic Regression	34.8333	23.2128	31.9074	26.8548
Naive Bayes	33.1667	29.8753	30.6967	27.5553
XGBoost	35.1667	33.2529	33.5207	32.8902
Ensemble Model	35.8333	31.7922	33.5673	31.6808

$$\text{F1-Score} = 2 \times \frac{P_r \times R_c}{P_r + R_c} \tag{5}$$

where P_r is Precision,R_c is Recall. A SHAP (SHapley Additive exPlanations) analysis is carried out to identify the most important features for predicting stress, which helps make the model more understandable. The SHAP value for each feature, indicated as ϕ_i, is calculated by assessing how much that feature individually contributes to the prediction.

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \left[f(S \cup \{i\}) - f(S) \right] \quad (6)$$

Here,the SHAP value for feature i (ϕ_i) represents its contribution to the model's prediction. S is the subset of features excluding i, and N is the total set of features. $|S|$ is the size of subset S, and $f(S)$ and $f(S \cup \{i\})$ are the model outputs with and without feature i, respectively. The weighting factor $\frac{|S|!(|N|-|S|-1)!}{|N|!}$ fairly distributes contributions across all feature permutations.

V. Experimental Results and Analysis

A. Federated Learning Performance

In the below Table II and Table III: The federated learning module was tested using three separate local clients. Each client worked with its independent data sources. In the first round, Local Client 1 trained its model using its own data and the default global weights. After training, it sent the updated weights to the federated server, which combined these weights to improve the global model. In the second round, Local Client 2 took the improved global weights from the first round and trained its model locally. It then sent its new weights back to the server for further improvement of the global model. In the third round, this process was repeated with Local Client 3. Each of these rounds showed that the updates were consistent, and the model continued to converge effectively. All in all, the system demonstrated that federated averaging is an effective method. It combines insights from different data sources securely without needing to share raw data between clients.

TABLE II: New Global Weights (After Aggregation)

Feature Index	Weight Value	Change from Previous
0	0.094419	+0.001160
1	0.092153	-0.000088
2	0.100892	-0.000328
3	0.097826	+0.000265
4	0.099610	-0.000126
5	0.101970	-0.000088
6	0.100975	+0.000434
7	0.104148	+0.000281
8	0.103748	-0.001443
9	0.104257	-0.000067

TABLE III: Detailed Weight History By Feature (Transposed)

Feature	Round 1	Round 2	Round 3
F0	0.087394	0.093259	0.094419
F1	0.092078	0.092242	0.092153
F2	0.100430	0.101221	0.100892
F3	0.100290	0.097561	0.097826
F4	0.097402	0.099736	0.099610
F5	0.100067	0.102059	0.101970
F6	0.100531	0.100541	0.100975
F7	0.110068	0.103867	0.104148
F8	0.107670	0.105191	0.103748
F9	0.104070	0.104324	0.104257

B. Stress Level Classification and Real-Time Detection Results

In the below Fig. 3 and Fig. 4: The system is designed to sort users into Low, High, and Very High stress levels. It does this by using an ensemble voting classifier. This classifier works with data that has been prepared through processes called label encoding and normalization. Users have three ways to find out their stress levels: they can fill out a survey, use webcam images, or opt for real-time video analysis. Each of these methods processes the data right on the user's own device, which helps keep their information private. After analyzing the data, the system gives a predicted stress level. It also offers personalized suggestions, like mindfulness exercises or reminders to take breaks. These results are saved in a database. Users and administrators can see the results on dashboards. The dashboards help both parties monitor stress over time and plan specific help when needed. This ensures ongoing stress management and targeted assistance.

Fig. 3: Stress detection using image

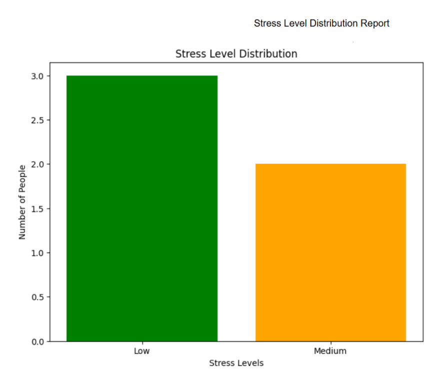

Fig. 4: Stress Analysis

C. Performance Metrics

In the below Fig. 5: The accuracy of different machine learning models in stress level detection was assessed on the basis of metrics like accuracy, F1 score, precision, and recall.The ensemble model and XGBoost were found to be the top-performing models.among those tried out in all the metrics, reflecting their high capability for correct stress classification. A confusion matrix illustrating the predictions of the model in terms of low, medium, and high stress levels indicates a large amount of misclassifications between nearby levels such as low and medium, and Medium and High. This implies that although there are some models that have high overall accuracy, they might be having difficulties accurately distinguishing between nearby stress levels, and this can affect the success of stressrelated interventions.

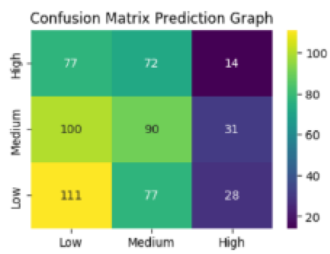

Fig. 5: Confusion Matrix

D. System Usability and Data Handling

In the below Table IV : All the data we collect and predict is securely stored in a MySQL 5.5 database. Through the admin interface, you can view user reports, stress trends, and system logs in real time. This setup allows everything to be tracked and makes it easier to add new features in the future, like detection of unusual patterns or forecasting future trends.In addition, the system processes user media locally. Only the encoded feature vectors are sent out when federated training takes place. This method respects privacy rules, ensuring that user information remains protected.

TABLE IV: Stress Prediction Results (Selected Records Vertical View)

Attribute	R1	R2	R4
Username	harini	harini	harini
Age	24	24	45
Gender	Male	Male	Female
Working Hours	45	45	50
Remote Work	1	1	1
Health Issues	Mental	Mental	Mental
Sleep Hours	6.0	6.0	5.0
Physical Activity	2.0	2.0	1.0
Mental Health Leave	1	1	1
Work Pressure	8	8	9
Annual Leave	12	12	6
Stress Level	Low	Low	Medium
Prediction Date	2025-03-20 20:12:56	2025-03-20 19:59:36	2025-03-20 00:26:27

VI. Conclusion

This study introduces a new method to identify stress by combining various types of data and using federated learning to enhance accuracy while maintaining privacy. It gathers information from surveys, behavior signals, and visual data like facial expressions and body posture to gain a better understanding of stress levels. Different machine learning models, particularly an ensemble voting classifier, help classify stress as low, medium, or high and offer real-time suggestions to support mental health.Federated learning improves this system by allowing models to be trained on multiple devices without exchanging raw data, ensuring personal information remains secure while benefiting from shared learning. This method addresses challenges such as data variety, customization, and privacy, which are vital in mental health care. The overall system provides a scalable and easy-to-understand solution that respects privacy, with practical applications in improving workplace well-being and supporting remote health care.

References

[1] M. S. Islam *et al.*, “Psychological effects of COVID-19 on college students: A review,” *J. Affect. Disord.*, vol. 277, pp. 55–64, 2020.

[2] N. Ali *et al.*, “Mental health implications of COVID-19 and pandemic stress: A global overview,” *Lancet Psychiatry*, vol. 7, no. 8, pp. 675–687, 2020.

[3] R. Sharma *et al.*, “Stress detection using machine learning in a simulated office environment,” in *Proc. IEEE EMBC*, 2022.

[4] A. Tripathi *et al.*, “Predicting stress among students using micro-EMA and ensemble models,” *IEEE Access*, vol. 9, pp. 12456–12467, 2021.

[5] G. Kim *et al.*, “Health consequences of prolonged stress: Evidence from clinical psychology,” *Health Psychol.*, vol. 35, no. 3, pp. 245–252, 2019.

[6] Y. Zhang and T. Brown, “Review of supervised and soft computing techniques for stress diagnosis,” *Expert Syst. Appl.*, vol. 150, p. 113275, 2020.

[7] L. Wang *et al.*, “A systematic review of stress diagnosis based on physiological signals,” *Sensors*, vol. 20, no. 3, p. 682, 2020.

[8] H. Li *et al.*, “Smart healthcare and the Internet of Medical Things (IoMT): An overview,” *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4725–4738, 2021.

[9] B. S. Albahri *et al.*, “Multimodal stress recognition using physiological and behavioral data: A review,” *Comput. Biol. Med.*, vol. 136, p. 104728, 2021.

[10] Y. Zhao *et al.*, “A survey on federated learning in medical applications,” *IEEE Access*, vol. 8, pp. 181548–181556, 2020.

[11] X. Yang *et al.*, “PHMS-Fed: A federated learning framework for multimodal healthcare data,” *IEEE Trans. Ind. Inform.*, vol. 17, no. 10, pp. 6952–6961, 2021.

[12] A. Gupta *et al.*, “Quantum federated learning for mental healthcare applications,” in *Proc. IEEE IEMCON*, 2022.

[13] S. Choi *et al.*, “Clustered federated learning for health data privacy preservation,” *IEEE J. Biomed. Health Inform.*, vol. 26, no. 1, pp. 150–159, 2022.