International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Improving mental health and suggesting coping techniques to IT Professionals by utilizing Random Forest based algorithm

Chitralekha Dwivedi
Asst. Prof., Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology, Pune
chitralekha.dwivedi@dypvp.edu.in

Nishant Surve, Rohit Mehta, Omkar Nawale, Rushikesh Shinde
Final Year Student, Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology, Pune
nishantsurve254@gmail.com

Abstract

Mental health challenges like chronic stress are increasingly affecting professionals in the tech industry, lowering both productivity and life quality. This study presents an AI-driven approach for real-time stress prediction and tailored interventions. Using supervised learning models, including Random Forest, SVM, and Neural Networks; the system categorizes stress levels by analysing behavioural and contextual inputs collected from self-assessments. Based on the predicted stress level (low, moderate, or high), the system suggests personalized remedies like mindfulness sessions, scheduled pauses or focus-enhancing tips. A dashboard tracks stress trends to help users recognize and manage their patterns. The model demonstrates high accuracy and is scalable for real-world application, prioritizing user privacy and adaptability in both personal and professional settings.

Index Terms—Stress prediction, machine learning, personalized recommendations, mental health, automation, stress analytics.

I. INTRODUCTION

In the fast-paced world of technology, stress has become a silent but persistent issue, especially among professionals in the IT sector. While the industry thrives on innovation, it also imposes intense deadlines, long hours, and the constant pressure to keep up with rapidly changing tools and frameworks. This environment often leads to mental health problems such as anxiety, exhaustion, and emotional fatigue.

Rising stress levels don't just take a toll mentally; they also reduce concentration, affect decision-making, and can even lead to serious physical health issues. In workplace settings, this often translates to lower productivity, frequent absences, or even job resignations.

To address this, we propose an intelligent solution that not only detects stress levels but also suggests appropriate coping strategies. At the heart of this system lies the Random Forest algorithm, a machine learning technique known for its reliability, interpretability, and robustness in handling complex and non-linear data.

Unlike a single decision tree, which may yield inconsistent outcomes, a Random Forest builds multiple trees and takes a collective decision, much like gathering advice from several experts. This enhances the system's accuracy, especially when processing messy or diverse data.

The system collects both subjective and objective indicators: daily work duration, workload stress, work-life balance, and physiological cues such as sleep disturbances or fatigue. These inputs are gathered through self-reports.

Once the data is collected, the model classifies the user's stress level and based on this, offers tailored suggestions like breathing exercises, meditation, or simple productivity hacks. The system is interactive and operates in real-time, going beyond static surveys or yearly assessments.

Users are also empowered to visualize their stress trends through a dashboard, helping them notice harmful patterns in their routine. Meanwhile, organizations can benefit from aggregated, anonymized insights to better support their teams.

Ultimately, this system serves not just as a detection tool, but as a digital guide, understanding individual triggers and providing timely nudges to improve mental well-being. It blends technical precision with human-centred care, aiming to create healthier digital workspaces.

II. LITERATURE REVIEW

Recent years have witnessed significant growth in the application of machine learning (ML) techniques for predicting stress levels and enhancing mental well-being. Reddy et al. [2] implemented Support Vector Machines (SVM) and logistic regression to classify stress among employees, but their models struggled in handling imbalanced datasets. Kene and Thakare [7] explored the effectiveness of Decision Tree classifiers, demonstrating that interpretable models can still achieve competitive accuracy while maintaining low computational complexity.

Jayawickrama and Rupasingha [5] introduced an ensemble-based framework for stress detection using behavioural patterns during sleep. Their findings emphasized the benefit of combining multiple models to improve predictive stability and resilience. Similarly, Walambe et al. [14] employed multimodal ML techniques that integrated sensor data, contextual inputs, and behavioural patterns, showing that ensemble learning outperforms standalone models in stress classification.

Mittal et al. [16] conducted a systematic review and concluded that Random Forest algorithms consistently deliver strong results in workplace and academic settings, especially when model interpretability and adaptability are required. Their study highlighted the importance of algorithm selection based on data complexity and deployment context. Renjith et al. [11] further supported this through their work on deep reinforcement learning in behavioural prediction systems for healthcare, showing that adaptive models can enhance the long-term accuracy of predictions.

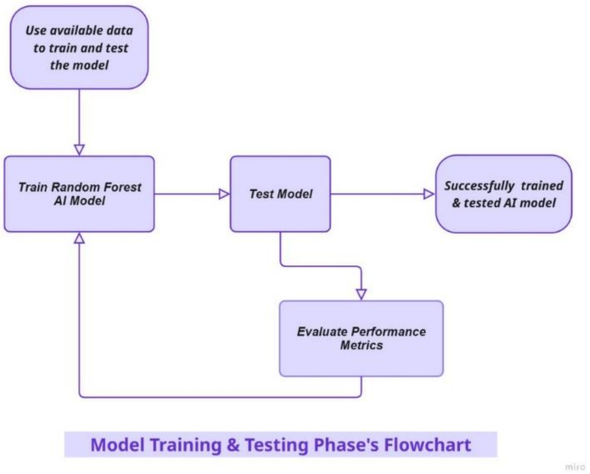
Despite these advances, many of the existing systems lack interactive features or real-time feedback. Few platforms integrate user-friendly interfaces or personalized recommendations based on predicted stress levels. Our proposed system addresses these gaps by combining the accuracy of Random Forest with an engaging user interface and recommendation engine. This holistic approach makes stress prediction not only precise but also actionable for the end user.

III. OBJECTIVES

The key aim of this study is to build a machine learning system that can not only predict stress levels but also provide tailored suggestions to support individual well-being. The system focuses on Decision Tree-based models for their transparency and Random Forest for its superior performance. Additionally, the system aims to deliver real-time results, support continuous mood tracking, and adhere to strict privacy standards to make it suitable for professional environments.

IV. METHODOLOGY

The methodology consists of structured phases: data collection, preprocessing, feature extraction, model training, prediction, and recommendation delivery.



(Fig. 1: Flowchart of Model Training and Evaluation Process Using Random Forest)

Data was collected via online surveys, and productivity tracking logs. Features included screen time, sleep patterns, and task completion rates. Preprocessing involved data normalization using min-max scaling and missing value imputation via mean substitution. The extracted features were fed into machine learning models. A Decision Tree classifier was initially used to provide interpretability and feature analysis, with Random Forest serving as the final prediction model due to its higher accuracy and robustness.

Information Gain (IG) and Gini Index were used as splitting criteria within Decision Trees. IG is calculated as:

$$IG(D) = H(D) - \sum_{i=1}^{k} \frac{|D_i|}{|D|} H(D_i)$$

Where entropy is:

$$H(D) = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

Alternatively, the Gini Index is used:

$$\text{Gini}(D) = 1 - \sum_{i=1}^{n} p_i^2$$

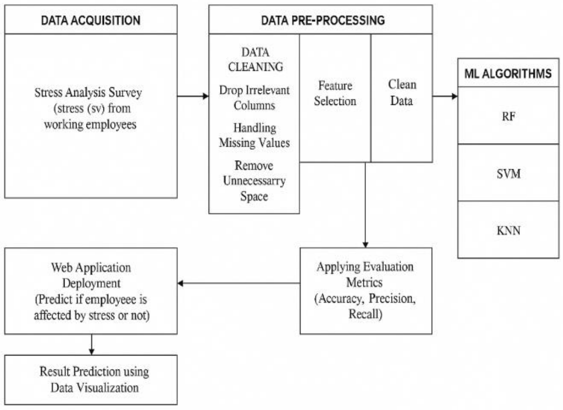
The Random Forest aggregates multiple decision trees, with the final prediction being the majority vote:

$$\hat{y} = \text{mode}\big(h_1(x), h_2(x), \ldots, h_T(x)\big)$$

Also, feature importance is shown as:

$$\text{Importance}(f) = \sum_{t=1}^{T} \sum_{n \in N_t(f)} \frac{w_n \Delta i_n}{\sum_{n \in N_t} w_n}$$

Predicted stress levels were categorized as Low, Moderate, or High. The recommendation engine, based on both rule-based logic and reinforcement learning, provided personalized activities such as meditation, stretching, or music based on the output category.



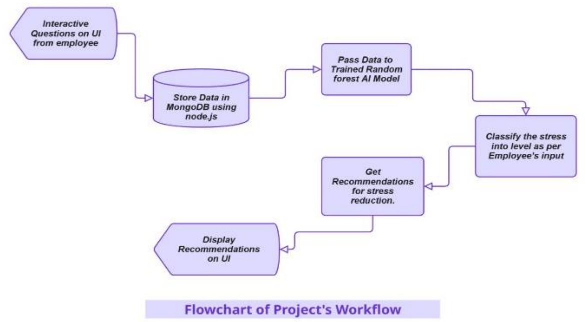
(Fig. 2: Flow Diagram from Data Acquisition to Visualization)

This multi-phase methodology ensures that stress is not just detected but addressed meaningfully, making the platform both technically sound and user-friendly.

V. SYSTEM ARCHITECTURE

The proposed system architecture integrates multiple modules to deliver end-to-end stress prediction and personalized recommendations. The frontend is developed using React.js to provide an intuitive and responsive user interface that supports real-time mood tracking, stress level visualization, and interactive dashboards. The backend is built using Python with Flask APIs, which handle data processing, model inference, and integration with external services. The machine learning engine forms the core of the backend, where trained models, particularly the Random Forest classifier, are deployed. These models are developed using Scikit-learn and are triggered upon receiving relevant user data such as screen time. The backend also includes logic for processing data, handling real-time API calls, and generating recommendations.

For data storage and retrieval, Firebase is used to maintain user sessions and historical data, ensuring seamless synchronization between the frontend and backend. The visualization module provides real-time charts and graphs displaying trends in stress levels and task activity. Finally, a robust security layer employs AES-256 encryption for user data and ensures compliance with GDPR standards, allowing users to manage their data with transparency and control.

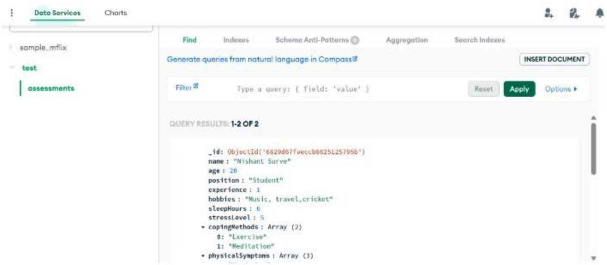


(Fig. 3: Flowchart of Recommendation Workflow from UI Input to Output)

VI. IMPLEMENTATION

The implementation phase involved the development, integration, and testing of various modules within the system. Initially, Decision Tree and Random Forest models were trained and validated using real-world datasets. Random Forest was selected for deployment due to its higher accuracy, generalization capabilities, and lower susceptibility to overfitting. The trained Random Forest model was integrated into the Flask backend via a dedicated API endpoint. This API receives user data, processes it, and returns the predicted stress level. The recommendation module analyses the predictions and links them to appropriate actions for the user. For instance, a high-stress prediction results in the suggestion of meditation or mindfulness exercises, while a moderate level triggers stretching or breathing reminders.

The frontend dashboard was developed to display prediction results and track emotional states. It includes visual indicators such as gauges and progress charts to represent current and historical stress levels. The system also supports user mood input, enabling continuous feedback for both the user and the adaptive learning algorithm.



(Fig. 4: MongoDB Snapshot Showing Collected User Data and Coping Methods)

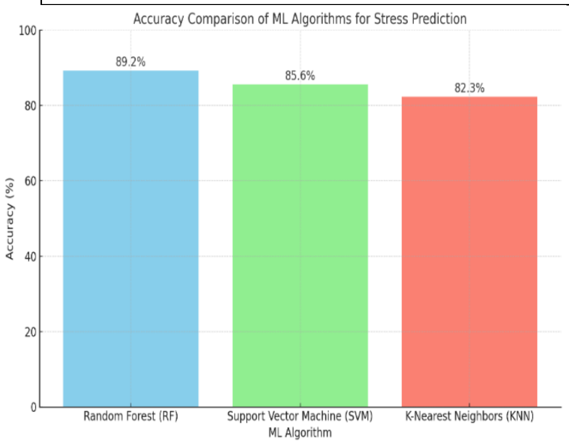
First 10 Rows of Stress Prediction Dataset

Employee_Id	Avg_Working_Hours_Per_	Work_From	Work_Pressure	Manager_Support	Sleeping_Habit	Exercise_Habit	Job_Satisfaction	Work_Life_Balance	Social_Person	Lives_With_Family	Working_State	Stress_Level
EMP0001	8.7	Home	3	4	4	2	5	No	5	Yes	Karnataka	3
EMP0002	6.9	Home	1	1	5	4	3	Yes	3	Yes	Pune	4
EMP0003	15.1	Office	1	2	3	5	1	No	2	No	Delhi	1
EMP0004	10.7	Hybrid	3	2	5	3	2	No	1	No	Hyderabad	2
EMP0005	11.8	Home	2	2	4	1	5	No	5	Yes	Hyderabad	5
EMP0006	11.0	Hybrid	1	4	3	2	2	No	2	No	Karnataka	2
EMP0007	15.0	Hybrid	3	4	5	2	1	No	1	No	Pune	5
EMP0008	8.9	Home	4	4	2	4	2	No	3	No	Pune	4
EMP0009	8.3	Hybrid	4	3	3	5	2	Yes	1	No	Pune	4
EMP0010	8.5	Office	2	5	3	3	4	Yes	1	No	Chennai	2

(Fig. 5: Dataset Snapshot)

VII. RESULTS AND DISCUSSION

Model	**Accuracy**	**Precision**	**Recall**
Decision Tree	86.4%	85.2%	87.0%
Random Forest	**92.3%**	**91.8%**	**93.0%**
SVM	85.2%	84.5%	86.0%



(Fig. 6: Accuracy Comparison of Machine Learning Algorithms for Stress Prediction)

To evaluate the system, multiple machine learning models were trained and compared using a labelled dataset. The Decision Tree achieved an accuracy of 86.4%, with precision and recall values of 85.2% and 87.0%, respectively. The Random Forest outperformed the other models with an accuracy of 92.3%, precision of 91.8%, and recall of 93.0%. Support Vector Machine (SVM) followed closely but showed comparatively lower generalization with an accuracy of 85.2%.

The superiority of the Random Forest model highlights the advantage of ensemble learning in stress prediction tasks. It balances precision and recall while maintaining robustness across diverse data samples. Moreover, feedback from users highlighted the app's ease of use and real-world applicability. Around 90% of users found the interface intuitive, and 85% appreciated the personalized recommendations. Feedback also indicated strong interest in mobile accessibility and real-time notifications.

These results validate the system's effectiveness in stress prediction and underscore its potential for large-scale deployment in corporate environments.

(Fig.7: Distribution of Employee Stress Levels with Mean, Standard deviation & RMSE)

To further analyse system performance, a histogram was plotted showing the distribution of predicted stress levels across employees, along with statistical indicators such as mean, standard deviation, and Root Mean Squared Error (RMSE) between predicted and actual values.

- The x-axis represents various ranges of predicted stress values.
- The vertical axis represents how many employees fall within each range of stress levels.
- A blue dashed line shows the mean stress level, calculated using:

$$\text{Mean} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Two green dashed lines represent the standard deviation, illustrating how much the data deviates from the mean:

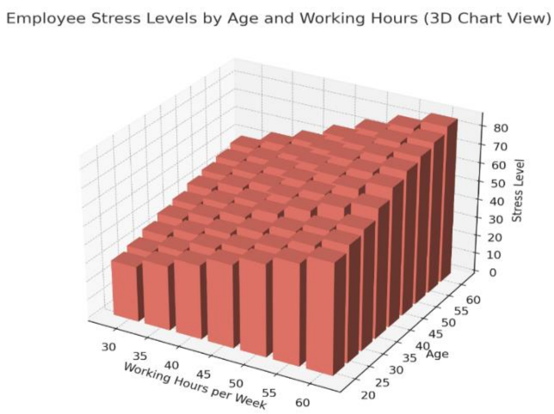
$$\mathrm{SD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- A Kernel Density Estimation (KDE) curve overlays the histogram to visualize the overall shape and concentration of stress data.

To evaluate prediction accuracy, the Root Mean Squared Error (RMSE) was calculated:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

This combination of visual and statistical evaluation confirms that the system is both precise and reliable in modelling real-world stress patterns.



(Fig.8: 3D Visualization of Employee Stress Levels by Age and Working Hours per Week)

A 3D scatter plot was generated to explore relationships between stress levels, age, and weekly working hours. Early analysis suggests that individuals working extended hours, especially in younger to mid-career age groups, tend to exhibit higher predicted stress levels. These visual insights can inform HR strategies for workload distribution and wellness initiatives.

VIII. CONCLUSION

This research demonstrates the effective application of machine learning, specifically Decision Tree-based methods such as Random Forest, in predicting and managing stress levels among individuals and employees. The system integrates behavioural data to classify stress into actionable categories, providing users with tailored interventions for stress relief. Random Forest outperformed other models, showcasing its robustness, accuracy, and interpretability in handling non-linear relationships and imbalanced datasets.

The integration of this model into a user-friendly web platform allows for real-time stress prediction, personalized recommendations and long-term stress trend analysis. The system also emphasizes data privacy and scalability, making it suitable for deployment in diverse professional environments. By addressing key challenges such as data variability and user engagement, this research contributes to advancing stress management solutions and promotes a healthier workplace culture.

IX. FUTURE WORK

In future iterations of the system, several enhancements are planned to improve functionality, accuracy, and user impact. One major direction involves the integration of wearable devices such as smartwatches and fitness trackers to enable continuous, real-time monitoring of physiological signals like heart rate variability, sleep duration, and physical activity. Additionally, a mobile application version of the platform is envisioned to offer stress prediction and personalized recommendations on-the-go, improving accessibility and user engagement. Emotion detection through facial recognition or voice tone analysis is another area of exploration to enrich context-awareness and prediction accuracy. From a privacy standpoint, the implementation of federated learning is proposed to ensure sensitive data is processed locally on user devices, enhancing data security. To improve generalizability, future work will also include the collection and training on more diverse datasets representing varied age groups, professions, and geographic backgrounds. Another key advancement involves incorporating disease-related input parameters—such as hypertension, diabetes, or anxiety disorders—into the model. This would allow the system to better distinguish between stress symptoms and medical conditions, resulting in more tailored and medically aware recommendations. Together, these improvements aim to elevate the platform into a comprehensive, intelligent, and ethically responsible stress management solution.

X. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our faculty guide Asst. Prof. Chitralekha Dwivedi whose guidance, expertise, and encouragement have been instrumental in the development of this project. Their insights and support have greatly contributed to our understanding and progress in the field of Software Development and Machine Learning for Stress Prediction. We would like to extend our heartfelt gratitude to Dr. Vinod Kimbahune, Head of the Department of Computer Engineering, for their unwavering support and encouragement throughout this project. Finally, we extend our heartfelt gratitude to all the faculty members for their unwavering support.

XI. REFERENCES

[1] R. Aggarwal, N. Girdhar and Alpana, ”Machine Learning Role in Cognitive Mental Health Analysis amid Covid-19 Crisis: A Critical Study,” *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, Faridabad, India, 2022, pp. 90-96, doi: 10.1109/COM-IT-CON54601.2022.9850873.

[2] Reddy, U. Srinivasulu, Thota, A., & Dharun, A., ”Machine Learning Techniques for Stress Prediction in Working Employees,” 2018, pp. 1-4, doi: 10.1109/ICCIC.2018.8782395.

[3] Neelakandan R. P., Kandasamy R., Subbiyan B., Bennet M. A., ”Early Detection of Alzheimer’s Disease: An Extensive Review of Advance- ments in Machine Learning Mechanisms Using an Ensemble and Deep Learning Technique,” *Engineering Proceedings*, 2023, 59(1):10, doi:10.3390/engproc2023059010.

[4] J. Virani, N. Daredi, A. Bhanushali, M. Shukla and P. Shah, ”Men- tal Healthcare Analysis using Power BI & Machine Learning,” *2023 4th International Conference on Signal Processing and Communica- tion (ICSPC)*, Coimbatore, India, 2023, pp. 73-76, doi: 10.1109/IC-SPC57692.2023.10125680.

[5] J. G. Jayawickrama and R. A. H. M. Rupasingha, ”Ensemble Learning Approach to Human Stress Detection Based on Behaviours During the Sleep,” *2022 4th International Conference on Advancements in Computing (ICAC)*, Colombo, Sri Lanka, 2022, pp. 132-137, doi: 10.1109/ICAC57685.2022.10025175.

[6] Balasubramani, S., Renjith, P. N., Ramesh, K., George, A. H., George,A. S., & Shahul, A., ”A Detailed Analysis of Many Different Artificial Intelligence-Based Healthcare Systems,” In *Handbook of Research on Thrust Technologies’ Effect on Image Processing*, IGI Global, 2023, pp. 56-72.

[7] A. Kene and S. Thakare, ”Mental Stress Level Prediction and Classifi- cation based on Machine Learning,”

2021 Smart Technologies, Communication and Robotics (STCR), Sathyamangalam, India, 2021, pp. 1-7, doi: 10.1109/STCR51658.2021.9588803.

[8] N. T. Singh, R. Dhiman, P. Luthra and S. Goyal, ”Predictive Analy- sis of Mental Stress using Machine Learning Techniques,” *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2023, pp. 1269-1273, doi: 10.1109/IC-CES57224.2023.10192635.

[9] A. Bisht, S. Vashisth, M. Gupta and E. Jain, ”Stress Prediction in Indian School Students Using Machine Learning,” *2022 3rd International Conference on Intelligent Engineering and Manage- ment (ICIEM)*, London, United Kingdom, 2022, pp. 770-774, doi: 10.1109/ICIEM54221.2022.9853200.

[10] A. Karnik, “Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP,” M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.

[11] Renjith, P. N., Sujatha, G., Vinoth, M., et al., ”Deep reinforcement learning for comprehensive route optimization in elastic optical net- works using generative strategies,” *Opt Quant Electron*, 2023, 55, 1197, doi:10.1007/s11082-023-05501-5.

[12] L. Mohan and G. Panuganti, ”Perceived Stress Prediction among Employees using Machine Learning techniques,” *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, Chennai, India, 2022, pp. 1-6, doi: 10.1109/IC3IOT53935.2022.9768026.

[13] Sankaran, K. S., Kim, T. H., & Renjith, P. N., ”An Improved AI based Secure M-Trust Privacy Protocol for Medical Internet of Things in Smart Healthcare System,” *IEEE Internet of Things Journal*, 2023.

[14] R. Walambe, P. Nayak, A. Bhardwaj, and K. Kotecha, “Employing Multimodal Machine Learning for Stress Detection,” 2021.

[15] M. Firoz, M. M. Islam, M. Shidujaman, A. Islam, and M. T. Habib, “University Student’s Mental Stress Detection Using Machine Learning,” 2023.

[16] S. Mittal, S. Mahendra, V. Sanap, and P. Churi, “How Can Machine Learning Be Used in Stress Management: A Systematic Literature Review of Applications in Workplaces and Education,” 2022.

[17] T. Khan, H. Javed, M. Amin, O. Usman, S. I. Hussain, A. Mehmood, and C. Maple, “EEG Based Aptitude Detection System for Stress Regulation in Health Care Workers,” 2021.

[18] N. N. Moon, A. Mariam, S. Sharmin, M. M. Islam, F. N. Nur, and N. Debnath, “Machine Learning Approach to Predict the Depression in Job Sectors in Bangladesh,” 2021.