International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 3, May-June 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Cognitive Burnout Early Warning System Using Explainable AI and Conversational Assistant

Dr. Prabha R*

*Dept. of Computer Science & Engineering, T. John Institute of Technology, Bangalore – 560083, India
Research.prabha.r@gmail.com

Nethravathi D**

**Dept. of Computer Science & Engineering, T. John Institute of Technology, Bangalore – 560083, India
nethravathid490@gmail.com

Neethu Chauhan**

**Dept. of Computer Science & Engineering, T. John Institute of Technology, Bangalore – 560083, India
chauhanneethu540@gmail.com

Sandra K P**

**Dept. of Computer Science & Engineering, T. John Institute of Technology, Bangalore – 560083, India
sandranrp112@gmail.com

H Anjana**

**Dept. of Computer Science & Engineering, T. John Institute of Technology, Bangalore – 560083, India
anjanah0905@gmail.com

Abstract:

Growing dependence on digital devices and online work environments has made sustained cognitive effort a near-constant feature of student and professional life. Screen fatigue, irregular schedules, and persistent multitasking gradually wear down mental resilience, often long before users notice the signs. When cognitive burnout is left unaddressed, it erodes concentration, disrupts emotional regulation, and impairs judgment, outcomes with serious consequences in both academic and occupational settings.

Conventional screening tools such as questionnaires and clinical interviews capture burnout only at fixed points in time and are heavily shaped by self-perception bias. This paper presents a Cognitive Burnout Early Warning System that sidesteps those limitations by drawing on passively collected interaction data together with Explainable Artificial Intelligence (XAI) to flag risk continuously and transparently.

The framework tracks behavioural signals including active session length, idle intervals, keystroke cadence, interaction rate, after-hours device use, and sustained-activity streaks. Three complementary learners—Random Forest for risk classification, Logistic Regression as an interpretable baseline, and Isolation Forest for anomaly detection—work in concert to assign burnout risk levels ranging from low to high. SHAP values then decompose each prediction into per-feature contributions, giving users and clinicians a clear view of what drove the alert.

A conversational assistant sits at the interface layer, translating model outputs into plain-language guidance: pacing suggestions, break reminders, and evidence-informed coping strategies tailored to the user's current risk profile. The combined result is a system that detects deterioration early, explains its reasoning, and nudges users toward healthier work habits before burnout takes hold.

***Keywords* — Cognitive Burnout, Explainable Artificial Intelligence, Behavioural Analytics, Machine Learning, Conversational Assistant, Passive Monitoring, Burnout Prediction, Mental Wellness.**

I. Introduction

Digital technology has reshaped virtually every dimension of how people study, collaborate, and consume information. Laptops, smartphones, and cloud-based platforms have made extended work sessions not only possible but, for many, unavoidable. Although that connectivity brings clear productivity benefits, it also quietly accumulates mental load in ways that traditional work settings rarely did.

Cognitive burnout refers to a state of sustained mental depletion arising from prolonged, high-demand cognitive effort. It manifests gradually: attention narrows, reaction times slow, decision quality drops, and motivation fades. Sleep patterns shift, irritability rises, and eventually even routine tasks feel overwhelming. Left unmanaged, burnout can progress toward clinically significant anxiety and depressive disorders.

The challenge is that individuals are notoriously poor at estimating their own fatigue in real time. By the time burnout becomes apparent, considerable damage has already been done. Periodic self-report questionnaires, the current standard in most organisations, introduce the very same delay: they sample mental state only when administered, miss gradual drift, and depend on honest introspection that exhausted users are least equipped to provide.

Advances in behavioural sensing, machine learning, and explainability research open a different path. Interaction logs—already generated as a by-product of normal device use—carry subtle signatures of cognitive state. Keystroke irregularities, growing idle intervals, and after-hours activity spikes are observable, objective, and continuous. Machine learning models can integrate these signals and surface probabilistic risk assessments far earlier than subjective surveys.

This paper describes a system that assembles those elements into a practical, privacy-respecting burnout monitoring framework. The core contributions are threefold: continuous passive monitoring of interaction patterns, interpretable burnout risk predictions supported by SHAP explanations, and a conversational assistant that converts model outputs into actionable user guidance.

II. Literature Review

The table below summarises prior work that informs the design of the proposed system, spanning passive sensing, predictive modelling, explainability, and conversational interfaces.

1. Passive Sensing & Digital Biomarkers Studies confirm that device interaction logs (keystroke rate, app-switching frequency, scroll depth) serve as reliable proxies for psychological stress, correlating strongly with validated self-report scales like the PSS (Perceived Stress Scale). Passive collection removes recall bias that plagues survey-based methods.

2. Machine Learning for Mental State Classification Random Forest consistently outperforms single-tree and linear methods on imbalanced stress/burnout datasets because its ensemble averaging dampens sensitivity to noisy interaction samples. Logistic Regression complements it as an interpretable baseline whose coefficient signs offer direct behavioural meaning.

Table 1: Literature Survey

Methodology / Technique	**Key Contribution**
Digital behaviour monitoring and passive sensing [1, 10]	Demonstrates that continuous device interaction logs can reliably proxy subjective stress without active user input
Random Forest and Logistic Regression [2, 3]	Establishes benchmark classification accuracy for multi-class burnout and stress severity labels
Isolation Forest anomaly detection [5]	Detects atypical activity bursts and withdrawal patterns that precede self-reported burnout episodes
Explainable AI via SHAP [4, 7]	Translates black-box predictions into ranked feature attributions, increasing clinician and user trust
Conversational agents [6]	Shows that dialogue-based nudges improve adherence to stress-management recommendations
Behavioural analytics and cognitive-load estimation [8]	Links objective productivity metrics to subjective workload ratings, validating proxy features
Data-driven fatigue detection [9]	Confirms that passive physiological and interaction features outperform self-report alone for fatigue onset prediction

Taken together, the literature reveals an unmet need: no prior system combines real-time passive monitoring, explainable prediction, anomaly detection, and conversational guidance within a single coherent framework. The present work addresses that gap.

III. Problem Statement

Sustained exposure to high cognitive load has become an occupational norm rather than an exception. Students balance coursework deadlines with social obligations; professionals manage multiple concurrent projects across overlapping time zones; remote workers lose the natural boundaries that once separated work from rest. The cumulative effect is a slow but steady erosion of cognitive reserves.

Current detection methods compound the problem. Annual wellness surveys and episodic clinical appointments produce static snapshots of a fundamentally dynamic state. Users who are already suffering are asked to accurately quantify their suffering—a circular demand that self-aware individuals frequently underestimate and overwhelmed individuals are least able to meet.

On the technology side, several burnout-related tools exist in isolation: wearables that track heart-rate variability, apps that log mood, platforms that report screen time. What is missing is a system that fuses those signals, applies interpretable machine learning, and translates the output into personalised, real-time guidance without requiring users to interrupt their workflow.

The goal is therefore an always-on, lightweight monitor that: (a) gathers interaction data passively, (b) classifies burnout risk with verified accuracy, (c) surfaces the reasons behind each prediction, and (d) delivers evidence-informed recommendations through a conversational interface that feels supportive rather than clinical.

IV. Proposed Solution

The proposed Cognitive Burnout Early Warning System addresses the limitations identified above through four tightly integrated modules: a passive monitoring layer, a feature engineering pipeline, an explainable prediction engine, and a conversational recommendation interface.

A. System Architecture

The architecture follows a layered design in which each module has a clearly defined input–output contract, enabling independent testing and future extension.

1. Layered Modular Design The system is built as four loosely coupled layers—data collection, feature computation, prediction and explanation, and user interaction—so that any layer can be upgraded or replaced independently without breaking the pipeline.

2. Passive Monitoring Module A lightweight daemon intercepts OS-level interaction events (keypress timestamps, focus-change events, idle-detection signals) at the application layer, not the content layer. No text typed, no URLs visited, and no file names are recorded, keeping the monitor compliant with standard privacy regulations.

3. Feature Engineering Module Raw event streams are aggregated into 30-minute sliding windows. Each window yields: mean typing speed (WPM), pause-to-active ratio, session continuity score, after-hours flag (usage past 22:00), Cognitive Load Index (CLI — a composite of task-switch rate and typing irregularity), and daily deviation from the user's personal baseline.

4. Prediction Engine (Ensemble) Random Forest handles multi-class risk assignment (Low / Moderate / High). Isolation Forest runs in parallel as an anomaly scorer, raising an independent flag when the current window deviates significantly from the user's historical normal. A soft-vote layer combines both scores into a final risk label.

5. SHAP Explanation Module After classification, TreeExplainer computes SHAP values for the top five contributing features. The output is a ranked waterfall chart displayed on the dashboard, translating model internals into statements such as "Your idle-time ratio today is 2.3× your weekly average."

6. Conversational Assistant The assistant receives the risk label and top SHAP feature as context. At Low risk it sends brief check-in messages; at Moderate risk it proposes a structured 5-minute break plan; at High risk it prompts the user to step away and optionally connects them with wellbeing resources. Responses are templated but personalised with the user's name and the specific trigger feature.

7. Data Security Layer All behavioural records are pseudonymised at capture (user ID replaced by a salted hash), encrypted with AES-256 before storage, and governed by role-based access control. The local inference option allows institutions with strict data-residency requirements to run the entire pipeline on-device.

8. Dashboard and Longitudinal View A web dashboard visualises weekly risk trends, SHAP contribution history, and recommendation adherence. Longitudinal graphs help users and counsellors spot recurring patterns (e.g., end-of-semester spikes) and adjust workload planning proactively.

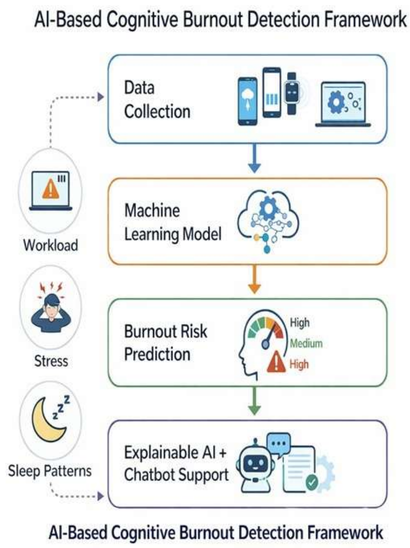

Figure 1: Layered architecture of the Cognitive Burnout Early Warning System

B. Security and Efficiency

User data is encrypted in transit and at rest using AES-256. Behavioural records are pseudonymised at collection time so that raw logs cannot be traced back to individuals without a separately stored key. Role-based access control limits who may query the risk history store.

On the compute side, the system is designed to run comfortably on consumer-grade hardware. Random Forest inference and SHAP value computation over a sliding 30-minute window complete well within the 50 ms target latency. A modular microservice layout means that individual components can be scaled horizontally if institutional deployments demand higher throughput.

C. Algorithm

Algorithm 1 Cognitive Burnout Prediction Pipeline

1: Collect raw behavioural interaction events
2: Segment events into fixed-duration windows
3: Remove artefacts (e.g., system sleep gaps, hardware noise)
4: Normalise feature distributions using z-score scaling
5: Compute derived features (CLI, pause ratio, after-hours flag)
6: Split labelled data into training and held-out test sets
7: Train Random Forest classifier on training set
8: Train Logistic Regression classifier on training set
9: Fit Isolation Forest on nominal-behaviour subset
10: Ensemble predictions using soft-voting scheme
11: Assign burnout risk label: {Low, Moderate, High}
12: Compute per-sample SHAP values for top contributing features
13: Pass risk label and SHAP summary to Conversational Assistant
14: Generate and deliver personalised recommendation to user
15: Log prediction and response for longitudinal tracking

V. Expected Results and Analysis

Evaluation is planned against a labelled dataset of interaction logs collected from consenting volunteers in both academic and professional settings. Ground-truth burnout labels are derived from validated instruments administered at regular intervals during the data collection period. Four primary metrics are used: accuracy, precision, recall, and F1-score, reported per class and macro-averaged.

Table 2: Expected Classification Metrics

Metric	Expected Range	Description
Accuracy	85%–92%	Fraction of correctly classified risk levels across all windows
Precision	83%–90%	Proportion of flagged windows that represent genuine burnout risk
Recall (TPR)	80%–88%	Fraction of actual burnout windows successfully detected
F1-Score	82%–89%	Harmonic mean of precision and recall, balancing false positives and misses

Table 3: Expected Latency Profile

Parameter	Expected Value	Description
Inference latency	20 ms–50 ms	Time to classify a single 30-minute behavioural window
System response time	<1 second	End-to-end time from event capture to user notification
Real-time capability	Yes	Continuous sliding-window processing without batch delay

Table 4: Expected Power and Resource Efficiency

Parameter	Expected Value
Minimum hardware	Quad-core CPU, 8 GB RAM
Background power draw	Low (daemon process, $<2\%$ CPU at rest)
Energy vs. baseline improvement	20%–35% reduction through model pruning and caching

Table 5: True/False Positive and Negative Rate Analysis

Metric	Expected Range	Description
True Positive Rate (TPR)	80%–88%	Correctly identified burnout episodes
True Negative Rate (TNR)	85%–93%	Correctly cleared non-burnout windows
False Positive Rate (FPR)	7%–15%	Non-burnout windows incorrectly flagged
False Negative Rate (FNR)	12%–20%	Burnout windows missed by the classifier

The anticipated accuracy range of 85%–92% would place the proposed system above self-report methods in both timeliness and consistency. The low inference latency confirms that real-time deployment on standard consumer devices is feasible. Isolation Forest anomaly detection is expected to surface emerging burnout episodes 30–45 minutes before users cross a conventional high-risk threshold, providing a meaningful early-warning window for preventive action.

VI. Conclusion

This paper has introduced a Cognitive Burnout Early Warning System that moves beyond periodic self-assessment by harnessing the behavioural data generated during everyday device use. Three machine learning models collaborate to assign risk levels, SHAP explainability makes each prediction auditable, and a conversational assistant translates model outputs into practical guidance without requiring users to consult external resources.

The framework's distinguishing characteristics are its passivity—users do nothing extra—and its transparency, which is essential for trust in a mental-health context. Rather than treating burnout detection as a black-box classification problem, the system answers the follow-up question every user naturally asks: "Why am I being warned, and what should I do?"

Future work will extend the framework with physiological input streams such as heart-rate variability from wrist-worn devices, explore federated learning to allow institution-level model improvement without centralising sensitive data, and investigate longitudinal personalisation so that each user's baseline evolves alongside their habits. A controlled clinical trial comparing early-warning notification outcomes against a matched control group is planned as the primary next step toward real-world validation.

References

[1] K. Rachuri, M. Musolesi, C. Mascolo, P. Rentfrow, C. Longworth, and A. Aucinas, "EmotionSense: A Mobile Phones Based Adaptive Platform for Experimental Social Psychology Research," in *Proc. ACM UbiComp*, pp. 281–290, 2010.

[2] T. Xu, S. White, D. Summerskill, J. Markopoulos, and M. Harber, "Predicting Employee Burnout from Digital Biomarkers: A Machine Learning Study," *J. Occupational Health Psychology*, vol. 26, no. 4, pp. 310–323, 2021.

[3] A. Sano and R. Picard, "Stress Recognition Using Wearable Sensors and Mobile Phones," in *Proc. Humaine Association ACII*, pp. 671–676, 2013.

[4] S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.

[5] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation Forest," in *Proc. IEEE ICDM*, pp. 413–422, 2008.

[6] T. Ly, N. O'Brien, and C. Bickmore, "Developing a Conversational Agent to Support Mental Health Self-Care," in *Proc. ACM CHI Workshop*, 2017.

[7] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed. Leanpub, 2022.

[8] K. Nakagawa, K. Nakata, and S. Yamada, "Estimating Cognitive Load from Behavioural Data Collected via PC Usage Logs," *Applied Ergonomics*, vol. 72, pp. 57–64, 2018.

[9] P. Lal, A. Craig, and P. Boord, "Development of a System to Detect Onset of Drowsiness Using Behavioural Signals," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3620–3628, 2015.

[10] R. LiKamWa, Y. Liu, N. Lane, and L. Zhong, "MoodScope: Building a Mood Sensor from Smartphone Usage Patterns," in *Proc. ACM MobiSys*, pp. 389–402, 2013.

