BEAM: Behaviour and Emotion-Aware Multimodal Framework for Digital Wellness

Ashly P Eldho[1], Nishanth R[2]
[1] *Department of Computer Science and Engineering PES University* Bangalore, India
[2] *M.E., M.B.A.(Ph.D), Senior Data Scientist Great Learning*Chennai, Tamilnadu, India
doi.org/10.64643/IJIRTV13I3-207580-459

***Abstract*—Excessive smartphone usage has become a growing concern due to its impact on mental wellbeing, sleep quality, productivity, and daily life. Existing digital wellbeing applications primarily focus on screen-time monitoring and often pro-vide limited support for personalised intervention and long-term behavioural assessment. To address these limitations, this paper presents BEAM, a multimodal and agentic AI framework for personalised digital wellness. The proposed framework com-bines behavioural analytics and facial emotion recognition to as-sess user wellbeing. Behavioural indicators, including screen time, phone unlock frequency, social media usage, sleep dura-tion, stress level, anxiety score, physical activity, and family in-teraction, are analysed using a Logistic Regression model to pre-dict addiction levels. Emotional context is obtained through an ensemble emotion recognition framework integrating a Convolutional Neural Network (CNN) and DeepFace. The system fur-ther incorporates Retrieval-Augmented Generation (RAG), persistent memory, and a LangGraph-based Agentic AI layer to generate personalised recommendations, analyse historical behavioural patterns, and provide contextual wellness insights. Experimental evaluation demonstrates that the behavioural prediction model achieves an accuracy of 91%, while the emo-tion recognition framework achieved approximately 77% grouped emotion classification accuracy and provides meaning-ful emotional context for recommendation generation. The integration of memory, knowledge retrieval, and autonomous reasoning enables personalised and longitudinal wellbeing assessment. The results indicate that BEAM provides a more comprehensive approach to digital wellness management than conventional screen-time monitoring systems and demonstrates the potential of multimodal and agentic AI techniques for personalised wellbeing support.**

***Index Terms*—Digital Wellness, Agentic AI, Multimodal Learning, Emotion Recognition, Behavioural Analytics, Retrieval-Augmented Generation.**

I. INTRODUCTION

The widespread adoption of smartphones has transformed the way individuals communicate, work, learn, and access in-formation. Although mobile technologies provide significant benefits in daily life, excessive and uncontrolled smartphone usage has been associated with several negative outcomes, including reduced productivity, sleep disturbances, increased stress, social isolation, and behavioural addiction [1], [4], [6], [10]. Recent studies have highlighted the growing need for digital wellbeing solutions capable of helping users maintain a healthier balance between technology usage and overall wellbeing [5], [8]. Several digital wellbeing applications have been intro-duced to address excessive smartphone usage through screen-time monitoring, application usage statistics, and usage re-strictions [5], [7], [8]. While these solutions provide useful behavioural information, they often operate as passive monitoring tools and offer limited support for personalised intervention. Most existing systems rely primarily on screen-time statistics and fail to consider important contextual factors such as emotional state, psychological wellbeing, and historical behavioural patterns. As a result, the recommendations generated by these systems may not adequately reflect the user's actual wellbeing condition.

Recent advances in Artificial Intelligence (AI), Machine Learning (ML), Large Language Models (LLMs), and affective computing present new opportunities for developing intelligent digital wellness systems [2], [3]. Behavioural analytics can be used to identify patterns associated with prob-lematic smartphone usage, while emotion recognition techniques can provide additional insight into the user's psycho-logical state [1], [2], [3]. Similarly, AIpowered recommendation systems can support the delivery of personalised guidance tailored to individual wellbeing needs [3], [8].

To address the limitations of existing approaches, this pa-per presents BEAM, a multimodal and agentic AI framework for personalised digital wellness. The proposed framework integrates behavioural addiction prediction, facial emotion recognition, alert computation, recommendation generation, persistent memory, Retrieval-Augmented Generation (RAG), and autonomous reasoning within a unified platform. Behavioural information and emotional context are analysed together to provide a more comprehensive assessment of user wellbeing. Historical session data are preserved through a persistent memory layer, enabling long-term behavioural monitoring and personalised insight generation.

The proposed framework combines machine learning, deep learning, large language models, and agentic AI tech-niques to move beyond traditional screen-time monitoring systems and provide intelligent, context-aware wellbeing as-sistance. Experimental evaluation demonstrates the effective-ness of the framework in supporting behavioural assessment, personalised recommendation generation, and longitudinal wellbeing monitoring.

The major contributions of this work are summarized as fol-lows:

- Development of a multimodal digital wellness framework that integrates behavioural analytics and facial emotion recognition.
- Implementation of a behavioural addiction prediction model using machine learning techniques for screen-time risk assessment.
- Development of an ensemble emotion recognition framework combining CNN-based classification and DeepFace analysis.
- Integration of Retrieval-Augmented Generation for knowledge-grounded recommendation generation.
- Implementation of a persistent memory system for long-term behavioural tracking and historical analysis.
- Development of a LangGraph-based Agentic AI framework capable of contextual reasoning and per-sonalised insight generation.
- Design of an intervention and reporting framework for proactive digital wellness management.

II. LITERATURE REVIEW

A. Digital Wellbeing Applications

Digital wellbeing has emerged as an important research area due to the increasing impact of excessive smartphone us-age on mental health, sleep quality, productivity, and social wellbeing. Several digital wellbeing applications have been developed to monitor screen time and support self-regulation; however, most provide limited personalised guidance and contextual understanding of user behaviour [5], [7], [8].

B. Behavioural Addiction Prediction

Machine learning techniques have been widely applied to predict smartphone addiction and problematic mobile usage patterns. Previous studies have shown that behavioural and psychological indicators can be effectively utilised for addic-tion classification and risk assessment [1], [6], [12]. These approaches demonstrate the usefulness of machine learning for behavioural analysis but generally focus only on usage-related features.

C. Emotion Recognition for Wellbeing Assessment

Emotion recognition has gained significant attention as a means of understanding user wellbeing and psychological state. Advances in facial emotion recognition and affective computing have enabled the identification of emotions from facial expressions, providing valuable information for wellbeing assessment [2], [3]. The FER2013 dataset continues to be one of the most commonly used benchmarks for emotion recognition research [17].

D. AI-Based Behaviour Analysis

Recent developments in AI-driven behaviour analysis and smartphone-based sensing have further demonstrated the potential of intelligent systems for personalised wellbeing sup-port [9], [10], [14]. These studies highlight the importance of combining behavioural information with advanced AI tech-niques to improve user understanding and engagement.

E. Research Gap

Despite these advances, limited research has explored the integration of behavioural analytics, emotion recognition, persistent memory, Retrieval-

Augmented Generation, and Agentic AI within a unified digital wellness framework. To address this gap, the proposed BEAM framework combines these components to provide personalised, context-aware, and long-term digital wellness support.

TABLE I.COMPARISON OF BEAM WITH EXISTING APPROACHES

Capability	Addiction Prediction Studies [1], [6], [12]	Digital Wellbeing Applica-tions [5], [7], [8]	BEAM
Behavioural Addiction Pre-diction	Yes	Limited	Yes
Emotion Recognition	No	No	Yes
Personalised Recommenda-tions	Limited	Limited	Yes
Context Aware Decision Mak-ing	No	No	Yes
Retrieval Au-gumented Gen-eration	No	No	Yes
Persistent Memory	No	No	Yes
Historical Be-haviour Analy-sis	Limited	No	Yes
Agentic AI Reasoning	No	No	Yes
Intervention Detection	No	Limited	Yes
Wellness Re-porting	No	Limited	Yes

Table I. compares the capabilities of BEAM with representative smartphone addiction prediction studies and existing digital wellbeing applications. While previous approaches primarily focus on behavioural prediction or screen-time monitoring, they generally lack emotional awareness, long-term memory, knowledge-grounded reasoning, and autonomous decision-making capabilities. The proposed BEAM framework addresses these limitations by integrating behavioural analytics, emotion recognition, RetrievalAugmented Generation, persistent memory, and Agentic AI within a unified digital wellness platform.

III. METHODOLOGY

A. Proposed BEAM Architecture

The BEAM framework integrates behavioural analytics, facial emotion recognition, Retrieval-Augmented Generation, persistent memory, and Agentic AI within a unified digital wellness platform. The architecture is designed to provide personalised wellbeing assessment and intervention by combining information obtained from both behavioural and emotional modalities.

The framework operates in two phases. The first phase performs behavioural addiction prediction and emotion recognition. Behavioural indicators such as screen time, phone unlock frequency, social media usage, sleep duration, stress level, anxiety score, physical activity, and family inter-action are analysed using a machine learning model to deter-mine the user's addiction level [1], [6], [12]. Simultaneously, facial images are processed through an emotion recognition framework to identify the user's emotional state [2], [3].

The outputs generated by these modules are combined to compute an alert level that reflects the overall wellbeing condition of the user. Based on the alert level and emotional con-text, personalised recommendations are generated using a Large Language Model [18].

The second phase extends the system through Agentic AI and long-term behavioural monitoring. User sessions are stored within a persistent memory layer, while a Retrieval-Augmented Generation framework retrieves relevant well-ness knowledge from a vector database. A LangGraph-based agent utilises memory retrieval, knowledge retrieval, and intervention analysis tools to generate personalised insights and recommendations. The framework further supports weekly reporting, historical trend analysis, and wellness ranking generation.

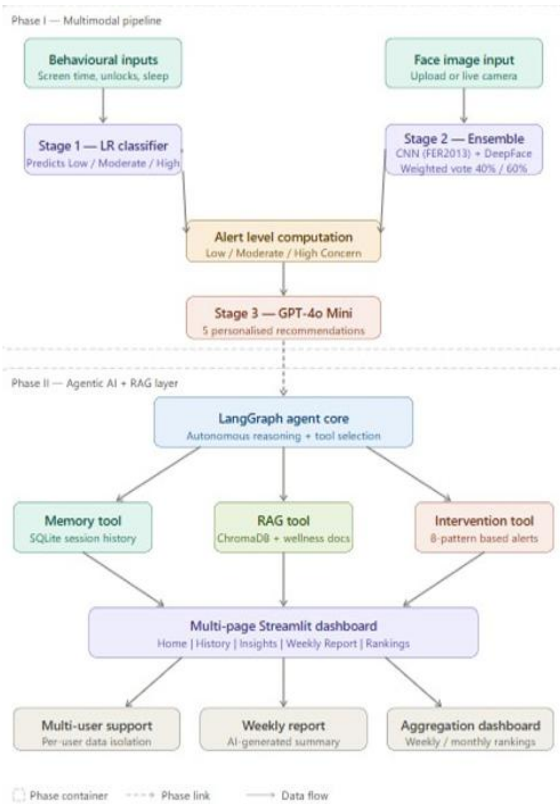

Fig. 1. *Architecture of Proposed BEAM Framework*

Fig. 1. illustrates the overall architecture of the BEAM framework, integrating behavioural prediction, emotion recognition, recommendation generation, persistent memory, Retrieval-Augmented Generation, and Agentic AI within a unified digital wellness platform. The framework combines behavioural and emotional information to provide personal-ised recommendations, long-term wellbeing monitoring, and intelligent intervention support.

B. Behavioural Prediction Framework

The behavioural prediction module is responsible for identifying the level of smartphone addiction based on behavioural and wellbeing-related indicators. A behavioural dataset containing screen-time characteristics, lifestyle factors, and psychological indicators was utilised for model development [1], [6], [12].

The input features include:

- Daily screen time
- Phone unlock frequency
- Social media usage
- Sleep duration
- Physical activity
- Family interaction time
- Stress level
- Anxiety score

Several machine learning algorithms were evaluated during model development, including Logistic Regression, Ran-dom Forest, Decision Tree, Support Vector Machine, XGBoost, and Gradient Boosting models. Experimental evaluation indicated that Logistic Regression achieved the high-est overall performance and was therefore selected as the fi-nal prediction model.

The model classifies users into three addiction categories:

- Low Addiction
- Moderate Addiction
- High Addiction

The predicted addiction level forms one of the primary inputs for alert computation and recommendation generation. The behavioural prediction module employs Logistic Regression to estimate the probability of a user belonging to a particular addiction category based on behavioural and psycho-logical indicators. The probability prediction is computed using the logistic function:

$$P(y=1|X) = \frac{1}{1+e^{-(\beta_0+\sum_{i=1}^{n}\beta_i x_i)}}$$

where X represents the feature vector containing behavioural indicators, β_0 denotes the intercept term, and β_i represents the coefficient associated with the corresponding feature $x_{i.}$ The model learns these coefficients during training and utilises them to estimate addiction risk levels.

C. Emotion Recognition Framework

To incorporate emotional context into the wellbeing assessment process, BEAM employs a facial emotion recognition framework based on an ensemble approach. The frame-work combines a custom Convolutional Neural Network (CNN) trained on the FER2013 dataset [17] with the Deep-Face emotion analysis framework [2].

The CNN model consists of multiple convolutional layers, batch normalisation layers, max-pooling operations, dropout regularisation, and fully connected classification layers. The model was trained to recognise emotional expressions from grayscale facial images.

DeepFace provides an additional emotion prediction using a pre-trained deep learning architecture. The outputs from both models are combined through a weighted voting mechanism to improve robustness and prediction reliability.

For digital wellness assessment, individual emotions are grouped into three categories:

- Positive
- Neutral
- Negative

The resulting emotion group provides emotional context for alert computation and recommendation generation.

1) Ensemble Emotion Fusion Strategy

To improve emotion recognition performance on real-world facial images, BEAM employs an ensemble strategy that combines predictions from the FER2013 CNN model and the DeepFace framework. The motivation for this approach arises from the observation that the standalone CNN model occasionally misclassifies images captured under unconstrained conditions due to the controlled nature of the FER2013 training dataset. In contrast, DeepFace demonstrates greater robustness to variations in lighting, pose, and facial appearance.

Both models independently generate probability distributions across the emotion classes. The CNN processes gray-scale facial images resized to 48 × 48 pixels, while DeepFace analyses the original RGB image using its pre-trained emotion recognition pipeline. The resulting probability distributions are combined through a weighted averaging mechanism, assigning a weight of 0.40 to the CNN model and 0.60 to DeepFace.

The ensemble probability for each emotion class is computed as:

$$P_{ensemble} = 0.40 \times P_{CNN} + 0.60 \times P_{DeepFace}$$

where P_{CNN} and $P_{DeepFace}$ represent the probability values generated by the CNN and DeepFace models, respectively. The emotion class with the highest ensemble probability is selected as the final predicted emotion and subsequently utilised for alert computation, recommendation generation, historical analysis, and intervention generation. The ensemble approach improves robustness and reduces the impact of in-dividual model errors, thereby enhancing overall emotion recognition reliability.

D. Alert Computation Framework

The alert computation framework combines behavioural addiction predictions and emotional information to determine the overall wellbeing status of the user. This process enables the system to identify situations that may require additional attention or intervention.

The alert level is determined using predefined decision rules based on addiction severity and emotional condition. Users with high addiction levels are categorised as High Concern irrespective of emotional state. Moderate addiction lev-els are categorised as Moderate Concern, while users with low addiction levels are further evaluated based on their emotional condition.

The framework generates three alert categories:

- Low Concern
- Moderate Concern
- High Concern

These alert levels serve as the foundation for recommendation generation and intervention analysis.

TABLE II. ALERT LEVEL CLASSIFICATION RULES

Addiction Level	Emotion Group	Alert Level
High	Any	High Concern
Moderate	Any	Moderate
Low	Negative	Moderate
Low	Positive/Neutral	Low Concern

These alert levels serve as the foundation for recommen-dation generation, intervention analysis, and personalised wellness support within the BEAM framework.

E. Retrieval-Augmented Recommendation Framework

The recommendation generation module combines Large Language Models with Retrieval-Augmented Generation techniques to provide personalised digital wellness guidance [18].

A wellness knowledge base consisting of research-backed documents and digital wellbeing resources is stored within a ChromaDB vector database. The documents are converted into vector embeddings and indexed to support semantic retrieval.

When recommendations are requested, relevant documents are retrieved based on the user's behavioural and emo-tional context. The retrieved information is incorporated into the language model prompt, enabling the generation of recommendations that are both personalised and grounded in ex-ternal knowledge sources.

This approach improves recommendation relevance while reducing the possibility of unsupported responses.

F. Agentic AI Framework

The BEAM framework incorporates a LangGraph-based Agentic AI layer to support contextual reasoning and personalised insight generation. Unlike conventional recommendation systems, the agent is capable of analysing historical in-formation, retrieving relevant knowledge, and selecting ap-propriate tools based on the user's current condition.

The agent operates using three primary tools:

- Memory Tool
- RAG Tool
- Intervention Tool

The Memory Tool retrieves historical session information from persistent storage. The RAG Tool accesses relevant wellness knowledge from the vector database, while the Intervention Tool evaluates behavioural patterns and identifies potential risk indicators.

By combining information from these sources, the agent generates personalised observations, behavioural insights, and wellness recommendations.

G. Persistent Memory and Intervention Framework

Long-term behavioural monitoring is achieved through a persistent memory system implemented using SQLite. Each user session is stored together with behavioural inputs, emotional information, addiction predictions, alert levels, and generated recommendations.

The stored information enables longitudinal analysis of behavioural patterns across multiple sessions. Historical trends are subsequently utilised by the Agentic AI framework to generate wellness insights and identify recurring behavioural patterns [9], [10], [14].

The intervention framework continuously evaluates behavioural history to detect potentially concerning trends such as increasing screen time, deteriorating emotional conditions, reduced sleep duration, or elevated stress levels. When such patterns are identified, personalised intervention messages and wellness recommendations are generated.

The integration of memory, intervention analysis, and agent-based reasoning enables BEAM to provide continuous and personalised digital wellness support rather than isolated session-based recommendations.

H. System Workflow

The operational workflow of BEAM consists of a se-quence of interconnected stages that collectively support personalised digital wellness assessment and intervention. Initially, users provide behavioural information, including screen time, phone unlock frequency, social media usage, sleep duration, physical activity, stress level, anxiety score, and family interaction. Simultaneously, facial images are analysed through the emotion recognition module to determine the user's emotional state.

The behavioural prediction model classifies users into ad-diction categories, while the emotion recognition framework identifies the corresponding emotional group. These outputs are combined by the alert computation module to determine the overall wellbeing status of the user. Based on the generated alert level and emotional context, personalised wellness recommendations are produced using a Large Language Model.

The generated recommendations, behavioural information, emotional data, and prediction results are subsequently stored within the persistent memory layer. The Retrieval-Augmented Generation framework retrieves relevant wellness knowledge from the vector database, while the Agentic AI layer analyses historical behavioural patterns and generates contextual insights. Finally, the framework sup-ports intervention detection, weekly wellness reporting, and long-term behavioural monitoring to encourage healthier digital habits.

IV. EXPERIMENTAL CONFIGURATION

A. Dataset Description

The BEAM framework utilises two datasets for behavioural prediction and emotion recognition. The behavioural prediction module was developed using the Global Mobile Addiction Dataset, which contains approximately 3,000 records and 34 attributes representing behavioural, lifestyle, and psychological factors associated with smartphone usage. Features such as daily screen time, phone unlock frequency, social media usage, sleep duration, physical activity, family interaction, stress level, and anxiety score were utilised for ad-diction-level classification [1], [6], [12].

Prior to model training, the behavioural dataset was subjected to preprocessing procedures including missing value handling, feature selection, label encoding, and feature normalisation. These steps ensured data consistency and im-proved model reliability during training and evaluation. The processed dataset was subsequently divided into training and testing subsets to facilitate objective model comparison and performance assessment across multiple machine learning algorithms.

For emotion recognition, the FER2013 dataset was employed to train the Convolutional Neural Network (CNN) model. The dataset contains 35,887 grayscale facial images representing multiple emotional expressions and is widely used as a benchmark for facial emotion recognition research [17]. Images were resized to 48 × 48 pixels and normalised prior to model training and validation.

To improve robustness in real-world scenarios, the CNN predictions were combined with DeepFace-based emotion analysis through an ensemble strategy. The resulting frame-work leverages both dataset-driven learning and pre-trained facial analysis capabilities to provide reliable emotional con-text for digital wellness assessment and recommendation generation.

B. Experimental Environment

The BEAM framework was implemented using Python and several open-source libraries. Behavioural prediction models were developed using Scikit-learn, while TensorFlow and Keras were utilised for deep learning-based emotion recognition. DeepFace was incorporated as an additional emotion analysis framework to support ensemble prediction. The recommendation generation module employed GPT-4o Mini [18], while ChromaDB was used for semantic knowledge retrieval and SQLite was utilised for persistent memory management. Furthermore, the Agentic AI layer was implemented using LangGraph to support contextual reasoning, behavioural analysis, and intervention generation.

TABLE III. EXPERIMENTAL TOOLS AND TECHNIQUES

Component	Technology Used
Programming Language	Python
ML Framework	Scikit-Learn
DL Framework	TensorFlow, Keras
Emotion Recognition	DeepFace
Behavioural Prediction Model	Logistic Regression
Emotion Recognition Model	CNN Ensemble
LLM	GPT-4o Mini
Vector Database	ChromaDB
Persistent Memory Storage	SQLite
Agent Framework	LangGraph
User Interface	Streamlit
Data Processing	NumPy, Pandas

Table III. summarises the major technologies and soft-ware tools utilised in the development of the BEAM frame-work. The integration of machine learning, deep learning, large language models, vector databases, and Agentic AI technologies enabled the implementation of a comprehensive digital wellness platform capable of behavioural prediction, emotion recognition, personalised recommendation generation, and long-term wellbeing monitoring.

C. Evaluation Metrics

The behavioural prediction models were evaluated using Accuracy, Precision, Recall, and F1-Score metrics. These evaluation measures provide a comprehensive assessment of classification performance and prediction reliability.

Emotion recognition performance was assessed using classification accuracy, training and validation curves, and confusion matrix analysis. In addition, the overall BEAM frame-work was evaluated through end-to-end system testing to assess recommendation generation, memory utilisation, behavioural trend analysis, and intervention capabilities.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed BEAM framework was evaluated using behavioural prediction experiments, facial emotion recognition analysis, and end-to-end system validation. The objective of the evaluation was to assess the effectiveness of the individual modules as well as the overall framework in supporting personalised digital wellness assessment and intervention.

A. Behavioural Prediction Results

The behavioural prediction module was evaluated using multiple machine learning algorithms including Logistic Regression, Random Forest, Decision Tree, Support Vector Ma-chine, XGBoost, Gradient Boosting, Extra Trees, and K-Nearest Neighbour. Performance was measured using classification accuracy.

TABLE IV. COMPARISON OF ML MODELS

MODEL	ACCURACY(%)	PRECISION(%)	RECALL(%)	F1-SCORE(%)
LOGISTIC REGRESSION	91	91	91	91
XGBOOST	86	87	86	86
GRADIENT BOOSTING	84	85	84	84
EXTRA TREES	80	80	80	80
RANDOM FOREST	79	80	79	79
DECISION TREE	67	66	67	67
SVM	39	39	39	37
KNN	38	38	38	37

The results presented in Table IV. indicate that Logistic Regression achieved the highest prediction accuracy of 91%, outperforming all other evaluated models. XGBoost and Gradient Boosting also demonstrated competitive performance, while Support Vector Machine and K-Nearest Neighbour produced comparatively lower accuracies. Based on these observations, Logistic Regression was selected as the final behavioural prediction model due to its superior performance, computational efficiency, and interpretability.

The results demonstrate that behavioural indicators such as screen time, social media usage, stress level, anxiety score, and family interaction provide sufficient information for reliable addiction-level prediction.

B. Behavioural Classification Analysis

To further evaluate the classification capability of the selected model, a confusion matrix was generated for Logistic Regression.

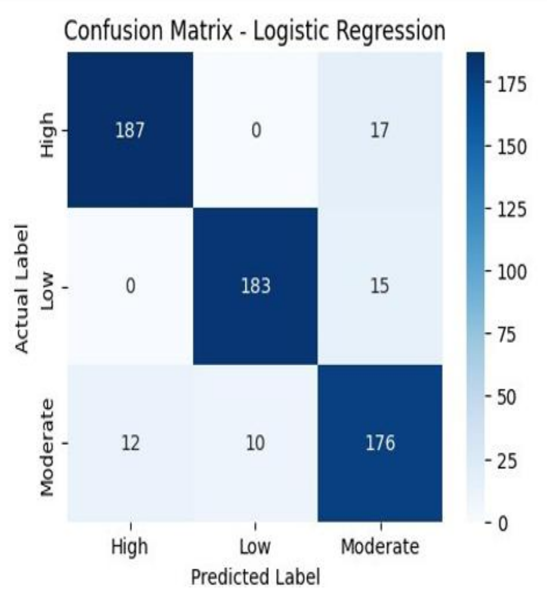

Fig. 2. *Logistic Regression Confusion Matrix*

The confusion matrix shows that the majority of samples were correctly classified into their respective addiction categories. High addiction, moderate addiction, and low addiction classes all exhibited strong classification performance with only a small number of misclassified instances.

The results indicate that the model can effectively distinguish between different levels of smartphone addiction and can therefore be utilised as a reliable component of the pro-posed digital wellness framework.

C. Feature Importance Analysis

Feature importance analysis was performed to identify the behavioural factors that contribute most significantly to ad-diction prediction.

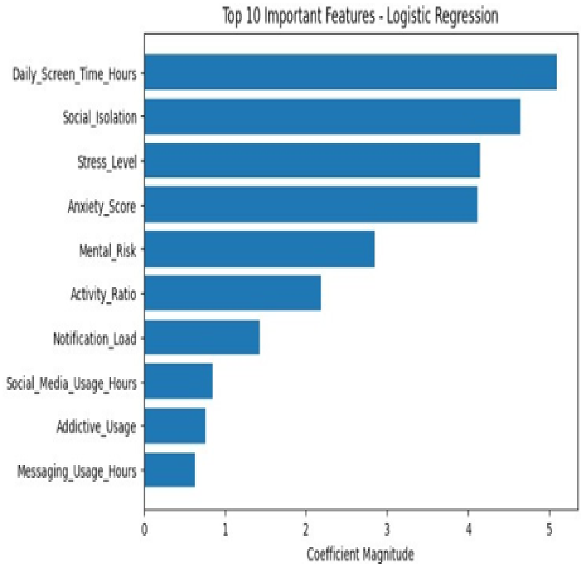

Fig. 3. *Feature Importance Analysis*

The analysis revealed that daily screen time, social isolation, stress level, anxiety score, and mental health indicators were among the most influential predictors. In contrast, certain demographic and lifestyle features contributed less to the final prediction.

These findings suggest that excessive smartphone usage is closely associated with psychological and behavioural wellbeing factors rather than screen-time duration alone. Similar observations have been reported in previous studies on smartphone addiction and behavioural analytics [1], [10]. The observations further justify the inclusion of emotional and contextual information within the BEAM framework.

D. Emotion Recognition Results

The emotion recognition module was evaluated using the FER2013 facial expression dataset [17]. The CNN model was trained using balanced emotion classes and standard preprocessing techniques.

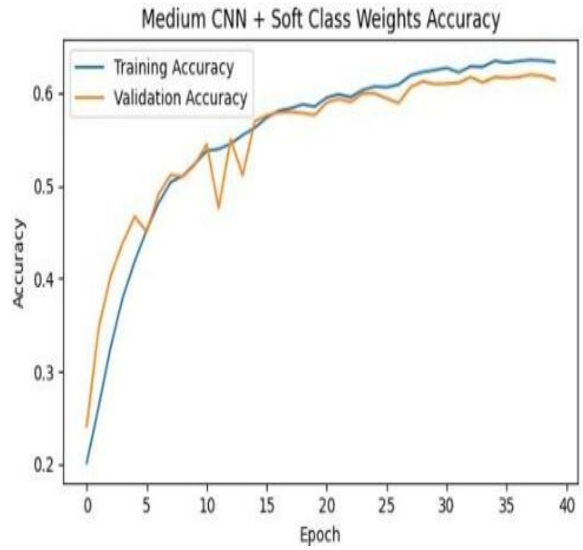

Fig. 4. *Training and Validation Accuracy Curve*

The training process demonstrated stable learning behaviour throughout the training period. The validation accuracy reached approximately 63%, while the training accuracy followed a similar trend. The relatively small difference be-tween the two curves indicates satisfactory generalisation and limited overfitting.

The CNN model achieved a seven-class emotion classification accuracy of approximately 63% on the FER2013 dataset. For the digital wellness application, the recognised emotions were subsequently grouped into three broader categories: Positive, Neutral, and Negative. This grouping re-duced class overlap and improved classification performance, resulting in an overall grouped emotion classification accu-racy of approximately 77%. The grouped emotion categories were utilised for alert computation, recommendation generation, and behavioural analysis within the BEAM framework. These results indicate that the emotion recognition frame-work provides meaningful emotional context for digital well-ness assessment and supports more reliable wellbeing analysis than relying solely on behavioural indicators

E. Emotion Classification Analysis

To analyse class-wise performance, a confusion matrix was generated for the emotion recognition framework.

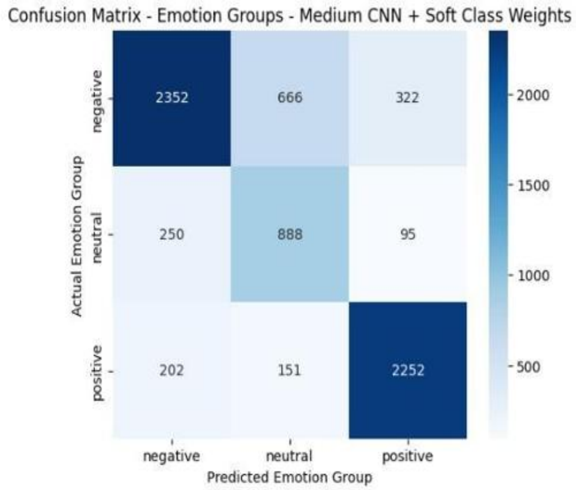

Fig. 5. *Emotion Recognition Confusion Matrix*

The confusion matrix demonstrates the ability of the model to classify emotional states into positive, neutral, and negative categories. Most samples were correctly assigned to their corresponding emotional groups, although some overlap was observed between neighbouring emotional classes.

The integration of DeepFace with the CNN model im-proved prediction robustness and reduced sensitivity to variations in image quality, facial orientation, and expression in-tensity [2]. The resulting emotional information provides an additional layer of context that complements behavioural addiction prediction.

F. BEAM Framework Evaluation

Beyond individual model performance, the complete BEAM framework was evaluated through end-to-end system testing. The evaluation focused on validating the interaction between behavioural prediction, emotion recognition, recommendation generation, persistent memory, Retrieval-Augmented Generation, and Agentic AI components.

The system successfully performed the following functions:

- Behavioural addiction prediction
- Facial emotion recognition
- Alert level computation
- Personalised recommendation generation
- Research-backed recommendation retrieval
- Historical session storage
- Behavioural trend analysis
- Intervention detection
- Weekly wellness report generation
- Wellness ranking generation

The integration of these components enabled the frame-work to provide comprehensive and personalised digital well-ness support beyond traditional screen-time monitoring ap-plications.

1) Recommendation Generation Evaluation

The recommendation generation component was evaluated based on its ability to produce personalised wellness guidance using behavioural information, emotional context, and alert levels. Recommendations were generated through GPT-4o Mini and further enhanced using the Retrieval-Aug-mented Generation framework. By incorporating information retrieved from the wellness knowledge base, the system was able to provide recommendations that were both context-aware and grounded in relevant digital wellbeing resources.

The generated recommendations varied according to the user's addiction level, emotional condition, and behavioural patterns. Users classified under higher risk categories received recommendations focused on screen-time reduction, stress management, sleep improvement, and healthier technology usage habits. In contrast, users with lower risk levels received preventive and motivational suggestions aimed at maintaining positive digital wellness practices.

2) Memory-Based Behaviour Analysis

The persistent memory component enabled the frame-work to retain information across multiple user sessions. Behavioural inputs, emotional states, alert levels, and generated recommendations were stored within the SQLite database and subsequently utilised for historical analysis.

The availability of historical information allowed the system to identify recurring behavioural patterns and monitor changes in user wellbeing over time. Trends such as increasing screen time, declining sleep duration, elevated stress lev-els, and repeated negative emotional states could be detected through memory-based analysis. This capability extended the framework beyond single-session recommendations and sup-ported long-term digital wellness monitoring.

Furthermore, the stored information provided valuable contextual information for the Agentic AI layer, enabling more personalised and consistent recommendation generation across multiple interactions.

3) Agentic AI Evaluation

The Agentic AI component was evaluated through end-to-end testing of its reasoning and intervention capabilities. The LangGraph-based agent successfully utilised memory retrieval, knowledge retrieval, and intervention analysis tools to generate contextual insights and personalised observations.

By combining current behavioural information with historical session data, the agent was able to identify behavioural trends and provide more informed recommendations than conventional rule-based systems. The integration of Retrieval-Augmented Generation further improved the quality of generated responses by supplying relevant wellness knowledge during the reasoning process.

The evaluation demonstrated that the agent could effectively support intervention detection, historical trend analysis, and personalised wellbeing guidance. These capabilities contributed significantly to the overall intelligence and adapt-ability of the BEAM framework.

G. Discussion

The experimental results demonstrate that multimodal wellbeing assessment can provide a richer understanding of user behaviour than screen-time monitoring alone. The behavioural prediction model achieved high classification accu-racy, while the emotion recognition framework contributed valuable emotional context for personalised decision-making.

The inclusion of persistent memory enabled longitudinal monitoring of behavioural patterns across multiple sessions. Furthermore, the Retrieval-Augmented Generation frame-work improved recommendation quality by incorporating knowledge retrieved from a curated wellness knowledge base. The Agentic AI layer extended the framework by sup-porting contextual reasoning, historical analysis, and personalised insight generation.

Unlike conventional digital wellbeing applications that primarily present usage statistics and basic usage controls [5], [7], [8], BEAM combines behavioural analytics, emotional assessment, memory-based reasoning, and knowledge-grounded recommendations within a unified framework. The results demonstrate the effectiveness of this integrated ap-proach for supporting healthier technology usage habits and long-term digital wellness management.

VI. CONCLUSION AND FUTURE WORK

This paper presented BEAM, a multimodal and agentic AI framework for personalised digital wellness that integrates behavioural analytics, emotion recognition, Retrieval-Aug-mented Generation (RAG), persistent memory, and autonomous reasoning within a unified platform. The framework was developed to address the limitations of conventional dig-ital wellbeing applications, which primarily focus on screen-time monitoring and provide limited personalised support.

The proposed system combines behavioural addiction prediction using machine learning with facial emotion recognition to generate a comprehensive assessment of user well-being. Experimental evaluation demonstrated that the behavioural prediction module achieved an accuracy of 91%, while the emotion recognition framework provided meaningful emotional context for personalised recommendation generation. The integration of GPT-based recommendations with Retrieval-Augmented Generation further improved the relevance and reliability of wellness guidance by incorporating knowledge retrieved from external wellness resources.

A key contribution of this work is the incorporation of persistent memory and Agentic AI capabilities for longitudinal wellbeing assessment. Historical session information is utilised to identify behavioural trends, support intervention analysis, and generate personalised insights. The LangGraph-based agent enhances the framework by enabling contextual reasoning through the integration of memory retrieval, knowledge retrieval, and intervention tools.

The experimental results indicate that combining behav-ioural data, emotional information, knowledge-grounded recommendations, and agent-based reasoning provides a more comprehensive approach to digital wellness assessment than traditional screen-time monitoring systems. The proposed framework demonstrates the potential of multimodal AI and Agentic AI techniques in supporting healthier technology us-age habits and long-term wellbeing management.

Although the proposed framework demonstrates promising performance, certain limitations remain. The behavioural prediction module currently relies on user-provided information, which may introduce inaccuracies due to subjective reporting. In addition, the performance of the emotion recognition framework may be influenced by factors such as image quality, lighting conditions, and facial orientation. The framework has been evaluated using publicly available da-tasets and controlled testing scenarios; therefore, further large-scale real-world validation is required to assess its generalisability across diverse user populations.

Future work will focus on mobile application deployment, automated behavioural data collection, wearable de-vice integration, and real-time wellbeing monitoring. Additional enhancements may include advanced multimodal emo-tion recognition, adaptive intervention strategies, personalised wellness coaching, and expanded knowledge bases for recommendation generation. These improvements can further strengthen the effectiveness and practical applicability of intelligent digital wellness systems.

REFERENCES

[1] V. Pawar, A. B. Patil, and A. Santra, "Predicting Smartphone Addiction Using Behavioral and Psychological Traits with Machine Learning," in Proc. 2025 Global Conf. Emerging Technology (GINOTECH), Pune, India,May 2025, doi: https://doi.org/10.1109/GINOTECH63460.2025.11077064

[2] R. Guo, H. Guo, L. Wang, M. Chen, D. Yang, and B. Li, "Development and Application of Emotion Recognition Technology: A Systematic Literature Review," BMC Psychology, vol. 12, no. 95, 2024, doi: https://doi.org/10.1186/s40359-024-01581-4

[3] M. Schlicher, Y. Li, S. M. K. Murthy, Q. Sun, and B. W. Schuller, "Emotionally Adaptive Support: A Narrative Review of Affective Computing for Mental Health," Frontiers in Digital Health, vol. 7, 2025, doi: https://doi.org/10.3389/fdgth.2025.1657031

[4] T. Wang, A. Seiger, A. Markowetz, I. Andone, K. Blaszkiewicz, and T. Penzel, "Smartphone Usage Patterns and Sleep Behavior in Demographic Groups: Retrospective Observational Study," Journal of Medical Internet Research, 2023

[5] S. Almoallim and C. Sas, "Toward Research-Informed Design Implications for Interventions Limiting Smartphone Use: Functionalities Review of Digital Well-being Apps," JMIR Formative Research.

[6] T. Si, A. Sharma, A. Singh, and I. Khan, "Trapped in the Mobile Screen: A Machine Learning Approach to Predict Nomophobia," in Proc. 2025 Int. Conf. Electronics, AI and Computing (EAIC), IEEE, 2025, doi: https://doi.org/10.1109/EAIC66483.2025.11101516

[7] D. Biedermann, S. Kister, J. Breitwieser, J. Weidlich, and H. Drachsler, "Use of Digital Self-Control Tools in Higher Education: A Survey Study," Education and Information Technologies, vol. 29, pp. 9645–9666, 2024, doi: https://doi.org/10.1007/s10639-023-12198-2

[8] D. A. Parry et al., "Digital Wellbeing Applications: Adoption, Use and Perceived Effects," Stellenbosch University, South Africa

[9] G. Gokulraj, S. Bose, S. Danya, N. Maheswaran, M. Pradeep Kumar, and T. Shruthi, "A Continuous Integration Framework for Machine Learning-Based Mobile Behaviour Analysis," in Proc. 2025 Innovations in Power and Advanced Computing Technologies (i-PACT), IEEE, 2025, doi: https://doi.org/10.1109/I-PACT65952.2025.11307944

[10] M. Archaya, D. P. Isravel, and J. P. M. Dhas, "Impact of Mobile Device Usage on User Behaviour: Predictive Analytics and Digital Health Perspectives," Journal of Information Systems Engineering and Management, vol. 10, no. 38s, 2025.

[11] S. A. Shifani et al., "A Novel Prediction Methodology to Detect Students Electronic Gadget Addiction Based on Deep Learning Principle," in Proc. 2025 Int. Conf. Recent Innovation in Science Engineering and Technology (ICRISET), IEEE, 2025, doi: https://doi.org/10.1109/ICRISET64803.2025.11252188

[12] H. Wu, "A Survey of Research on Anomaly Detection for Time Series," Xi'an, China.

[13] V. P. Cornet and R. J. Holden, "Systematic Review of Smartphone-Based Passive Sensing

for Health and Wellbeing," Journal of Biomedical Informatics, vol. 77, pp. 120–132, 2018.

[14] G. Chittaranjan, J. Blom, and D. Gatica-Perez, "Who's Who with Big-Five: Analyzing and Classifying Personality Traits with Smartphones," EPFL, Lausanne, Switzerland.

[15] A. Wang, "Research on App Usage Behavior Analysis and Prediction Algorithms for Mobile Users," in Proc. 2019 11th Int. Conf. Measuring Technology and Mechatronics Automation (ICMTMA), IEEE, 2019.