ISSN: 2788–7669 *Journal of Machine and Computing 5(3)(2025)*

Cognitive Emotion Aware Systems Using Multimodal Signals and Reinforcement Learning

[1]Ezil Sam Leni A, [2]Revathi T and [3]Niranchana Radhakrishnan
[1,2,3]Deaprtment of Computer Science and Engineering, Alliance School of Advanced Computing, Alliance University, Anekal, Bengaluru, Karnataka, India.
[1]lenisatish@gmail.com, [2]revathigiri13@gmail.com, [3]niranchana.r14@gmail.com

Correspondence should be addressed to Niranchana Radhakrishnan : niranchana.r14@gmail.com

Article Info
Journal of Machine and Computing (https://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202505106.
Received 28 November 2024; Revised from 04 January 2025; Accepted 18 March 2025.
Available online 05 July 2025.

Abstract – Predicting human behaviour is a complex task. Traditional methods often rely on explicit user input or external observation, which can be restrictive and impractical in real-world scenarios. As an alternative, Brain-Computer Interfaces (BCIs) offer a more direct and specific means of accessing cognitive and emotional states, providing valuable insights into human intentions and decision-making processes. This paper proposes a novel method that predicts and suggests personalised emotion-based activities for individual users based on multi-modal sensory data collected from the brain, body, and environment. Our method overcomes the limitations of conventional systems by incorporating a multi-modal data collection set throughout the day to understand user context and intent better. By analysing this data, we predict the emotions-based practice of the user's day. We train our method using state-of-the-art, nature-inspired reinforcement learning algorithms and agent technology to optimise its optimisations and personalised continuously. The performance evaluation shows that the accuracy and F1 score for the proposed method achieved 95.6% and 84%, respectively, achieving 2 to 3% more accuracy than AI-based emotion state-of-the-art detection methods.

Keywords – Agent Technology, Brain-Computer Interfaces, Human Behavior, Personalized Daily Activities, Multi-Modal Sensory.

I. INTRODUCTION

Our social life has become increasingly integrated with Artificial Intelligence (AI) due to its ability to communicate and integrate naturally with humans in the most effective and trustable manner. AI has gained this potential due to its intelligence in understanding and managing emotions with humans, referred to as emotional intelligence (EI). Emotion is a complex reaction involving physiological, behavioural, and cognitive changes due to internal or external stimuli. It accurately assesses the self-state and the emotional states of others to manage and regulate social communication suitable to the environment. It is critically intertwined with our decisions in complex situations, demonstrating that the cognitive processes such as perception, memory and learning are complex to separate. However, understanding and predicting human emotions are very demanding tasks that have motivated researchers to explore innovative approaches beyond traditional methods and develop computational models of emotions [1]. While past efforts relied heavily on explicit user inputs or external observations, these approaches often need to be revised in real-world scenarios. As a promising alternative, emerging Brain-Computer Interfaces (BCIs) provide more valuable insights into direct and specific means of accessing cognitive and emotional states [2].

Artificial Emotional Intelligence (AEI) is very attractive in healthcare decision-support systems. As AI plays a growing role in healthcare, the need for EI systems capable of interacting effectively with patients becomes paramount. AI-powered healthcare solutions can revolutionise patient care by recognising and understanding human emotions, expressing empathy, and potentially even experiencing emotions internally.

The Emotion recognition (ER) and Emotion classification (EC) technologies are rapidly evolving, fuelled by a real-time data explosion and advancements in artificial intelligence (AI). A 2022 study by Gartner found that 85% of customer interactions will be managed without a human by 2030. Understanding customer emotions in real-time will be critical for success in this automated future. Meanwhile, a 2021 study in the Journal of Personality and Social Psychology revealed that emotional intelligence – the ability to perceive, understand, and manage emotions – is a strong predictor of job performance, with emotionally intelligent employees generating 33% more revenue and experiencing a 12% increase in productivity. These technologies typically analyse a combination of data points streamed in real time, including facial expressions, speech patterns, and even written text. Facial recognition software might track subtle movements in eyebrows, eyes, and the mouth to identify emotions like happiness, anger, or sadness at a frame rate of over 30 frames per second. Speech analysis can detect variations in pitch, tone, and volume that can signal emotional states in real-time conversations. Even written language can offer clues – the frequency of certain words or the use of exclamation points can indicate sentiment, with AI models analysing text streams as they're typed.

Emotional classification is the process of sorting emotions into distinct categories. There are two main approaches: categorising emotions as essential and distinct or placing them on a spectrum. The influential theory suggests six universal basic emotions: happiness, sadness, anger, fear, surprise, and disgust. Other emotions might be more complex combinations of these or arise from situations. This study area is still evolving, but it has applications in fields like psychology, computer science, and marketing. While there's some agreement on basic emotions like happiness or anger, expressed through facial features and biology, classifying emotions gets complicated [4]. Further complicating things, cultural experiences can influence how emotions are perceived and expressed. A smile might signify happiness in one culture but politeness in another [5]. This is why researchers also consider the context in which emotions arise. Machines can analyse text, speech, and even facial expressions to try to identify emotions. This has applications in sentiment analysis for social media, where companies can gauge customer satisfaction, or in developing AI that can better understand and respond to human emotions.

From personalised diagnoses and empathetic care to mental health support and decision support, AEI has immense potential to transform healthcare delivery, improving patient experiences, better clinical outcomes and a more compassionate healthcare system. This paper delves into this exciting realm by proposing a novel method capable of predicting emotions. This experiment leverages the power of multi-modal sensory data collected from the brain, body, and environment, providing a comprehensive understanding of individual context and intent. Through a sophisticated deep learning model, the proposed method analyses this data to predict the trajectory of the emotion in a user's day.

The following contributions are included in the proposed emotion prediction method:

- We have developed a non-invasive system for continuous Emotional Intelligence (EI) monitoring by identifying a robust multi-modal data set that will seamlessly capture data from various biomarkers based on emotion generated through text conversation, visual contact, and speech modulation and variations.
- We have developed an enriched deep-learning personalised model through nature-inspired reinforcement learning, which continuously optimises its recommendations and provides highly customised future casting.
- We have conducted experiments by training the model on relevant features using the feedback and suggestion impact mechanism. The models have been tested and continuously monitored according to the proposed reward-penalty system, which has been optimised effectively.
- We provide a critical and comprehensive evaluation based on metrics, such as relevance, engagement, and user satisfaction, to measure the effectiveness of the proposed recommendations from the developed EI system.
- We expose the challenges and existing problems in analysing EI using cognitive agents tailored to an individual.

This paper is organized as follows. Section 1 highlights the need for this work as an introduction, while Section 2 discusses the related works, also underscoring the research gap. Section 3 introduces the proposed prediction methodology. Experimentations, results and analysis in comparisons with state-of-the-art methods are discussed in detail in Section 4, just before the conclusions of our work are briefly depicted in Section 5.

II. RELATED WORK

This research paper aims to understand the significance of the role played by AI in cognitive processes and its application in emotional intelligence. This section focuses on the studies made in existing works related to developing the architecture for cognitive agents and modelling of emotions in the agents. This section sheds light on related computational models of emotions and their analytical evaluation methods.

Several cognitive architectures have been developed in the past decades, and this illustrates a shift from symbolic cognitive architectures to neurally inspired architectures. Nowadays, there is also interest in developing hybrid architectures. This rapid growth has excellent potential for the development of intelligent systems. The evolved architectures include diverse perspectives of disciplines, ranging from psychoanalysis to neuroscience. It reflects the multifaceted nature of human cognition based on seven abilities: perception, attention, action selection, memory, learning, reasoning, and meta-reasoning. [1, 2]

In [3], the author introduces the novel CAIO architecture for improved social interaction between humans and robots. It focuses on cognitive and affective aspects, leading to more engaged interactions. It enables the robots to process sensory information, reason about the situation, and generate responses accordingly. Emma is an innovative chatbot that incorporates emotional awareness and personalised micro-interventions to promote mental well-being. She tailors her responses and recommendations based on user interactions and smartphone data inputs. She is designed to place empathetic and supportive responses, which create trust and connection with the users through short activities and exercises such as breathing techniques and social interaction prompts. Emma is under development as she is trained with limited samples. She must be explored for different populations and settings. It is crucial to study and investigate the long-term effects of her interventions [4].

The authors of [5] strongly criticise defining emotional intelligence as it fixes a set of abilities and proposes dynamic models that manage through the emotional response cycle. This model emphasises the need to design and develop customised, effective regulation strategies to highlight individual needs.

A comprehensive overview of categorising emotions and effects outlining the methods to measure and detect them is presented in [6]. It describes the taxonomy to categorise the theories for integrating emotions in human-computer interaction. In [7], a general framework for emotion modelling in cognitive agents is discussed. The model encompasses four components: emotion generation, emotion experience, emotion regulation, and emotional modulation. This paper introduces the classification of evaluation methods used in various studies on agents with emotional intelligence. The amygdala is an emotional centre in human beings, and its studies are critical for cognitive and decision-making papers. The inputs from the amygdala will help researchers to have a better understanding and processing of social cues, which will guide social interactions. It facilitates the readers to understand emotional learning, which intertwines emotional stimuli with future behaviours and responses [8]. The research work carried out in [9] discusses the neural basis of emotions, which plays a critical role in understanding mental health status like anxiety and depression and emotional processing. Insights from this research are very significant to refine our understanding of developing an emotional intelligence system. In [10], the proposed framework mimics human-like cognitive complexity by integrating large language models (LLMs) with autonomous agents. The modular mind theory proposes certain LLMs as alternates for the cognitive modules of the brain. Autonomous agents represent the human cognition modules, which are additionally strengthened by LLMs.

A conversational agent is a virtual agent capable of conversing, expressing, and obtaining information from the environment. It infers and acts according to the given scenario. Emotion and cognitive process integration is necessary to develop such agents that can be evaluated based on believability and social acceptance [11]. Adaptive control of thought-rational (ACT-R) [12] is a traditional cognitive architecture comprising a set of rule-based information-processing modules to regulate their behaviour on a central system. The emotional module of ACT-R determines the positive and negative emotions to derive inferences on learning and problem-solving processes. The computational model of this architecture evaluates affective valence and arousal in decision-making. It needs more coverage of many emotional intelligence requirements [13, 14]. ALMA [15, 16], a computational model of emotion, regulates nonverbal and verbal expressions. It demonstrates enriched communication skills with humans. This model presents the affective response derived from emotion, mood, and personality. E-VOX [18] is inspired by ALMA and a Soar-based cognitive architecture [17]. It assists in capturing information from Wikipedia. It is an effective amalgamation of SOAR and ALMA. Integrative framework (InFra) [19] presents an intertwined emotional and cognitive architecture for autonomous agents. It generates and identifies emotions based on environmental stimuli. It includes personality and culture as parameters that influence the emotions generated in the system.

The MAMID architecture [20] employs emotions based on personality traits and cognitive signals. It uses complex cognitive elicitors, such as personal history and the convergence of expectations and goals, to perform a discrete stimulus assessment (for four basic emotions). The Fuzzy logic adaptive model of emotions (FLAME) [21] is based on appraisal theory and Rossmann's theory. It uses fuzzy logic as a behaviour selection strategy. This model observes the event that occurs and generates a corresponding emotional value based on the model's history. The appropriate emotion is selected according to the mood of the scenario, which is then used by the decision-making system. The Ethical Emotion Generation System (EEGS) [22] is a four-stage computational model of emotions. These are emotion elicitation, cognitive appraisal, affects generation, and effect regulation. The model is evaluated and assessed based on the positive and negative emotional values, personalities and cognitive mental states such as goals and attitudes. In [23], the authors propose a model that designs robot personalities reflecting natural and human-like interactions with it. It has been achieved by integrating social psychology and fuzzy logic. It defines six categories of emotional dimensions with twelve emotions. These emotional impacts created by external stimuli are studied based on the effect created and shown in the robot's responses.

The authors of [24] present a detailed illustration of adopting cognitive-affective architectures as affective user models in behavioural health technologies. It is based on internal processing and reflections upon user state changes. It stresses using such models to support mental health and well-being through technologies. The scope of integrating dynamic representations of emotions, semantic maps, and moral schemas of the human brain's structure with AI to understand emotional intelligence, enhance social interactions, improve decision-making, and foster transparency is presented in [25].

This section addresses one of the fundamental challenges in designing and developing cognitive agents with emotional intelligence. These agents are evaluated based on computational models of emotion, as discussed in the previous section. The evaluations are carried out in predefined scenarios and are subjective. It is accepted that if the agent performs well in such conditions for the stimuli, then the computational model satisfies the metrics. The paper evaluates the agent's response based on the text conversation. The agent is trained for a case study where the cancer diagnosis is announced. The same scenario is mimicked with a scenario involving a human being, and the emotional reaction of the agent is studied.

EBDI uses the Tileworld system to simulate a multiagent environment and evaluates them by tuning different parameters and determining the fitness of the agent architecture. The collection of agents represents different emotions, and their impact on decision-making is studied. Bourgeois et al. experimented with and evaluated the GAMA based multiple agents environment for crucial situations and resultant emotions. In the framework it identifies and interprets emotional cues from facial expression, tone of voice and body language and assigns them a meaning. Accordingly, appropriate emotional consequences are observed in an interview scenario and assessed for integration with other cognitive functions. Computational models developed to simulate human emotions in artificial agents are evaluated based on believability and social acceptance. Social acceptance assesses the resultant emotion from the cognitive agent based on the degree of matching the expectations that humans perform in the same scenario as the social setup. Social acceptance supports the believability criteria as it is always unacceptable in a real-time social environment.

The authors propose a novel approach to training AI through collaborative learning games. This algorithm would leverage deep reinforcement learning, where the AI learns by playing the game and receiving rewards for successful actions. Uniquely, it incorporates natural language processing, allowing the human player to guide the AI through natural language instructions. The paper likely explores the design of this algorithm, its effectiveness in training the AI, and the potential benefits of human-machine collaboration in learning environments.

The research paper highlights an EEG-based emotion recognition technique, which analyses scalp electroencephalogram (EEG) recordings to categorise a person's emotional state. This has applications in healthcare and human-computer interaction. The paper highlights the role of brain region interactions in emotion processing. To capture these interactions, the researchers propose a novel method using a graph convolutional network (GCN). This approach treats EEG channels as nodes in a graph, with connections between them representing the relationships between brain regions. By processing the EEG data through this GCN, the model can learn the complex interplay between brain regions that underlie different emotions.

The research by authors suggests our brains process rewards and emotional surprises differently. Traditionally, reinforcement learning places emphasis on reward prediction errors (PEs) to shape behaviour. This study introduces the concept of affective PEs, which gauges the difference between anticipated emotions and what actually occurs. The researchers used electroencephalography (EEG) to measure brain activity during social learning tasks. Their findings revealed distinct neural signatures for both reward PEs and affective PEs. This implies separate neural mechanisms underlie how we learn from rewards and emotional surprises, suggesting emotions play a crucial role in shaping our behaviour, even beyond the pull of external rewards.

This research work in delves into using reinforcement learning to optimise interventions and personalise feedback in real time using data from wearable sensors. Wearable sensors provide a constant stream of information about a user's state, which the reinforcement learning model can analyse. The model then determines the most effective intervention or personalised feedback to deliver based on the user's situation and goals. This approach holds promise for enhancing various applications, such as healthcare and fitness, by providing tailored support and improving user outcomes.

The fascinating research in the paper investigates a method to improve the accuracy of recognising emotions from facial expressions. This approach utilises ensemble learning, which combines multiple classifiers (algorithms that categorise data) instead of relying on a single one. The paper likely explores how the ensemble is created, what kind of individual classifiers are used, and how they analyse facial features. By combining the strengths of various classifiers, the ensemble method aims to achieve better accuracy in recognising emotions like happiness, sadness, or anger from human faces compared to traditional single-classifier approaches

The challenges addressed in these works are integrating cognitive architectures with real-world sensors and actuators and developing more effective learning algorithms. They also highlight the gaps between theoretical models and human-level intelligence. In this paper, we highlight the gaps in developing systems that are more engaging, easier to use, and more easily learned in the human-computer interaction (HCI) domain.

III. PROPOSED PREDICTION METHODOLOGY

The proposed system overcomes the limitations of current recommendation systems, which often rely on generic suggestions that fail to capture the nuances of individual emotions. By incorporating a combination of human activity, biopotential sensors, and environmental data, the system gains a deeper understanding of the user's emotional state, physical activity levels, and surrounding context. Reinforcement learning further elevates this EI system through state-of-the-art algorithms and agent technology. The system learns to associate specific emotional patterns in the data based on stimuli and enhances the emotions with increased accuracy. It represents a significant leap forward in the HCI field, promising to revolutionise how we interact with technology. By gaining deeper insights into the intricate relationship between brain activity, physical state, and human behaviour, this paper paves the way for developing more intuitive and responsive interfaces. These interfaces will adapt to each user's needs and preferences, leading to a more natural and seamless interaction with technology, as illustrated by the high-level architecture of **Fig 1**. The proposed emotion prediction system employs reinforcement learning to learn the user's emotions based on multi-layered sensor information, namely human activity, speech, and visual information. The proposed emotion detection and prediction system is shown in **Fig 2**. This section discusses the agent design for the proposed motion prediction method. The agent has been designed using artificial rabbit optimisation algorithms to identify possible personalised emotions.

Improved Artificial Rabbit Optimisation Algorithm

The Artificial Rabbit Optimization (ARO) algorithm is designed based on the two laws of survival by the rabbit from the everyday world: the process of detour foraging and the process of random hiding. The process of detour foraging is the strategy of exploration to take food from near to the current place of stay. Random hiding is the strategy of making a move taken by a rabbit to other burrows, primarily for hiding further, and this strategy is known as the exploitation process.

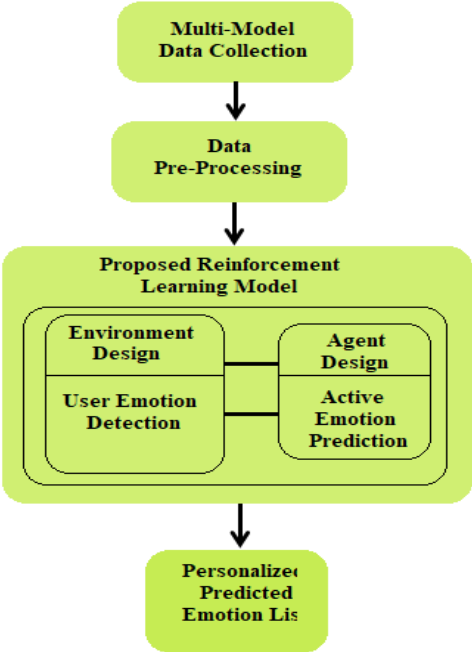

Fig 1. Proposed Architecture for Emotion Prediction.

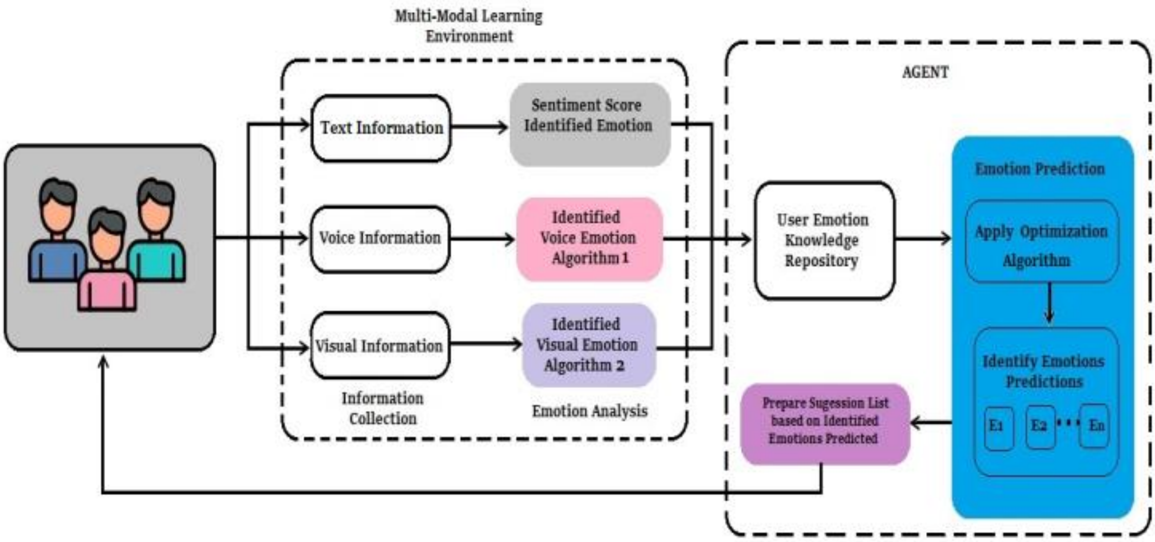

Fig 2. Proposed Emotion Detection and Predication Method.

The starting point of any optimum target search algorithm depends on initialisation. In the ARO algorithm, the following parameters are mandatory, considering the design variable size with the dimension of d: N is the number of the artificial rabbit colony, and the lower and upper limits are mentioned as UL and LL. The process of initialisation is performed as follows:

$$\vec{Y}_{i,k} = r.(UL_k - LL_k) + LL_k \rightarrow \quad (1)$$

Here, $\vec{Y}_{i,k}$ is the position indicating the i^{th} rabbit in the k^{th} dimension, and the value r is selected randomly.

Process of Exploration

Metaheuristic algorithms are designed based on the concept of two main activities: the process of exploration and the process of exploitation. In the ARO algorithm, the process of detour foraging is considered a process of exploration.

Detour foraging is the affinity of each rabbit to stimulate in and around the discovered food sources and randomly discover another rabbit location selected from the group to collect sufficient food. The equations used in the process of detour foraging are given as:

$$\vec{W}_i(t+1) = \vec{Y}_j(t) + R.\left(\vec{Y}_i(t) - \vec{Y}_j(t)\right) + \mathrm{rand}\left(0.5(0.05 + r_1)\right).n_1 \quad \rightarrow \tag{2}$$

$$R = L.C \quad \rightarrow \tag{3}$$

$$L = l_i - l_j \quad \rightarrow \tag{4}$$

$$D = \left[1 + \frac{\mathrm{ReqEmotionY_i}}{\mathrm{AvaiEmotion}}\right] \quad \rightarrow \tag{5}$$

$$l_i = \left(e - e^{D_i^2}\right).\sin(2\pi r_1) \quad \rightarrow \tag{6}$$

$$l_i = \left(e - e^{D_j^2}\right).\sin(2\pi r_1) \quad \rightarrow \tag{7}$$

$$C(k) = \begin{cases} 1 & \text{if } k = G(l), lk = 1, \ldots d \text{ and } l = 1, \ldots [r_3.d] \\ 0 & \text{OtherWise} \end{cases} \quad \rightarrow \tag{8}$$

$$G = \mathrm{randp}(d) \quad \rightarrow \tag{9}$$

$$n_1 \sim N[0,1] \quad \rightarrow \tag{10}$$

Here, $\vec{W_i}(t+1)$ indicates the updated position of the rabbit. The $\vec{Y_i}(t)$ indicates the location of the i^{th} rabbit, and $\overrightarrow{the\ Y_j}(t)$ represents artificial rabbits from a random position.

Algorithm 1: Initial Stage of IAROA
1: Initialised populations
2: For each $U_i \in N$ begin
3: Apply random process for selecting food source
4: Emotion list $EL_{U_i} = \{Text_{U_i}, Speech_{U_i}, Visual_{U_i}\}$
5: Compute Fitness value for EL_{U_i}
6: End
7: Assign $B_{solution} = EL_{U_i}[0]$
8: For each $j \in N\ and\ i \neq j$
9: $If\ \left(EL_{U_i}[j] > B_{solution}\right)\ then$
10: $B_{solution} = EL_{U_i}[j]$
11: End
12: Return $(B_{solution})$

Process of Exploitation

In the process of random hiding, rabbits usually select burrows in and around their nests and arbitrarily select one to hide to reduce the probability of being predated. The procedure for arbitrarily creating burrows by the rabbits is given as follows. The i^{th} rabbit produces the j^{th} burrow as in (11):

$$\vec{B}_{i,j}(t) = \vec{Y}_i(t) + H.g.\vec{Y}_i(t) \quad \rightarrow \tag{11}$$

$$H = \left[1 - \frac{\text{Required Emotion}+1}{\text{Available Emotion}}\right].n_2 \quad \rightarrow \tag{12}$$

$$g(k) = \begin{cases} 1 & \text{if } k == j, \quad lk = 1, \ldots, d \\ 0 & \text{OtherWise} \end{cases} \quad \rightarrow \tag{13}$$

$$\vec{W}_i(t+1) = \vec{Y}_i(t) + R.\left(r_4.B_{i,j}(t) - \vec{Y}_i(t)\right) \quad \rightarrow \tag{14}$$

$$G_r(k) = \begin{cases} 1 & \text{if } k == [r_5.d], \quad lk = 1, \ldots, d \\ 0 & \text{OtherWise} \end{cases} \rightarrow \tag{15}$$

$$\vec{B}_{i,j}(t) \leftarrow \vec{Y}_i(t) + H.G_r.\vec{Y}_i(t) \quad \rightarrow \tag{16}$$

Here $i = 1, \cdots, N$ and $j = 1, \cdots, d$, and n_2 are designed with a standard normal distribution function. The value of H means the parameter of hidden, which will linearly decrease from 1 to $\frac{ReqEmotionY_i}{AvaiEmotion}$ with stochastic perturbations. The value of H will tend to decrease in general, and this will maintain stable changeovers from the exploration phase to the exploitation phase during the iterations.

$$\vec{Y}_i(t+1) = \begin{cases} \vec{Y}_i(t) & If\ f\left(\vec{Y}_i(t) \le f\left(\vec{W}_i(t+1)\right)\right) \\ \vec{W}_i(t+1) & If\ f\left(\vec{Y}_i(t) > f\left(\vec{W}_i(t+1)\right)\right) \end{cases} \rightarrow \tag{17}$$

Where $\vec{W}_i(t+1)$ is the newly updated position for the rabbit, $\vec{B}_{i,j}(t)$ denotes an arbitrarily selected warren among the d number of burrows created for hiding the rabbit, and r_4 and r_5 are random numbers taken from the limit of 0 to 1. R is calculated by using (4) – (8). The new position for the i^{th} rabbit is computed by using (17).

Equation (17) illustrates an adaptive update for artificial rabbits. The rabbit repeatedly applied the selection process to stay in the current position or move to a new one based on the sustainable score and the value to be adopted. Generally, the working principle of the optimisation algorithm for the exploration phase depends on population preference in the early stages and in the middle and final stages, an exploitation phase will participate.

The artificial Rabbit Optimisation algorithm depends on the level of emotion of rabbits, and the emotion will decrease over time. This will initiate the transition from the phase of exploration to exploitation. The proposed algorithm calculates additional emotion factors for rabbits' sustainability. This will measure the sustainability of a rabbit based on the current emotion level. The sustainable score for the rabbits is measured as follows:

$$SFA(t) = \frac{1}{Freq_t} \times \left[1 - \frac{Required\ Emotion}{Available\ Emotion}\right].\ln\frac{1}{r} \rightarrow \tag{18}$$

$$Freq_t = \frac{1}{\sigma\sqrt{2.\pi}} e^{-\frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2} \rightarrow \tag{19}$$

Here, $Required\ Emotion$ and $Available\ Emotion$ mean required success for everyday survival and total available sustainability in emotion handling. The $Freq_t$ is the frequency of rabbit exploration that occurred, and this is calculated based on the normal distribution over [0 to 1]. The r is an arbitrary number from the range of [0 to 1]; the σ and μ are the mean and standard deviation for the frequency of exploration rabbits.

Algorithm 2: Process of Exploration and Exploitation
1: For each iteration from 1 to MAX begin
2: For each $U_i \in N$begin
3: Compute emotion fitness value using equation (1) and (2)
4: Compute the process of exploration or exploitation
$$\vec{Y}_i(t+1) = \begin{cases} \vec{Y}_i(t) & If\ f\left(\vec{Y}_i(t) \le f\left(\vec{W}_i(t+1)\right)\right) \\ \vec{W}_i(t+1) & If\ f\left(\vec{Y}_i(t) > f\left(\vec{W}_i(t+1)\right)\right) \end{cases}$$
5: Compute New Fitness value for EL_{U_i} according to $\vec{Y}_i(t+1)$
6: $B_{Solution} = FindBest\left(EL_{U_i}, 0 \le i \le N\right)$
7: End
8: End
9: Return $(B_{Solution})$

Proposed Reinforcement Learning Algorithm

This section discusses the proposed reinforcement algorithm for the personalised emotion prediction and recommendation method. The initial stage is constructed with predicted user emotions from the individual users, and the emotion recommendations will be prepared based on the artificial rabbit optimisation technique. The following algorithm 1 explains the working principle of the proposed reinforcement learning algorithm.

Algorithm 3: Proposed Reinforcement Learning Algorithm
1. For each $E_{U_i}, 1 \leq i \leq N$ begin 2. Identified list of emotion instances as $$E_{U_i} = \{(x_i, y_i, z_i)_{U_i}, 1 \leq i \leq N\}$$ 3. End 4. For each $(x_i, y_i, z_i)_{U_i}, 1 \leq i \leq N$ begin 5. Initial stage segmentation based on the probability $$\overrightarrow{W_i}(t+1) = \overrightarrow{Y_j}(t) + R.\left(\overrightarrow{Y_i}(t) - \overrightarrow{Y_j}(t)\right) + rand\big(0.5(0.05 + r_1)\big).n_1$$ 6. The emotion list for a user U_K will be updated as $EL_{U_K} = \{Text_{U_K}, Speech_{U_K}, Visual_{U_K}\}$ 7. Based on probability and fitness values, prepare a predicated list as follows $U_i \in U_K$. $$H = \left[1 - \frac{Required\ Emotion + 1}{Available\ Emotion}\right].n_2$$ 8. Apply ***Algorithm 1*** for initial population 9. Apply improved Artificial Rabbit Colony Algorithm (Algorithm 2) over predicted emotion list EL_{U_i} 10. End 11. Update emotion prediction list with selected emotion. 12. Expected Reward point calculated as follows: $$ER_T = \sum_{i=0}^{T} \gamma^i SAF(i) \rightarrow (20)$$ Here, γ^iweight factors the particular course. The weight factors are assigned within the interval of [0, 1]

IV. INTRODUCTIONEXPRIMENTATION, RESULTS, AND ANALYSIS

This section discusses the experimental setup for the performance evaluation and result analysis. The proposed method uses two datasets to recognise the user's emotions using human activity, speech, and visual expressions.

Experimental Setup

The proposed method was experimented using the Python 3.4.2 tool kit (PyTorch 1.12.1) with the basic hardware configuration of the i7 Intel core system. The cross-entropy loss function is used to measure the Adam optimiser model. The dropout rate is set as 0.3 and 0.4 to avoid overfitting in the proposed method. During the experimentation, the model has been trained with 50, 80, and 100 epochs each with different batch sizes of 40, 50, and 60. The experiment is carried out in three different modes wrt dataset: (*i*) only with MELD dataset, (*ii*) only with human activity dataset, (*iii*) combination of MELD and human activity dataset. The proposed method has been experimented in three phases: training, validation, and testing. The proposed method was trained with a training dataset, and the validation dataset was used to evaluate the trained model. The testing model has been used to predict the results from the test dataset. The training and evaluation phases are conducted with the above-given batch sizes for each dataset. The testing results are measured using confusion matrix method, and the results are evaluated by using Accuracy, F1 score, Precision, and Recall.

Results and Analysis

The performance of the proposed model is evaluated using different metrics namely weighted average accuracy (W_{AAC}), weighted average F1 score (W_{AF1}), weighted average precision (W_{APr}) and weighted average recall (W_{ARe}). The mathematical representations of the metrics are presented in the following equations:

$$\mathrm{W_{AAC}} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{W_i}.\frac{\mathrm{TP_i+TN_i}}{\mathrm{TP_i+TN_i+FP_i+FN_i}} \rightarrow \tag{21}$$

$$\mathrm{W_{AF1}} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{W_i}.\frac{\mathrm{2.Precision_i.Recall_i}}{\mathrm{Precision_i.Recall_i}} \rightarrow \tag{22}$$

Here, $\mathrm{Precision_i}$ and $\mathrm{Recall_i}$ can be calculated by using equation (23) and (24),

$$\mathrm{Precision_i} = \frac{\mathrm{TP_i}}{\mathrm{TP_i+FP_i}} \rightarrow \tag{23}$$

$$\mathrm{Recall_i} = \frac{\mathrm{TP_i}}{\mathrm{TP_i+FN_i}} \rightarrow \tag{24}$$

$$W_{APr} = \frac{1}{N}\sum_{i=1}^{N} W_i.\frac{TP_i}{TP_i+FP_i} \rightarrow \tag{25}$$

$$W_{ARe} = \frac{1}{N}\sum_{i=1}^{N} W_i.\frac{TP_i}{TP_i+FN_i} \rightarrow \tag{26}$$

The experimental evaluation for the proposed method calculates the category-wise weighted average accuracy and F1 score. **Table 1** and **Fig 3** present the average accuracy, F1 score, precision, and recall values. **Table 1** shows the top 5 emotions (anger, fear, joy, neutral, and sadness) taken for the evaluation. **Tables 2** presented comparisons with state-of-the-art methods, with the testing batch sizes of 40, 50, and 60. The testing dataset contains 40, 50, and 60 dialogue instances with the same epoch size of 100. The proposed method has been evaluated with 2610 dialogue instances collected from the MELD dataset of which 42 % are natural emotions and the remaining 58% are the other six emotions.

Table 1. Category-Wise Performance Evaluation for The Proposed Method

Evaluation metrics	Categories				
	Anger	**Fear**	**Joy**	**Neutral**	**Sadness**
W_{AAC}	74.3	81.2	79.2	84.3	72.7
W_{AF1}	77.5	79.4	82.4	81.5	83.2
W_{APr}	72.5	76.3	78.4	83.1	78.4
W_{ARe}	75.7	84.2	79.7	77.2	81.4

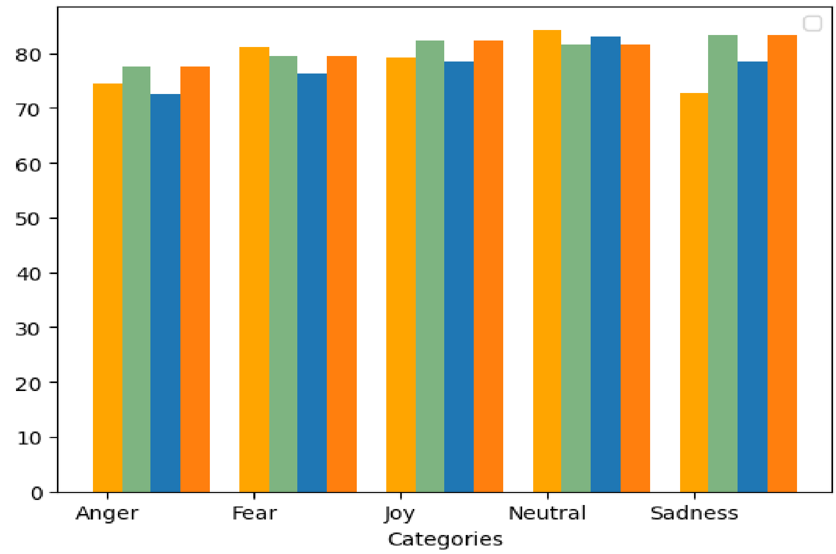

Fig 3. Evaluation Results for The Proposed Method with Different Categories.

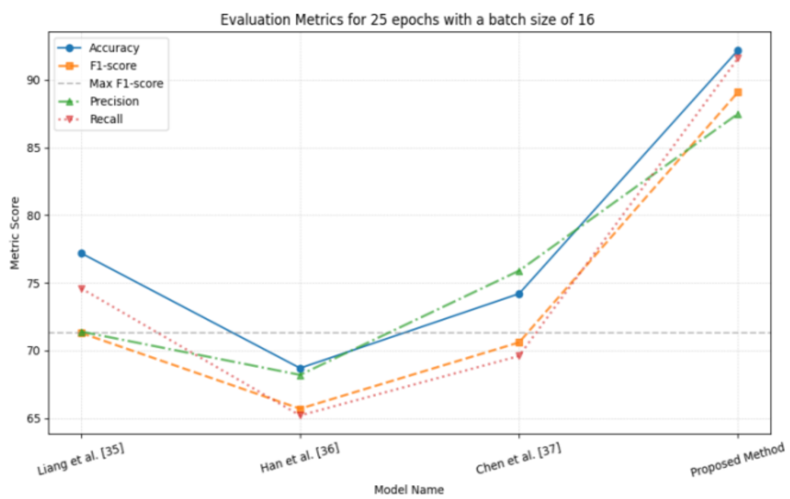

Fig 4 Performance Evaluation for The Proposed Method with 25 Epochs for Batch Size of 16.

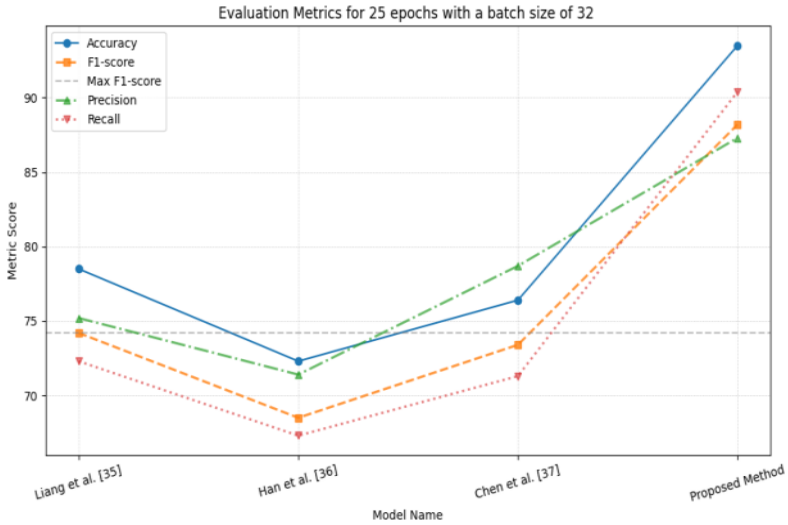

Fig 5. Performance Evaluation for The Proposed Method with 25 Epochs for Batch Size of 32.

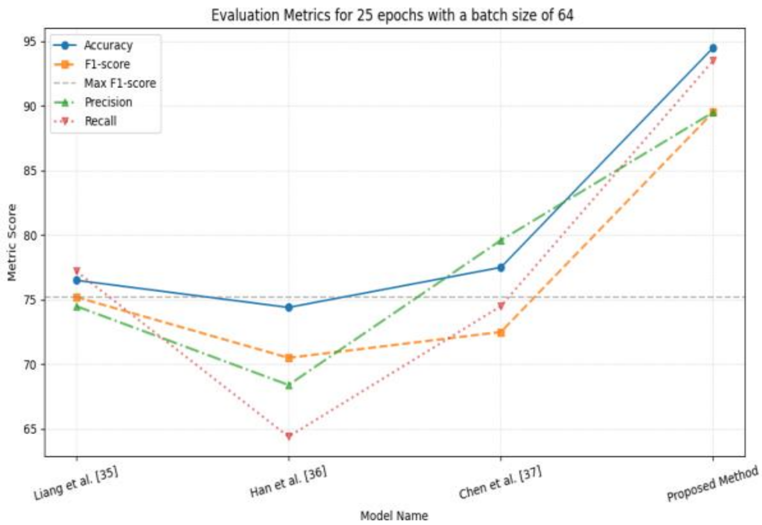

Fig 6. Performance Evaluation for The Proposed Method with 25 Epochs for Batch Size of 64.

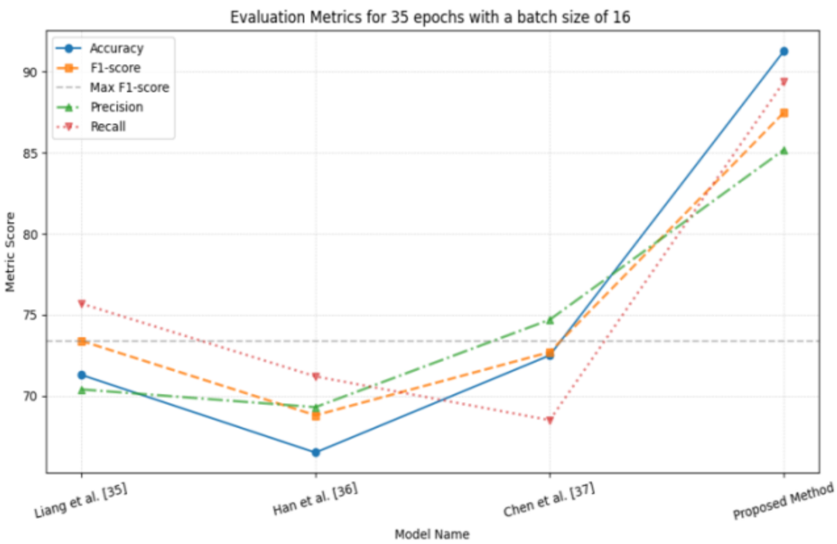

Fig 7. Performance Evaluation for The Proposed Method with 35 Epochs for Batch Size of 16.

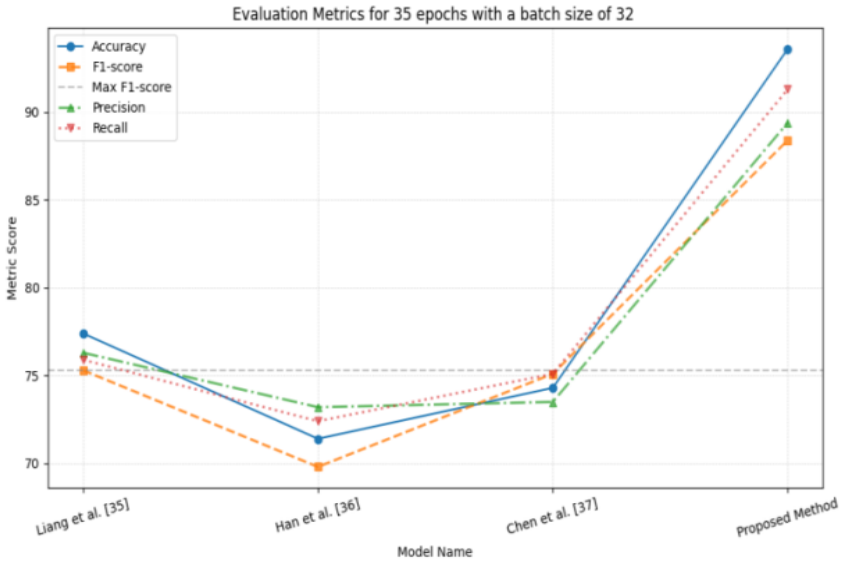

Fig 8. Performance Evaluation for The Proposed Method with 35 Epochs for Batch Size of 32.

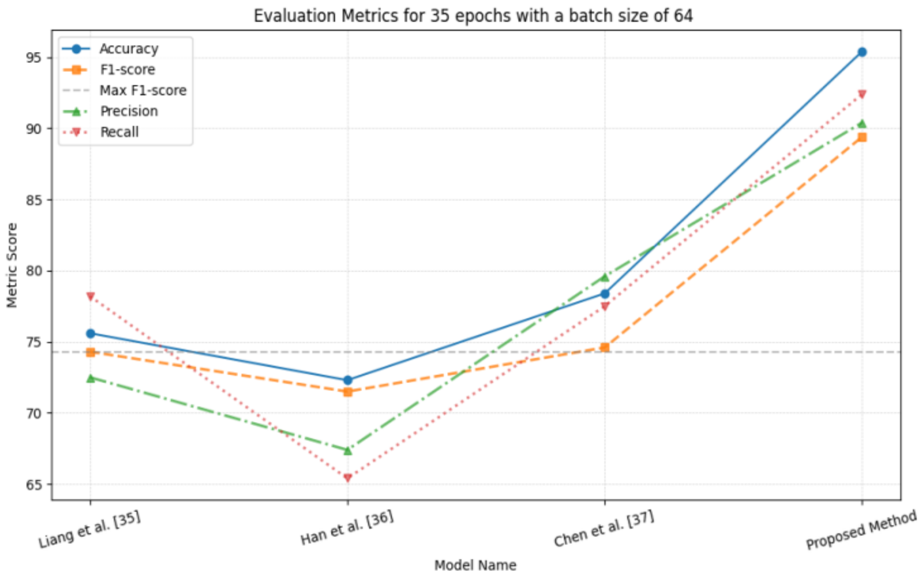

Fig 9. Performance Evaluation for The Proposed Method with 35 Epochs for Batch Size of 64.

Performance Evaluation

The performance evaluation for the proposed method is compared with the following AI recommendation systems used for E-Learning Environment discussed in Liang et al. Han et al. and Chen et al. Liang et al. presented a semi supervised method for emotion recognition using cross validation model and performance evaluation is minimum due to cross validation. Han et al. proposed a sentiment analysis method using multi model fusion technique. Chen et al. presented a multi model fusion technique for sentiment analysis using reinforcement learning method and this method accuracy rate is less compare to proposed emotion prediction method. The evaluation of other existing emotion classification approaches is measured based on 25, 35, and 45 epochs with different batch sizes of 16, 32, and 64 of dialogue instances. This evaluation is taken as an average of dropout rate within 0.3 to 0.4 to avoid overfitting in the proposed method. According to evaluation results, the proposed method achieves a maximum of 93% accuracy with the weighted average method and an F1 score of 83% **Fig 4**. Other emotion classification methods achieve less than 12% of accuracy and 14% of F1 score. The average accuracy for the proposed method is high by combining different batch sizes **Fig 7.**

Discussions

RQ1: To what extent can the proposed multi-modal data approach relying on text and video input improve the accuracy of emotion prediction compared to traditional methods?

The proposed emotion prediction system uses multi-modal information fusion approach based on text, video, and voice achieves high prediction rate by using improved artificial rabbit optimisation bio-inspired optimisation technique **Fig 5**.

The accuracy for the proposed method is improved based on reinforcement learning with agent model. The proposed agent model is designed based on the rewarding mechanism by computing Expected Reward point as given in equation (20)

Hyper-parameter	**Evaluation Metrics**	**Prediction Method for 25 Epochs**			
		Liang et al.	Han et al.	Chen et al.	Proposed Method
16 Batch Size	Accuracy	77.2	68.7	74.2	92.2
	F1 Score	71.3	65.7	70.6	89.1
	Precision	71.4	68.2	75.9	87.5
	Recall	74.6	65.2	69.6	91.6
32 Batch Size	Accuracy	78.5	72.3	76.4	93.5
	F1 Score	74.2	68.5	73.4	88.2
	Precision	75.2	71.4	78.7	87.3
	Recall	72.3	67.3	71.3	90.4
64 Batch Size	Accuracy	76.5	74.4	77.5	**94.5**
	F1 Score	75.2	70.5	72.5	**89.5**
	Precision	74.5	68.4	79.6	**89.5**
	Recall	77.2	64.4	74.5	**93.5**
Hyperparameter	**Evaluation Metrics**	**Prediction Method for 35 Epochs**			
		Liang et al.	Han et al.	Chen et al.	Proposed Method
16 Batch Size	Accuracy	71.3	66.5	72.5	91.3
	F1 Score	73.4	68.8	72.7	87.5
	Precision	70.4	69.3	74.7	85.2
	Recall	75.7	71.2	68.5	89.4
32 Batch Size	Accuracy	77.4	71.4	74.3	93.6
	F1 Score	75.3	69.8	75.1	88.4
	Precision	76.3	73.2	73.5	89.4
	Recall	75.9	72.4	75.1	91.3
64 Batch Size	Accuracy	75.6	72.3	78.4	**95.4**
	F1 Score	74.3	71.5	74.6	**89.4**
	Precision	72.5	67.4	79.6	**90.4**
	Recall	78.2	65.4	77.5	**92.4**
Hyperparameter	**Evaluation Metrics**	**Prediction Method for 45 Epochs**			
		Liang et al.	Han et al.	Chen et al.	Proposed Method
16 Batch Size	Accuracy	72.4	71.4	73.4	93.4
	F1 Score	73.5	69.7	69.5	90.5
	Precision	71.4	68.5	72.9	91.8
	Recall	73.6	72.3	71.6	94.5
32 Batch Size	Accuracy	74.6	71.5	72.7	95.4
	F1 Score	72.5	69.5	74.4	92.4
	Precision	73.3	72.5	76.5	92.4
	Recall	71.3	69.5	74.3	93.5
64 Batch Size	Accuracy	78.5	72.5	78.6	**96.3**
	F1 Score	76.3	71.6	74.5	**93.6**
	Precision	74.6	69.5	73.6	**93.4**
	Recall	79.4	70.4	71.9	**95.8**

Table 2. Performance Evaluation for The Proposed Method with A Batch Size of 32 With 25

RQ2: Does the personalised deep learning model with nature-inspired reinforcement learning lead to a more significant improvement for emotion prediction?

The proposed emotion prediction system employs reinforcement learning to learn the user's emotions based on the collected multi-modal information. Individual emotions are collected as for designing an efficient emotion prediction system for more personalised mode. Nature-inspired based reinforcement learning mechanism achieves excellent results and this can be extended for human sensitive based emotion detection.

RQ3: What are the limitations of the current deep learning-based model for predicting emotion?

The current deep learning models are using single emotion attribute for predicting the emotion and this will not provide an optimal accuracy **Fig 6**.

Deep learning has shown promise in emotion prediction, but it definitely has some hurdles to overcome such as data dependence, multi-channel of data input, black-box in nature and limited generality.

The deep learning models need massive amounts of labelled data to train on, which can be expensive and time-consuming to collect and accurate labelling can be subjective. In reality, emotions are conveyed through a combination of factors, including body language, tone of voice, and the situation leading to focus on multiple input channels and deterioration in the performance **Fig 8**. The interpretation of their results and identify biases is also a challenging task. Finally, deep learning models can be very good at recognising patterns in the data they are trained on, but they may not generalise well to unseen data. This means a model that works well on staged emotional expressions might struggle with real-world scenarios **Fig 9**.

V. CONCLUSION

This paper has proposed a novel method that predicts and suggests personalised emotions for individual users based on multi-modal sensory data collected from the brain, body, and environment. Our method uses reinforcement learning to achieve optimal performance and deliver personalised emotional experiences. Current recommendation systems often need to be improved, offering generic suggestions that fail to capture the nuances of individual preferences. This paper aims to overcome these limitations by incorporating a multi-modal data collection set throughout the day to better understand user context and intent. By analysing this data, the proposed method is able to predict the practice of the user's daily emotions. The proposed method has been trained using state-of-the-art, nature-inspired reinforcement learning algorithms and agent technology to continuously optimise its emotional recommendations. The agent has learnt to associate specific patterns in the data with specific actions, ultimately developing the ability to predict emotions with increasing accuracy. Relevant metrics have assessed the system's effectiveness and ensured it delivers tangible user benefits. The performance evaluation showed that the proposed method achieves 95.6% accuracy and 84% for F1 score. When compared to the existing AI-based emotion detection methods, the proposed method is 2 to 3% more accurate.

CRediT Author Statement

The authors confirm contribution to the paper as follows:
Conceptualization: Ezil Sam Leni A, Revathi T and Niranchana Radhakrishnan; **Methodology:** Ezil Sam Leni A and Revathi T; **Data Curation:** Niranchana Radhakrishnan; **Writing- Original Draft Preparation:** Ezil Sam Leni A and Revathi T; **Visualization:** Revathi T and Niranchana Radhakrishnan; **Investigation:** Ezil Sam Leni A, Revathi T and Niranchana Radhakrishnan; **Supervision:** Revathi T and Niranchana Radhakrishnan; **Validation:** Ezil Sam Leni A and Revathi T; **Writing- Reviewing and Editing:** Ezil Sam Leni A, Revathi T and Niranchana Radhakrishnan; All authors reviewed the results and approved the final version of the manuscript.

Data availability statement:

No data was used in this research.

Conflict of Interest:

There is no potential conflict of interest that was reported by the authors.

Funding Statement:

This research is not funded by any government or private bodies.

Competing Interests

There are no competing interests

References

[1]. P. Ye, T. Wang, and F.-Y. Wang, "A Survey of Cognitive Architectures in the Past 20 Years," IEEE Transactions on Cybernetics, vol. 48, no. 12, pp. 3280–3290, Dec. 2018, doi: 10.1109/tcyb.2018.2857704.

[2]. I. Kotseruba and J. K. Tsotsos, "40 years of cognitive architectures: core cognitive abilities and practical applications," Artificial Intelligence Review, vol. 53, no. 1, pp. 17–94, Jul. 2018, doi: 10.1007/s10462-018-9646-y.

[3]. C. Adam, W. Johal, D. Pellier, H. Fiorino, and S. Pesty, "Social Human-Robot Interaction: A New Cognitive and Affective Interaction-Oriented Architecture," Social Robotics, pp. 253–263, 2016, doi: 10.1007/978-3-319-47437-3_25.

[4]. A. Ghandeharioun, D. McDuff, M. Czerwinski, and K. Rowan, "EMMA: An Emotion-Aware Wellbeing Chatbot," 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 1–7, Sep. 2019, doi: 10.1109/acii.2019.8925455.

[5]. Barrett LF, Gross JJ. Emotional intelligence: a process model of emotion representation and regulation. 2001

[6]. M. Jeon, "Emotions and Affect in Human Factors and Human–Computer Interaction: Taxonomy, Theories, Approaches, and Methods," Emotions and Affect in Human Factors and Human-Computer Interaction, pp. 3–26, 2017, doi: 10.1016/b978-0-12-801851-4.00001-x.

[7]. R. Zall and M. R. Kangavari, "Comparative Analytical Survey on Cognitive Agents with Emotional Intelligence," Cognitive Computation, vol. 14, no. 4, pp. 1223–1246, May 2022, doi: 10.1007/s12559-022-10007-5.

[8]. E. A. Phelps, "Emotion and Cognition: Insights from Studies of the Human Amygdala," Annual Review of Psychology, vol. 57, no. 1, pp. 27–53, Jan. 2006, doi: 10.1146/annurev.psych.56.091103.070234.

[9]. R. Smith and R. D. Lane, “The neural basis of one’s own conscious and unconscious emotional states,” Neuroscience & Biobehavioral Reviews, vol. 57, pp. 1–29, Oct. 2015, doi: 10.1016/j.neubiorev.2015.08.003.

[10]. J. Ratican, J. Hutson, and D. Plate, “Synthesizing Sentience: Integrating Large Language Models and Autonomous Agents for Emulating Human Cognitive Complexity,” Journal of Artificial Intelligence, Machine Learning and Data Science, vol. 1, no. 4, pp. 135–141, Oct. 2023, doi: 10.51219/jaimld/jeremiah-ratican/17.

[11]. J. Pérez, E. Cerezo, F. J. Serón, and L.-F. Rodríguez, “A cognitive-affective architecture for ECAs,” Biologically Inspired Cognitive Architectures, vol. 18, pp. 33–40, Oct. 2016, doi: 10.1016/j.bica.2016.10.002.

[12]. J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, “An Integrated Theory of the Mind.,” Psychological Review, vol. 111, no. 4, pp. 1036–1060, 2004, doi: 10.1037/0033-295x.111.4.1036.

[13]. Belavkin RV. The role of emotion in problem solving. In Proceedings of the AISB'01 Symposium on Emotion, cognition and affective computing, Heslington, York, England. 2001, Citeseer, pp. 49–57.

[14]. I. Juvina, O. Larue, and A. Hough, “Modeling valuation and core affect in a cognitive architecture: The impact of valence and arousal on memory and decision-making,” Cognitive Systems Research, vol. 48, pp. 4–24, May 2018, doi: 10.1016/j.cogsys.2017.06.002.

[15]. C. Flavián-Blanco, R. Gurrea-Sarasa, and C. Orús-Sanclemente, “Analyzing the emotional outcomes of the online search behavior with search engines,” Computers in Human Behavior, vol. 27, no. 1, pp. 540–551, Jan. 2011, doi: 10.1016/j.chb.2010.10.002.

[16]. P. Gebhard, “ALMA,” Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems, pp. 29–36, Jul. 2005, doi: 10.1145/1082473.1082478.

[17]. R. R. McCrae and O. P. John, “An Introduction to the Five-Factor Model and Its Applications,” Journal of Personality, vol. 60, no. 2, pp. 175–215, Jun. 1992, doi: 10.1111/j.1467-6494.1992.tb00970.x.

[18]. Laird JE. The Soar cognitive architecture. MIT press. 2019.

[19]. L.-F. Rodríguez, J. O. Gutierrez-Garcia, and F. Ramos, “Modeling the interaction of emotion and cognition in Autonomous Agents,” Biologically Inspired Cognitive Architectures, vol. 17, pp. 57–70, Jul. 2016, doi: 10.1016/j.bica.2016.07.008.

[20]. Hudlicka E. Beyond cognition: modeling emotion in cognitive architectures. In ICCM. 2004, pp. 118–123.

[21]. M. S. El-Nasr, J. Yen, and T. R. Ioerger, Autonomous Agents and Multi-Agent Systems, vol. 3, no. 3, pp. 219–257, 2000, doi: 10.1023/a:1010030809960.

[22]. Ojha S, Vitale J, Williams M-A. EEGS: a transparent model of emotions. 2020. arXiv preprint. arXiv: 2011.02573.

[23]. G. Fernández-Blanco Martín et al., “An Emotional Model Based on Fuzzy Logic and Social Psychology for a Personal Assistant Robot,” Applied Sciences, vol. 13, no. 5, p. 3284, Mar. 2023, doi: 10.3390/app13053284.

[24]. E. Hudlicka, “The Case for Cognitive-Affective Architectures as Affective User Models in Behavioral Health Technologies,” Augmented Cognition. Human Cognition and Behavior, pp. 191–206, 2020, doi: 10.1007/978-3-030-50439-7_13.

[25]. A. V. Samsonovich, “Socially emotional brain-inspired cognitive architecture framework for artificial intelligence,” Cognitive Systems Research, vol. 60, pp. 57–76, May 2020, doi: 10.1016/j.cogsys.2019.12.002.