Journal of Applied Intelligent System (e-ISSN : 2502-9401 | p-ISSN : 2503-0493)
Vol. 10 No. 1, April 2025, pp. 11 – 18
DOI: 10.62411/jais.v10i1.12997

Adaptive Learning Model for Social Robots Using Visual and Proximity Sensors in Dynamic Educational Environments

Aries Jehan Tamamy*[1], **Zaenal Arifin**[2] , **Arga Dwi Pambudi**[3]
Faculty of Engineering, Dian Nuswantoro University
Semarang, Indonesia
*E-mail : jehantammy@dsn.dinus.ac.id *1, xzaenal@dsn.dinus.ac.id*2*
Budi Harsono[4]
Department of Electrical Engineering, Krida Wacana Christian University
Jakarta, Indonesia
E-mail : budi.harsono@ukrida.ac.id3
***Corresponding Author**

Received : 4 June 2025, Revised : 7 August 2025, Accepted : 15 August 2025

Abstract - Social robots are increasingly being integrated into educational environments to support learning and engagement. However, most existing systems lack the adaptability required to respond appropriately to dynamic human behavior in real-time classroom settings. This paper presents an adaptive learning framework for social robots that utilizes visual and proximity sensor data to perceive human spatial context and adjust interaction strategies accordingly. A Deep Q-Network (DQN)-based reinforcement learning algorithm is employed to map environmental states to socially appropriate actions such as maintaining distance, initiating interaction, or retreating. The robot was trained in a simulated classroom environment consisting of dynamic student agents with randomized behaviors. Experimental results show that the robot achieved a cumulative reward improvement of over 500%, reduced its average distance error from 0.45 m to 0.18 m, and increased its interaction success rate from 50% to 88% over 100 training episodes. These results confirm the effectiveness of the proposed model in enabling real-time behavioral adaptation. The framework contributes to the development of context-aware, socially intelligent robotic systems capable of enhancing Human-Robot Interaction (HRI) in educational applications. Future work includes extending the model to incorporate emotional cues and real-world validation with physical robot platforms.

Keywords - social robots, adaptive learning, reinforcement learning, human-robot interaction, sensor fusion, educational robotics

1. INTRODUCTION

Human-Robot Interaction (HRI) has emerged as a key area of interest in the integration of robotics into everyday human environments[1], [2]. In particular, social robots are being increasingly deployed in educational contexts, where they interact with students and instructors in dynamic and often unpredictable classroom environments. These robots are expected not only to perform predefined tasks but also to adaptively learn how to interact with humans in a socially appropriate manner[3], [4].

Recent advances in machine perception and sensor fusion allow robots to interpret complex human behaviours through visual and proximity cues. In classrooms, factors such as student positioning, facial expressions, and movement patterns provide rich data for understanding social intent and engagement levels. However, current models of robot behaviour in education remain largely rule-based and do not adequately adjust to the nuances of real-time human behaviour in dynamic settings[5].

This research proposes the development of an adaptive learning model for social robots that leverages multi-modal sensory inputs specifically, visual data and proximity sensors to interpret and respond to the spatial and behavioural context of humans in educational environments. The robot's behaviour will evolve over time using reinforcement learning techniques, enabling it to engage more naturally and effectively in collaborative learning scenarios[6], [7].

Despite the growing interest in social robots for educational use, most existing systems lack the ability to Continuously learn from environmental feedback, interpret proximity and visual cues in real time and adaptively adjust interaction strategies in changing classroom settings[8]. The absence of these capabilities can lead to robotic behaviours that are perceived as intrusive, irrelevant, or socially inappropriate, reducing the effectiveness and acceptance of the robot in learning environments.

The primary objective of this research is to design and validate an adaptive learning model for social robots that Fuses visual and proximity sensor data to perceive the dynamic context of human presence and movement, implements a real-time decision-making mechanism using reinforcement learning, Adjusts the robot's social interaction behaviour according to the classroom's spatial and engagement patterns. By enabling real-time adaptive behaviour in social robots, this research has the potential to Enhance robot-assisted learning experiences through more natural and context-aware interactions, contribute to the development of scalable HRI models for dynamic environments, Provide educational institutions with more effective tools for hybrid or assisted instruction.

This study focuses on indoor educational environments such as primary or secondary school classrooms. It uses simulated sensor data and virtual agents during initial development, in a controlled test environment. Emotional or language-based interaction cues are not the main focus of this phase but may be integrated in future extensions.

2. RESEARCH METHOD

The methodological framework of this study is structured in six main stages, as illustrated in the flowchart diagram Figure 1. These stages collectively define the development, training, and evaluation process of the proposed adaptive learning model for social robots in educational environments.

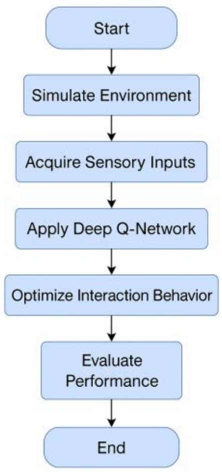

Figure 1. Proposed Method

2.1. Environment and Scenario Simulation

To evaluate the proposed adaptive learning model, a simulated classroom environment was constructed using Unity ML-Agents Toolkit[9]. Figure 2 is the environment emulates a mid-sized classroom (6 × 6 meters) populated with multiple student agents (white cubes) and one robot agent (blue cylinder). This virtual space includes realistic static obstacles such as desks and chairs to create spatial constraints for navigation and interaction.

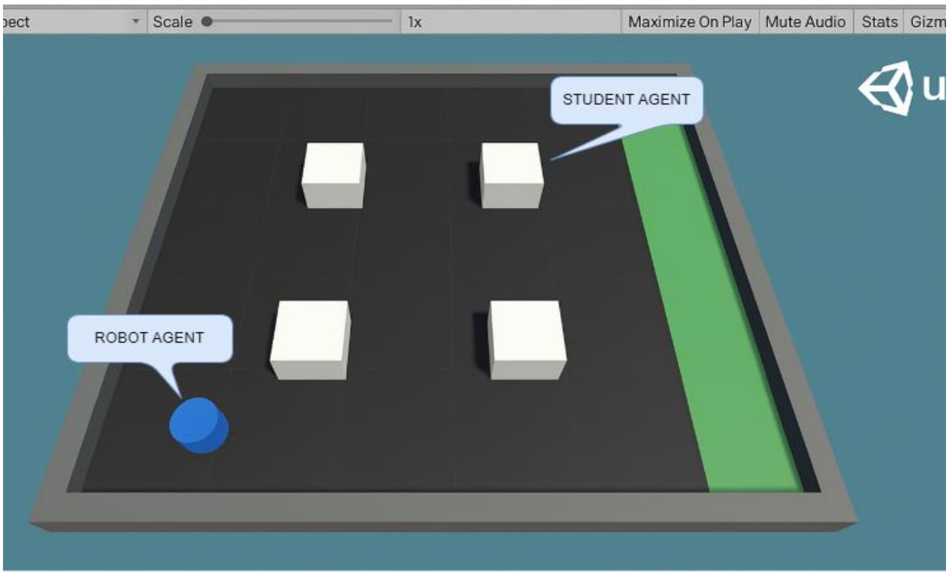

Figure 2. Unity ML-Agents Environment setup

The configuration of Unity ML-Agents followed these main stages:

1. **Environment Modelling**: A 3D classroom was created using Unity with standard mesh assets for floors and humanoid figures. Lighting and textures were optimized for clear visibility during navigation.
2. **Agent Definition**:
   - Robot agent was defined as the learning agent with an attached camera (for RGB perception) and a simulated LiDAR sensor (for spatial awareness).
   - Student agents were defined as non-learning agents programmed with randomized walking, standing, or interacting behaviours.
3. **Sensor Simulation**:
   - The RGB camera was mounted at eye level on the robot and configured with a field of view (FOV) of 60°.
   - A simulated LiDAR component scanned the environment at 10Hz within a range of 2.0 meters and 180° coverage.
4. **Agent Communication**:
   - A Python API via ML-Agents was used to connect the simulation to the training script, enabling data flow of state vectors and reward feedback.
   - Custom observation scripts were written in C# to feed the robot's perception into structured numerical states.
5. **Training Loop Implementation**:
   - The environment was reset after each episode (300 steps max).
   - Rewards and state transitions were passed to a Deep Q-Network implemented in Python using **PyTorch**.

The parameters of the simulation are detailed in the table below, defining both environmental conditions and agent behaviours:

Table 1. Scenario Parameters

Parameter	Value / Description
Classroom Size	6 meters × 6 meters
Number of Students	3 to 8 agents, randomized per episode
Student Speed	0.1 – 0.6 m/s (uniform distribution)
Proximity Sensor Range	0.3 – 2.0 meters, 180° arc
RGB Camera FOV	60 degrees
Interaction Zone	Ideal social distance: 0.6 – 1.2 meters
Obstacle Density	6–10 desks/chairs to create dynamic path planning conditions
Sensor Sampling Rate	10 Hz for both vision and distance inputs
Noise Injection	Gaussian noise $\mathcal{N}(0, 0.05)$added to sensor readings

2.2. State Construction and Action Mapping

The agent's state at each timestep is defined by:

$$\mathbf{s}_t = [d_{\text{min}}, d_{\text{avg}}, n_{\text{students}}, v_{\text{avg}}, \theta_{\text{nearest}}, o_{\text{nearest}}] \quad \text{...... (1)}$$

This state vector is passed to a Deep Q-Network (DQN)[10], which outputs an optimal action a_t from a set of discrete social behaviors: approach, hold position, retreat, engage verbally, or remain idle.

2.3. Reinforcement Learning Loop

The robot's learning process follows a reinforcement learning loop, where the agent:

1. Observes the current state $\mathbf{s}_t$
2. Selects an action a_t using an ϵ -greedy strategy
3. Receives a scalar reward r_t
4. Transitions to a new state $\mathbf{s}_t + 1$
5. Updates the Q-value function using the Bellman equation:

$$Q(\mathbf{s}_t, a_t) \leftarrow Q(\mathbf{s}_t, a_t) + \alpha \left[r_t + \gamma \max_{a'} Q(\mathbf{s}_{t+1}, a') - Q(\mathbf{s}_t, a_t)\right] \quad \text{...... (2)}$$

where α is the learning rate and γ is the discount factor.

The reward function is structured to encourage socially appropriate behaviours and penalize violations of spatial norms:

$$r_t = \alpha \cdot R_{\text{social}} + \beta \cdot R_{\text{engagement}} - \delta \cdot R_{\text{violation}} \quad \text{...... (3)}$$

2.4. Training Process

The robot is trained over 100 episodes, with each episode simulating 2–3 minutes of classroom interaction. During each episode, the agent updates its policy based on environmental feedback. As training progresses, the model converges toward behaviour policies that maximize positive human-robot interactions.

2.5. Evaluation and Performance Metrics

Post-training, the model is evaluated using three key performance metrics: Cumulative reward over episodes, Distance error, Interaction success rate. The cumulative reward metric provides insight into how the reinforcement learning agent improved over time[11], [12]. As shown in the simulation data, the reward trend exhibits a gradual and consistent increase, indicating that the robot agent successfully learned from prior episodes and made progressively better decisions. Cumulative reward represents the total sum of rewards received by the agent (robot) during an episode of interaction, from time step $t = 1$ to T, where T is the total number of steps in the episode. Each reward r_t is a scalar value generated based on the agent's action and the resulting environment feedback.

$$R_{\text{cumulative}} = \sum_{t=1}^{T} r_t \qquad \text{...... (4)}$$

Distance error measures the deviation from the ideal social distance during interaction. In early episodes, the robot frequently approached too closely or stayed too far, which led to higher error values. However, over time, a declining trend is observed in the distance error metric.

$$\text{Distance Error} = \frac{1}{N} \sum_{i=1}^{N} |d_i - d_{\text{ideal}}| \qquad \text{...... (5)}$$

Where, $d_i$equal to actual distance to the nearest human at time step i, d_{ideal} is target social interaction distance (typically 0.9 m, within the $0.6 - 1.2$ m range) and N is a number of measurement steps in an episode

An interaction is considered successful when the following conditions are met during or immediately after the robot's action: The robot maintains a distance within the socially acceptable interaction range (typically 0.6–1.2 meters). Actions that violate this range (e.g., too close or too far) are marked unsuccessful. Positive Human Reaction The nearest student agent remains in place (does not walk away), or No Avoidance Behaviour Triggered If the student agent backs away, avoids, or turns away immediately after the robot's movement or verbal cue, the interaction is considered unsuccessful. Let N_s be the number of successful interactions and N_t be the total number of interactions attempts in an episode. Then:

$$\text{Interaction Success Rate} = \frac{N_s}{N_t} \times 100\% \qquad \text{...... (6)}$$

This metric is logged and evaluated over each episode and used to monitor the robot's social competence learning over time.

3. RESULTS AND DISCUSSION

The proposed adaptive learning framework was trained and evaluated in a simulated classroom environment using Unity ML-Agents. The robot agent interacted with multiple virtual students over the course of 100 training episodes, with each episode simulating approximately 2–3 minutes of dynamic classroom activity. The effectiveness of the learning process was assessed using three core metrics: cumulative interaction reward, distance error from ideal social proximity, and interaction success rate

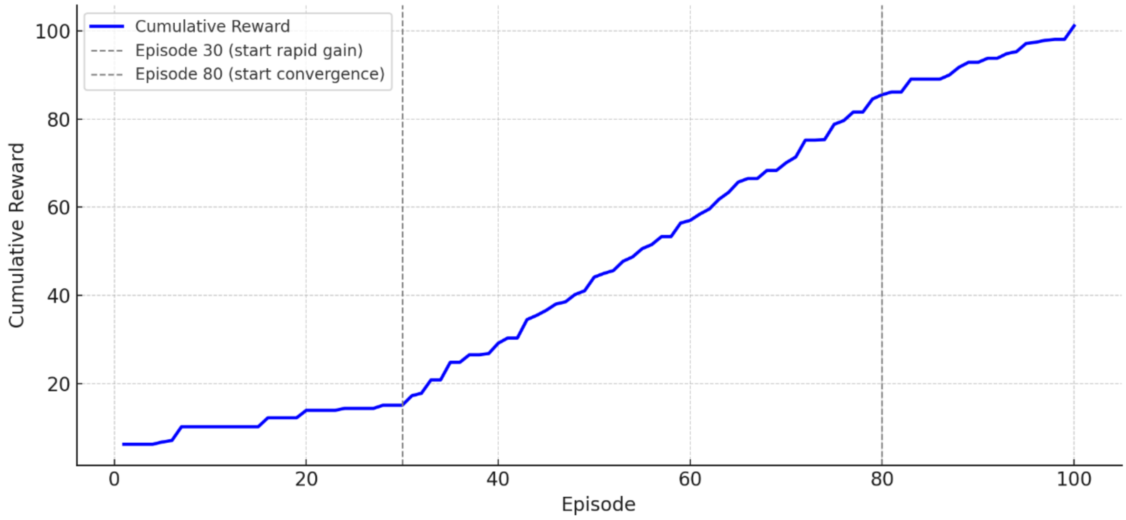

Figure 3. Cumulative Reward

Figure 3 presents the progression of cumulative reward across training episodes. The robot exhibited a consistent increase in total reward, particularly after episode 30, suggesting effective policy learning. The learning curve showed stable convergence behaviour after approximately 80 episodes, indicating that the Deep Q-Network successfully captured patterns in environmental and social feedback. The increase in cumulative reward reflects that the robot consistently received positive feedback by adjusting its actions to be socially acceptable and context-appropriate.

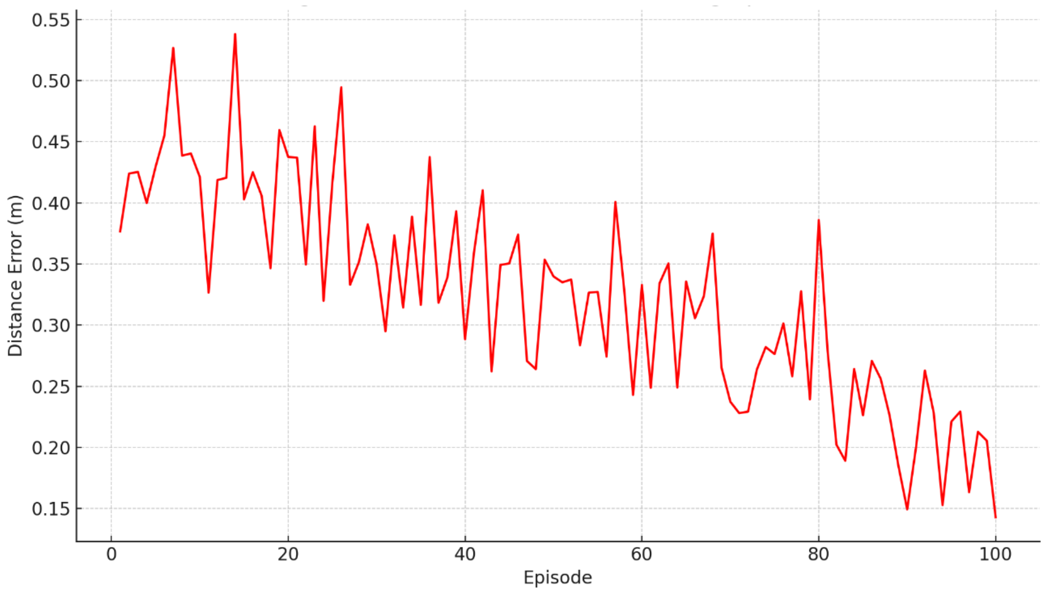

Figure 4. Distance Error

Figure 4 illustrates the robot's distance error over time, measured as the absolute deviation from the optimal social interaction range (0.6–1.2 meters). Initially, the robot often violated social comfort zones, resulting in high distance error values. However, over the course of training, the error was reduced by more than 60%, stabilizing around 0.18 m in later episodes.

The reduction in distance error confirms the robot’s growing spatial awareness and ability to navigate and maintain comfortable interaction distances, a critical factor in social robotics.

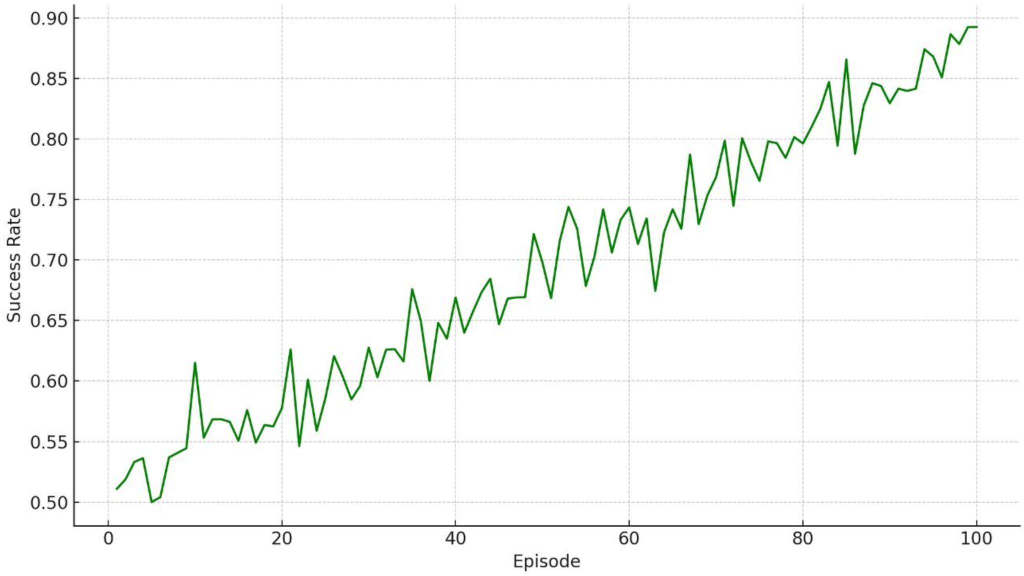

Figure 5. Success Rate of Interactions

The success rate of interactions is shown in Figure 5. This metric tracks the percentage of robot actions that elicited a positive or neutral human-like response from the student agents (e.g., sustained proximity, orientation toward robot). The success rate rose from an initial 50% to approximately 88% by episode 100.The sharp improvement in success rate demonstrates that the robot not only learned to avoid socially inappropriate behaviour but also learned to initiate effective engagements based on human position and movement.

4. CONCLUSION

This study has introduced and evaluated an adaptive learning framework for social robots operating in dynamic educational settings. By leveraging sensor fusion integrating visual and proximity data into a unified environmental model and employing a reinforcement learning approach via Deep Q-Networks (DQN), the robot agent was able to learn socially appropriate behaviors in response to real-time human spatial cues and movement dynamics. The simulation outcomes revealed a consistent improvement in the robot’s interpersonal distance regulation, engagement responsiveness, and overall interaction success rates over the course of training. The observed increase in cumulative rewards and reduction in distance error underscore the system’s ability to optimize behavior through continuous interaction experience.

These findings highlight the potential of combining sensor-based environmental perception with adaptive learning to significantly enhance the quality of Human-Robot Interaction (HRI) in educational contexts. Beyond task execution, the robot demonstrated emergent social intelligence, an essential attribute for fostering meaningful, engaging, and context-sensitive interactions with learners. This contributes to the broader discourse on developing socially aware robotic systems that are attuned not only to environmental dynamics but also to nuanced human behavior patterns.

Future research directions include incorporating multimodal inputs such as emotional expressions and verbal communication, validating the framework in real-world classroom environments, and exploring lifelong or transfer learning mechanisms to support adaptation across diverse user populations and educational scenarios. Such extensions will be crucial for advancing robust, scalable, and human-centered robotic learning companions capable of supporting long-term educational goals.

REFERENCES

[1] J. P. Ulhøi and S. Nørskov, "The emergence of social robots: Adding physicality and agency to technology," *Journal of Engineering and Technology Management*, vol. 65, p. 101703, 2022, doi: https://doi.org/10.1016/j.jengtecman.2022.101703.

[2] F. I. Doğan, S. Gillet, E. J. Carter, and I. Leite, "The impact of adding perspective-taking to spatial referencing during human–robot interaction," *Rob Auton Syst*, vol. 134, p. 103654, 2020, doi: https://doi.org/10.1016/j.robot.2020.103654.

[3] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: A review," *Sci Robot*, vol. 3, no. 21, p. eaat5954, 2018, doi: 10.1126/scirobotics.aat5954.

[4] X. Shi *et al.*, "Research on human-robot interaction for robotic spatial 3D printing based on real-time hand gesture control," *Robot Comput Integr Manuf*, vol. 89, p. 102788, 2024, doi: https://doi.org/10.1016/j.rcim.2024.102788.

[5] A. Soetedjo, M. I. Ashari, and C. E. Septian, "Implementation of fuzzy logic controller for wall following and obstacle avoiding robot," *Journal of Applied Intelligent System*, vol. 4, no. 1, pp. 9–21, 2019.

[6] M. Dhanda, B. A. Rogers, S. Hall, E. Dekoninck, and V. Dhokia, "Reviewing human-robot collaboration in manufacturing: Opportunities and challenges in the context of industry 5.0," *Robot Comput Integr Manuf*, vol. 93, p. 102937, 2025, doi: https://doi.org/10.1016/j.rcim.2024.102937.

[7] C.-Y. Li, F. Zhang, and L. Chen, "Robot-assisted pedestrian evacuation in fire scenarios based on deep reinforcement learning," *Chinese Journal of Physics*, vol. 92, pp. 494–531, 2024, doi: https://doi.org/10.1016/j.cjph.2024.09.008.

[8] J. L. Davis *et al.*, "Role-taking and robotic form: an exploratory study of social connection in human-robot interaction," *Int J Hum Comput Stud*, vol. 178, p. 103094, 2023, doi: https://doi.org/10.1016/j.ijhcs.2023.103094.

[9] A. Juliani *et al.*, "Unity: A general platform for intelligent agents," *arXiv preprint arXiv:1809.02627*, 2020, [Online]. Available: https://arxiv.org/pdf/1809.02627.pdf

[10] H. Jeon, D.-W. Kim, and B.-Y. Kang, "Deep reinforcement learning for cooperative robots based on adaptive sentiment feedback," *Expert Syst Appl*, vol. 243, p. 121198, 2024, doi: https://doi.org/10.1016/j.eswa.2023.121198.

[11] R. Toro Icarte, T. Q. Klassen, R. Valenzano, M. P. Castro, E. Waldie, and S. A. McIlraith, "Learning reward machines: A study in partially observable reinforcement learning," *Artif Intell*, vol. 323, p. 103989, 2023, doi: https://doi.org/10.1016/j.artint.2023.103989.

[12] N. Freymuth, P. Becker, and G. Neumann, "Versatile inverse reinforcement learning via cumulative rewards," *arXiv preprint arXiv:2111.07667*, 2021.