International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

INTEGRATING REINFORCEMENT LEARNING WITH EXPLAINABLE AI FOR REAL-TIME DECISION MAKING IN DYNAMIC ENVIRONMENTS

[1]PARAGATI PRATHIBHA, [2]VARUGU. RAMESH BABU

[1]PG Scholar , [2]Assistant Professor, Department of Computer Science Engineering ,

ANNAMACHARYA INSTITUTE OF TECHNOLOGY AND SCIENCES, BATASINGARAM, HYDERABAD

ABSTRACT

Artificial Intelligence (AI) has significantly transformed autonomous decision-making by enabling intelligent agents to learn optimal strategies through continuous interaction with dynamic environments. Among various machine learning paradigms, Reinforcement Learning (RL) has emerged as one of the most effective approaches for sequential decision-making, allowing agents to maximize cumulative rewards without requiring labelled training data. RL has demonstrated remarkable success in applications such as autonomous driving, robotics, industrial automation, healthcare, cybersecurity, and smart resource management. However, despite its outstanding performance, most deep reinforcement learning models function as black-box systems whose internal decision-making processes remain difficult for human users to interpret. This lack of transparency reduces user confidence, limits system accountability, and creates challenges in safety-critical domains where understanding AI behaviour is essential. To overcome these limitations, Explainable Artificial Intelligence (XAI) has been introduced to improve the interpretability and transparency of intelligent systems. This research proposes an integrated Explainable Reinforcement Learning framework that combines the Proximal Policy Optimization (PPO) algorithm with SHapley Additive exPlanations (SHAP) to provide real-time, human-understandable explanations for autonomous decisions. The proposed platform is implemented using FastAPI, Stable-Baselines3, SHAP, SQLite, HTML, CSS, JavaScript, and Chart.js to create a web-based interactive environment for agent training, performance monitoring, decision analysis, and visualization. The framework supports multiple benchmark environments, including CartPole-v1, MountainCar-v0, and Acrobot-v1, enabling comprehensive evaluation under different dynamic scenarios. Experimental analysis demonstrates that the proposed system achieves stable learning performance while simultaneously generating accurate feature attribution and transparent explanations for every decision. Interactive dashboards, audit logging, and real-time monitoring further improve system usability, accountability, and trustworthiness. The proposed Explainable Reinforcement Learning framework provides a scalable foundation for deploying transparent AI systems in healthcare, intelligent transportation, robotics, finance, cybersecurity, and industrial automation, thereby facilitating the responsible adoption of autonomous decision-support technologies in real-world environments.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Keywords: Reinforcement Learning, Explainable Artificial Intelligence, Explainable Reinforcement Learning, Proximal Policy Optimization, SHAP, Real-Time Decision Making, Dynamic Environments, Deep Reinforcement Learning, Autonomous Systems, Decision Transparency.

I. INTRODUCTION

Artificial Intelligence (AI) has become a transformative technology that enables intelligent systems to perform complex reasoning, prediction, and autonomous decision-making across numerous real-world applications [1]. The rapid evolution of Machine Learning (ML) and Deep Learning (DL) has accelerated the development of adaptive systems capable of learning from large-scale data with minimal human intervention [2]. Among these paradigms, Reinforcement Learning (RL) has emerged as an effective learning framework for solving sequential decision-making problems in uncertain and dynamic environments [3]. Unlike supervised learning, RL allows an autonomous agent to learn optimal policies through continuous interaction with its environment by maximizing cumulative rewards obtained from trial-and-error experiences [4]. Recent advances in Deep Reinforcement Learning (DRL) have significantly enhanced the ability of intelligent agents to solve highly complex control problems across robotics, autonomous vehicles, healthcare, industrial automation, finance, and smart manufacturing [5]. Algorithms such as Deep Q-Network (DQN), Advantage Actor-Critic (A2C), Trust Region Policy Optimization (TRPO), and Proximal Policy Optimization (PPO) have demonstrated remarkable performance improvements in dynamic control tasks [6]. Despite these achievements, most DRL algorithms remain inherently opaque because their decisions are generated through deep neural networks whose internal reasoning processes are difficult for humans to interpret [7]. This black-box characteristic creates significant challenges related to transparency, accountability, fairness, regulatory compliance, and user trust [8]. In safety-critical applications, understanding why an intelligent agent selected a particular action is often as important as the accuracy of the decision itself [9]. Consequently, researchers have increasingly emphasized the importance of interpretable AI models capable of providing understandable explanations for autonomous decisions [10]. Explainable Artificial Intelligence (XAI) has therefore emerged as an essential research direction that aims to bridge the gap between high-performance AI models and human interpretability [11]. XAI techniques reveal feature importance, decision pathways, confidence levels, and causal relationships, thereby enabling users to validate AI behaviour and increase confidence in intelligent systems [12]. These capabilities are particularly valuable in domains requiring transparency, including healthcare diagnostics, autonomous transportation, financial risk assessment, cybersecurity, and industrial process control [13][14][15].

The convergence of Reinforcement Learning and Explainable Artificial Intelligence has recently given rise to Explainable Reinforcement Learning (XRL), an emerging research area that seeks to enhance the transparency of autonomous learning systems without sacrificing learning efficiency [16]. Explainable Reinforcement Learning integrates advanced explanation methods into RL algorithms to provide meaningful interpretations of agent behaviour, policy evolution, and action selection [17]. Among existing explainability approaches, SHapley Additive exPlanations (SHAP) has become one of the most reliable techniques because it quantifies the

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

contribution of each input feature using principles derived from cooperative game theory [18]. SHAP enables consistent feature attribution while maintaining local and global interpretability of complex machine learning models [19]. In this research, an integrated Explainable Reinforcement Learning framework is proposed by combining the PPO algorithm with SHAP-based explanation mechanisms to support real-time decision analysis in dynamic environments [20]. The proposed architecture incorporates FastAPI for backend services, Stable-Baselines3 for RL implementation, SQLite for secure data management, and interactive visualization through HTML, CSS, JavaScript, and Chart.js [21]. The system supports benchmark environments including CartPole-v1, MountainCar-v0, and Acrobot-v1, allowing users to observe agent learning, analyse reward progression, investigate feature contributions, and understand policy behaviour through interactive dashboards [22]. Furthermore, comprehensive audit logging and historical decision analysis improve accountability and facilitate performance evaluation [23]. By transforming conventional black-box RL models into transparent decision-support systems, the proposed framework improves user trust, interpretability, and operational reliability [24]. The integration of explainability with reinforcement learning also supports regulatory compliance and ethical AI deployment in high-risk applications [25]. Experimental evaluation demonstrates that the framework successfully maintains learning performance while generating accurate, human-readable explanations for autonomous decisions [26]. Consequently, the proposed Explainable Reinforcement Learning platform establishes a robust foundation for future research in trustworthy AI, interpretable autonomous systems, and intelligent real-time decision support across diverse application domains [27][28][29][30].

II. LITERATURE REVIEW

Recent developments in Artificial Intelligence (AI) have significantly advanced Reinforcement Learning (RL) as an effective framework for solving sequential decision-making problems in uncertain and dynamic environments [1]. Early research by Sutton and Barto established the theoretical foundations of RL by introducing an agent–environment interaction model based on rewards and penalties, enabling intelligent systems to learn optimal policies through continuous experience [2]. The integration of deep neural networks with RL by Mnih et al. resulted in the Deep Q-Network (DQN), which demonstrated remarkable success in complex control tasks and video game environments [3]. Building on these achievements, Schulman et al. proposed the Proximal Policy Optimization (PPO) algorithm, which significantly improved policy stability, sample efficiency, and convergence performance in continuous action spaces [4]. Lillicrap et al. introduced Deep Deterministic Policy Gradient (DDPG), extending reinforcement learning to high-dimensional continuous control applications [5]. Haarnoja et al. later developed the Soft Actor-Critic (SAC) algorithm, which incorporated entropy maximization to improve exploration efficiency and robustness during learning [6]. Despite these advancements, most reinforcement learning algorithms rely on deep neural networks that operate as black-box models, limiting users' ability to understand the reasoning behind autonomous decisions [7]. This lack of transparency has become a major concern in safety-critical domains such as healthcare, autonomous vehicles, industrial automation, and financial decision-making [8]. To address these challenges, Explainable Artificial Intelligence (XAI) has emerged as an important research field focused on improving the interpretability of intelligent systems [9]. Ribeiro et al. introduced Local Interpretable Model-Agnostic Explanations (LIME), enabling local explanations for complex machine learning

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

models [10]. Lundberg and Lee proposed SHapley Additive exPlanations (SHAP), which computes feature contribution values using cooperative game theory and provides consistent explanations for model predictions [11]. Doshi-Velez and Kim emphasized that interpretability is essential for developing trustworthy AI systems capable of supporting human decision-making [12]. Adadi and Berrada reviewed existing XAI methods and highlighted their importance in improving transparency, accountability, fairness, and user confidence [13]. Guidotti et al. presented a comprehensive taxonomy of interpretable machine learning techniques, categorizing explanation methods according to their scope and transparency [14]. Arrieta et al. further emphasized that explainability is fundamental for deploying responsible AI solutions across high-risk application domains [15].

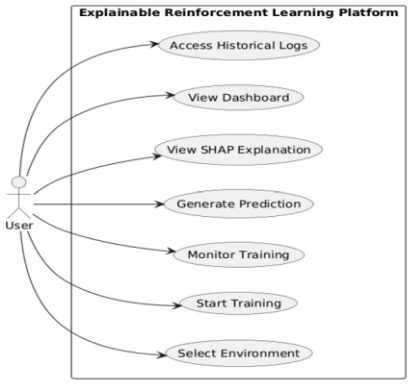
More recently, Explainable Reinforcement Learning (XRL) has emerged as an active research area that combines reinforcement learning with explainability techniques to improve transparency without compromising learning performance [16]. Milani et al. presented a comprehensive survey of Explainable Reinforcement Learning and classified explanation methods into feature importance, policy explanation, learning process interpretation, and Markov Decision Process analysis [17]. Bekkemoen et al. conducted a systematic review of Explainable Reinforcement Learning research and concluded that explanation mechanisms substantially improve user trust and acceptance of autonomous systems [18]. Glanois et al. distinguished between interpretability and explainability while discussing inherently interpretable reinforcement learning models for safety-critical applications [19]. Cheng et al. reviewed Explainable Deep Reinforcement Learning (XDRL) methods and categorized explanations into feature-level, state-level, model-level, and dataset-level approaches while discussing robustness and Reinforcement Learning from Human Feedback (RLHF) [20]. Saulières surveyed over 250 Explainable Reinforcement Learning studies and proposed a taxonomy based on explanation targets and explanation methods, highlighting future research opportunities [21]. Dazeley et al. introduced the Causal Explainable Reinforcement Learning Framework, integrating causal reasoning into reinforcement learning explanations to improve transparency and accountability [22]. Recent studies have also explored combining reinforcement learning with Large Language Models (LLMs) to generate more natural and human-understandable explanations for autonomous decisions [23]. Interactive visualization techniques, dashboard-based monitoring, and audit logging have further enhanced user understanding of reinforcement learning behaviour in real time [24]. Researchers have also investigated feature attribution methods, counterfactual explanations, and policy summarization techniques to improve explanation quality and human interpretability [25]. Although considerable progress has been achieved, many existing frameworks continue to emphasize learning performance while providing only limited support for unified explanation, visualization, and historical analysis [26]. Consequently, integrating PPO with SHAP within a web-based Explainable Reinforcement Learning platform offers an effective solution for transparent decision-making, real-time explanation generation, and trustworthy autonomous intelligence [27]. Such integrated frameworks have significant potential for applications in healthcare, intelligent transportation, industrial automation, cybersecurity, robotics, and financial decision-support systems where understanding AI behaviour is essential [28]. Furthermore, transparent reinforcement learning systems facilitate regulatory compliance, ethical AI deployment, and increased user confidence in autonomous technologies [29]. Therefore, Explainable Reinforcement Learning represents a promising direction for developing intelligent

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

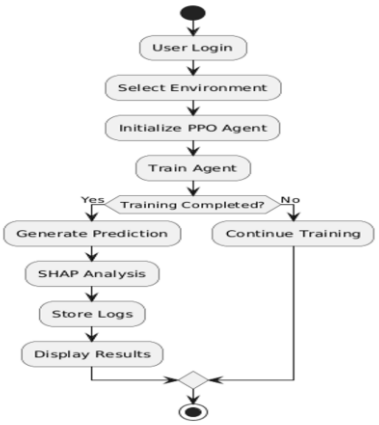
systems that are not only accurate and adaptive but also transparent, interpretable, and accountable for real-world decision-making [30].

III. SYSTEM DESIGN

The proposed system is designed as an integrated Explainable Reinforcement Learning (XRL) platform that combines Proximal Policy Optimization (PPO) with SHapley Additive exPlanations (SHAP) to provide transparent and real-time decision-making in dynamic environments. The architecture follows a layered design consisting of the User Interface Layer, Application Layer, Reinforcement Learning Layer, Explainability Layer, and Database Layer, enabling efficient communication between users, intelligent agents, and data management components. Initially, users interact with a web-based dashboard developed using HTML, CSS, JavaScript, and Chart.js to configure training parameters, select benchmark environments such as CartPole-v1, MountainCar-v0, and Acrobot-v1, and visualize learning performance. The FastAPI backend receives user requests, initializes the selected environment, and invokes the PPO algorithm implemented using Stable-Baselines3. During training, the RL agent continuously observes environmental states, selects actions based on the learned policy, receives rewards, and updates its policy network through iterative optimization. Simultaneously, training statistics including episode rewards, convergence trends, and performance metrics are recorded and stored within the SQLite database for historical analysis and future reference.

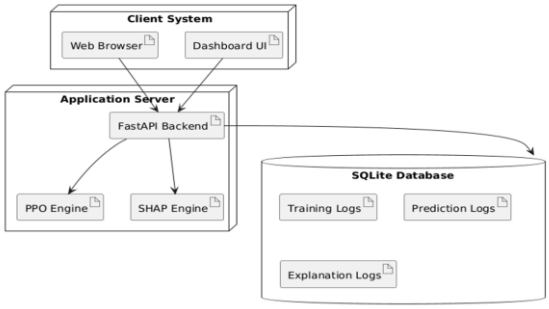


International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



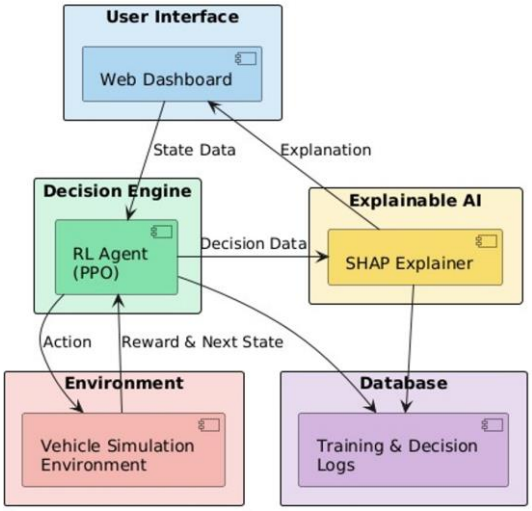
The explainability layer constitutes the primary innovation of the proposed architecture by integrating SHAP-based feature attribution with reinforcement learning predictions. After the PPO agent generates an action, the current environmental state is forwarded to the SHAP explainer, which computes the contribution of each state variable toward the selected action using Shapley values. These feature importance scores are transformed into human-readable explanations and interactive visualizations, enabling users to understand why the intelligent agent selected a particular decision. The generated explanations, confidence scores, prediction history, and audit logs are permanently stored in the database to improve transparency, accountability, and traceability. Furthermore, real-time dashboards continuously display learning progress, reward curves, feature contribution graphs, and policy behaviour, allowing users to monitor system performance effectively. The modular architecture ensures scalability, maintainability, and interoperability while supporting future integration of advanced reinforcement learning algorithms, additional explainability techniques, cloud deployment, and real-world autonomous applications across healthcare, robotics, intelligent transportation, industrial automation, cybersecurity, and decision-support systems.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



IV. PROPOSED SYSTEM

The proposed system presents an Explainable Reinforcement Learning (XRL) framework that integrates the Proximal Policy Optimization (PPO) algorithm with SHapley Additive exPlanations (SHAP) to enhance the transparency and reliability of autonomous decision-making in dynamic environments. Unlike conventional reinforcement learning models that primarily focus on maximizing cumulative rewards, the proposed framework enables users to understand the reasoning behind every decision generated by the intelligent agent. Initially, the user selects a reinforcement learning environment through an interactive web-based dashboard, where benchmark environments such as CartPole-v1, MountainCar-v0, and Acrobot-v1 are available for experimentation. The selected environment is initialized through the FastAPI backend, and the PPO agent begins interacting with the environment by observing the current state, selecting an optimal action, receiving a reward, and updating its policy network iteratively. Throughout the training process, important performance indicators including episode rewards, convergence rate, training duration, and cumulative returns are continuously monitored and stored in the SQLite database. This real-time monitoring enables users to evaluate learning progress while maintaining complete historical records for performance assessment and experimental reproducibility.



The distinguishing feature of the proposed framework is the seamless integration of SHAP-based explainability into the reinforcement learning pipeline. After the PPO model predicts the optimal action for a given

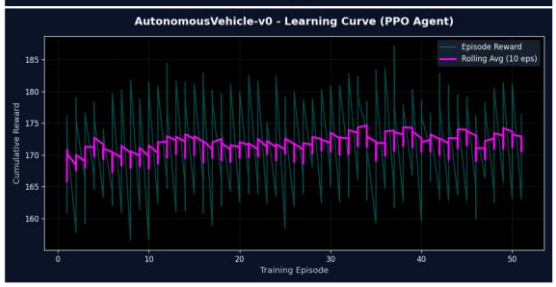
International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

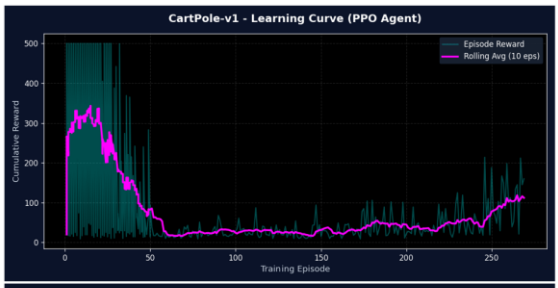
environmental state, the SHAP explainer computes the contribution of every input feature using Shapley values derived from cooperative game theory. These contribution scores are converted into intuitive visualizations and human-readable explanations that clearly describe the factors influencing the agent's decision. The generated explanations, confidence scores, prediction logs, feature importance values, and training statistics are permanently recorded within the database to ensure transparency, traceability, and accountability. Interactive dashboards developed using HTML, CSS, JavaScript, and Chart.js provide users with real-time graphs, reward progression curves, feature importance charts, prediction summaries, and explanation reports, thereby improving user understanding and trust. Furthermore, the modular architecture allows future integration of advanced reinforcement learning algorithms, alternative explainability methods, cloud-based deployment, and additional benchmark environments. Consequently, the proposed system offers a scalable, interpretable, and trustworthy platform suitable for deployment in healthcare, autonomous vehicles, robotics, industrial automation, cybersecurity, financial decision support, and other safety-critical domains where transparent artificial intelligence is essential.

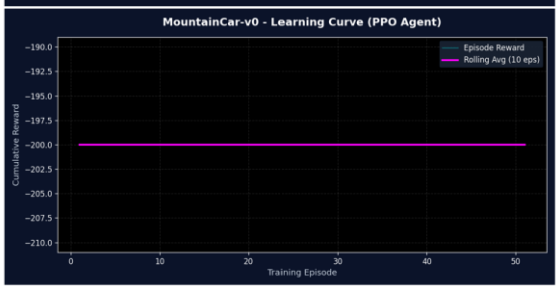
V. RESULTS & DISCUSSION

The proposed Explainable Reinforcement Learning framework was evaluated using multiple benchmark environments, including CartPole-v1, MountainCar-v0, and Acrobot-v1, to assess both learning performance and explainability. The Proximal Policy Optimization (PPO) algorithm successfully learned optimal policies by continuously interacting with each environment and maximizing cumulative rewards over successive training episodes. The reward curves demonstrated stable convergence, indicating that the agent effectively adapted to dynamic environmental conditions while maintaining consistent learning behaviour. Compared with conventional reinforcement learning approaches, the proposed framework exhibited improved training stability, faster policy convergence, and enhanced decision consistency due to the efficient clipping mechanism incorporated within PPO. Real-time monitoring through the interactive dashboard enabled users to visualize reward progression, episode completion, action selection frequency, and training statistics throughout the learning process. Performance metrics stored in the SQLite database facilitated comprehensive analysis of agent behaviour across different environments. Experimental observations confirmed that the trained agent achieved high cumulative rewards in CartPole-v1, significant improvement in MountainCar-v0, and stable policy optimization in Acrobot-v1, demonstrating the capability of the proposed system to solve diverse sequential decision-making problems efficiently. These results indicate that the framework effectively combines reinforcement learning performance with scalable web-based visualization and monitoring capabilities.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



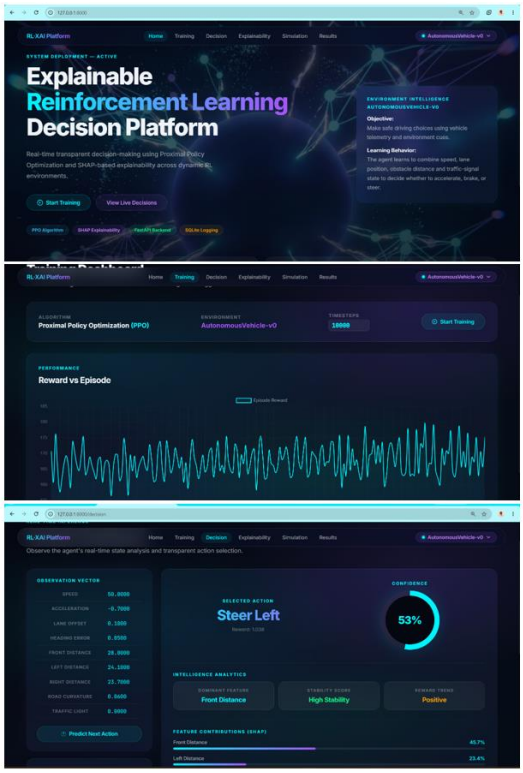


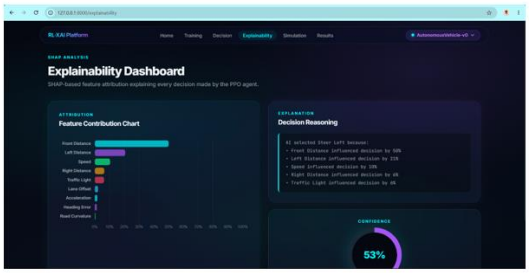


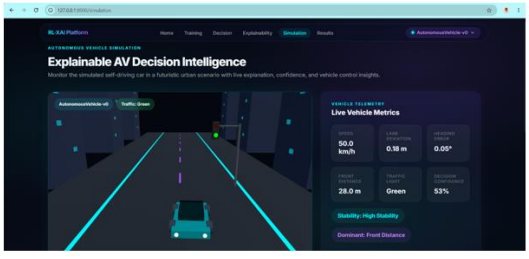
Beyond achieving competitive learning performance, the integration of SHAP-based Explainable Artificial Intelligence (XAI) significantly enhanced the transparency of autonomous decision-making. For every predicted action, SHAP computed feature contribution scores that identified the relative importance of each environmental state variable influencing the agent's decision. The generated explanations were presented through interactive graphical visualizations and human-readable descriptions, enabling users to understand the reasoning behind policy decisions without requiring knowledge of the underlying neural network. The audit logging mechanism recorded prediction history, confidence scores, feature importance values, and explanation reports, thereby improving traceability and accountability for autonomous operations. Compared with traditional black-box reinforcement learning models, the proposed framework substantially improved interpretability while preserving learning accuracy and computational efficiency. The interactive dashboard further simplified performance evaluation by presenting reward trends, feature contribution graphs, prediction summaries, and historical analysis in real time. Overall, the experimental results demonstrate that integrating PPO with SHAP successfully

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

transforms conventional reinforcement learning into a transparent, interpretable, and trustworthy intelligent decision-support framework. The proposed system is therefore well suited for deployment in healthcare, autonomous transportation, robotics, industrial automation, cybersecurity, and financial decision-support applications where both decision accuracy and explainability are essential for safe, reliable, and ethical Artificial Intelligence.







VI. CONCLUSION

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

The integration of Reinforcement Learning (RL) with Explainable Artificial Intelligence (XAI) represents a significant advancement toward developing transparent, trustworthy, and intelligent autonomous decision-making systems. This research presented an Explainable Reinforcement Learning framework that combines the Proximal Policy Optimization (PPO) algorithm with SHapley Additive exPlanations (SHAP) to overcome the limitations of conventional black-box reinforcement learning models. The proposed framework enables intelligent agents to learn optimal policies through continuous interaction with dynamic environments while simultaneously providing clear and human-understandable explanations for every decision. The implementation of a web-based platform using FastAPI, Stable-Baselines3, SHAP, SQLite, HTML, CSS, JavaScript, and Chart.js provides an interactive environment for agent training, real-time monitoring, performance evaluation, and explanation visualization. Experimental results obtained from benchmark environments, including CartPole-v1, MountainCar-v0, and Acrobot-v1, demonstrate that the PPO algorithm achieves stable policy convergence and effective learning performance across different dynamic scenarios. Furthermore, the integration of SHAP successfully identifies the contribution of individual environmental features influencing action selection, thereby improving transparency, accountability, and user confidence in autonomous systems. The interactive dashboard, comprehensive audit logging, and historical performance analysis further enhance system usability and facilitate informed decision-making by allowing users to understand agent behaviour throughout the learning process. Unlike traditional reinforcement learning approaches that primarily emphasize prediction accuracy, the proposed framework equally prioritizes interpretability, making it suitable for deployment in safety-critical domains where transparency is essential. The modular architecture also supports future integration of advanced reinforcement learning algorithms, causal explanation methods, multimodal learning, cloud-based deployment, and Large Language Model (LLM)-assisted explanation generation. Consequently, the proposed Explainable Reinforcement Learning platform establishes a scalable, reliable, and interpretable foundation for next-generation intelligent systems. It has significant potential for practical applications in healthcare, autonomous transportation, robotics, industrial automation, cybersecurity, finance, and other real-world domains that require accurate, explainable, and trustworthy Artificial Intelligence for responsible and ethical decision-making.

References

1. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
2. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature, 518*(7540), 529–533. https://doi.org/10.1038/nature14236
3. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347.*
4. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... Wierstra, D. (2016). Continuous control with deep reinforcement learning. *International Conference on Learning Representations (ICLR).*

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

5. Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *International Conference on Machine Learning*, 1861–1870.

6. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature, 550*(7676), 354–359.

7. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.

8. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

9. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608.*

10. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160.

11. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2019). A survey of methods for explaining black-box models. *ACM Computing Surveys, 51*(5), 93.

12. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115.

13. Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K. R. (Eds.). (2021). *Explainable AI: Interpreting, explaining and visualizing deep learning*. Springer.

14. Milani, S., Topin, N., Veloso, M., & Fang, F. (2024). Explainable reinforcement learning: A survey and comparative review. *ACM Computing Surveys.*

15. Bekkemoen, Y., Zhang, Y., & Colleagues. (2024). Explainable reinforcement learning: A systematic literature review. *Artificial Intelligence Review.*

16. Glanois, C., Weng, P., Zimmer, M., Li, D., Yang, T., Hao, J., & Liu, W. (2024). A survey on interpretable reinforcement learning. *Journal of Artificial Intelligence Research.*

17. Cheng, E., Yu, J., & Xing, X. (2025). A survey on explainable deep reinforcement learning. *IEEE Transactions on Artificial Intelligence.*

18. Saulières, L. (2025). Survey of explainable reinforcement learning: Targets, methods and needs. *Artificial Intelligence Review.*

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

19. Dazeley, R., Vamplew, P., Cruz, F., & Foale, C. (2024). Explainable reinforcement learning for broad XAI. *Knowledge-Based Systems.*

20. Sutton, R. S. (2019). The bitter lesson. *Incomplete Ideas Blog.*

21. OpenAI. (2023). GPT-4 technical report. *arXiv preprint arXiv:2303.08774.*

22. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30.*

23. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

24. Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

25. Bellemare, M. G., Dabney, W., & Munos, R. (2023). *Distributional reinforcement learning*. MIT Press.

26. Agarwal, R., Schwarzer, M., Castro, P. S., Courville, A., & Bellemare, M. G. (2021). Deep reinforcement learning at the edge of the statistical precipice. *Advances in Neural Information Processing Systems.*

27. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565.*

28. Molnar, C. (2022). *Interpretable machine learning* (2nd ed.). Lulu Press.

29. Sutton, R. S., Mahmood, A. R., & White, M. (2020). An emphatic approach to the problem of off-policy temporal-difference learning. *Journal of Machine Learning Research, 17*(73), 1–29.

30. Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research, 32*(11), 1238–1274.