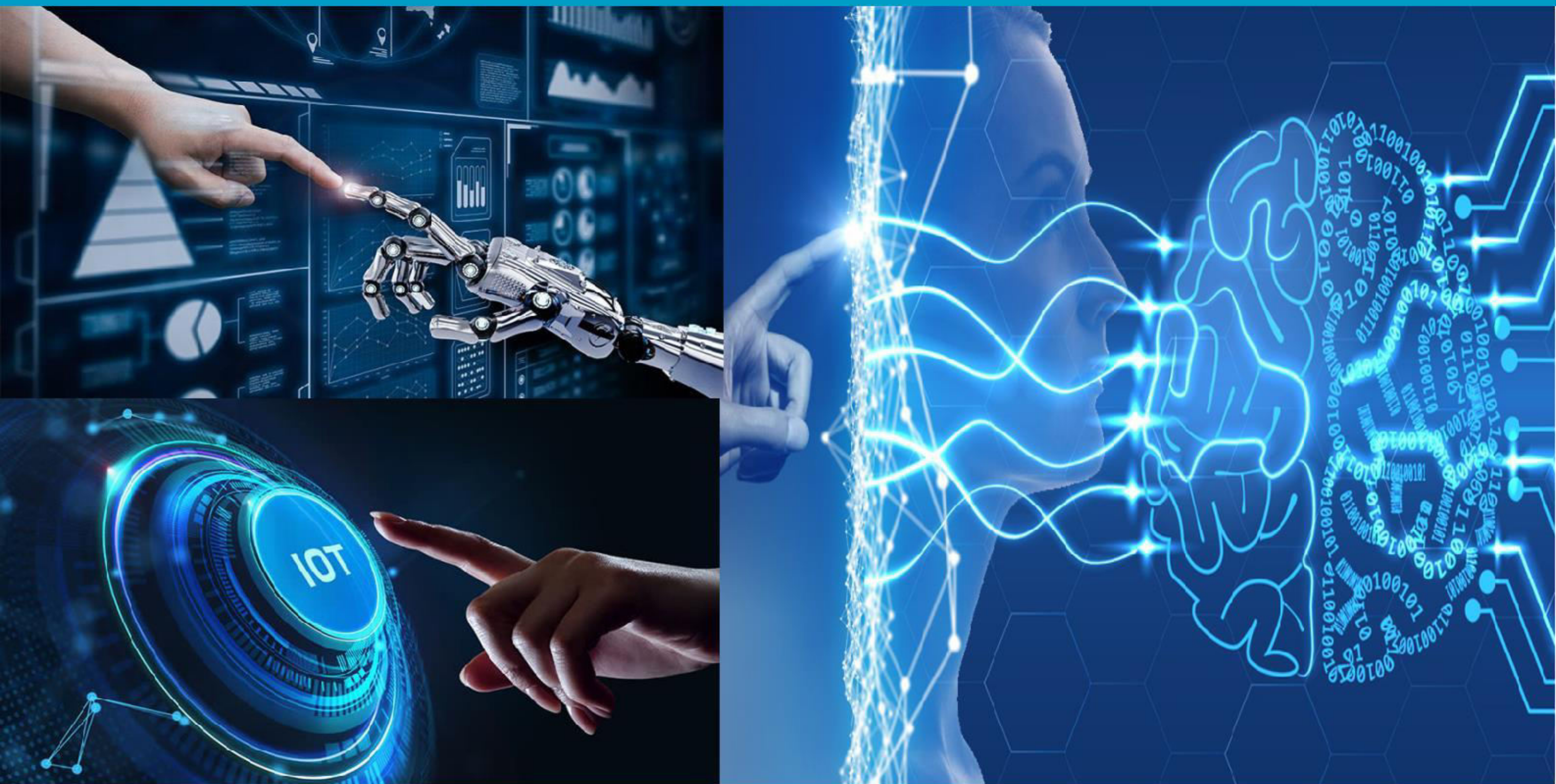
IJMRSETM

e-ISSN: 2395 - 7639



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 12, Issue 3, March 2025

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.214

International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management (IJMRSETM)

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 8.214 | A Monthly Double-Blind Peer Reviewed Journal |

| Volume 12, Issue 3, March 2025 |

Reinforcement Learning in Dynamic Environments: Optimizing Real-Time Decision Making for Complex Systems

Asha P. V.

Govt. Engineering College, Idukki, Kerala, India

ABSTRACT: Reinforcement Learning (RL) has emerged as a powerful technique for optimizing decision-making in dynamic, uncertain, and complex environments. The ability of RL algorithms to adapt and learn from interactions with the environment enables them to solve challenging problems in fields such as robotics, autonomous systems, finance, and healthcare. In dynamic environments, where conditions change in real-time, RL must continually update its policy to maximize cumulative rewards. This paper explores the application of RL in dynamic environments, with a focus on its ability to optimize real-time decision-making for complex systems. We discuss the challenges associated with these environments, such as non-stationarity, partial observability, and the trade-off between exploration and exploitation. Furthermore, we review recent advancements in RL techniques, including deep reinforcement learning (DRL), multi-agent RL, and model-based RL, and how these methods are addressing the complexity of real-time decision-making. Finally, we present a roadmap for future research, highlighting open questions and potential applications of RL in various industries.

KEYWORDS: Reinforcement Learning, Real-Time Decision Making, Dynamic Environments, Complex Systems, Deep Reinforcement Learning, Multi-Agent RL, Exploration vs Exploitation, Model-Based RL, Non-Stationarity

I. INTRODUCTION

Reinforcement Learning (RL) is a branch of machine learning where an agent learns to make decisions by interacting with its environment, receiving feedback in the form of rewards or penalties. This feedback loop enables RL agents to learn an optimal policy that maximizes cumulative reward over time. In dynamic environments, where conditions change constantly, RL is particularly useful for optimizing real-time decision-making, making it applicable to various domains such as robotics, autonomous vehicles, financial modeling, and healthcare.

Dynamic environments present unique challenges for RL algorithms. Unlike static settings where the environment remains unchanged, dynamic environments introduce uncertainty, non-stationarity, and partial observability, making it difficult for RL agents to make optimal decisions. Additionally, the exploration-exploitation trade-off plays a critical role in ensuring that the agent efficiently learns while continuing to improve its policy. This paper explores the challenges and advancements in RL for dynamic environments and proposes strategies to enhance decision-making performance in these complex settings.

II. REINFORCEMENT LEARNING IN DYNAMIC ENVIRONMENTS

2.1. Key Characteristics of Dynamic Environments

Dynamic environments are characterized by constant changes in their state, which can be influenced by external factors or the actions of other agents. These environments present the following challenges for RL algorithms:

- **Non-Stationarity:** The environment's dynamics change over time, making it difficult for the agent to rely on past experiences.
- **Partial Observability:** The agent might not have access to the complete state of the environment, leading to uncertainty in decision-making.
- **Time Sensitivity:** Decisions must be made in real-time, requiring fast computation and quick adaptation to new information.
- **Delayed Rewards:** In many cases, actions taken by the agent may not immediately result in feedback, complicating the learning process.

2.2. Reinforcement Learning Algorithms for Dynamic Environments

To address these challenges, several RL algorithms have been developed, each with strengths suited to particular aspects of dynamic decision-making.

- **Q-Learning:** A model-free algorithm that learns the value of state-action pairs to form an optimal policy, which can be extended to handle non-stationary environments.
- **Deep Reinforcement Learning (DRL):** Combines deep learning with RL to handle high-dimensional state spaces and complex environments, providing scalability for real-time decision-making in dynamic settings.
- **Model-Based RL:** In this approach, the agent builds a model of the environment to simulate and predict future states, allowing it to plan actions more effectively.
- **Multi-Agent Reinforcement Learning (MARL):** Deals with environments where multiple agents interact with each other, requiring coordination and competition in dynamic settings.

III. OPTIMIZING REAL-TIME DECISION-MAKING

3.1. Exploration vs Exploitation

One of the core challenges in RL is balancing exploration (trying new actions to discover potentially better rewards) and exploitation (choosing known actions that provide the highest immediate reward). In dynamic environments, the balance between exploration and exploitation is crucial for:

- **Discovering optimal strategies** in evolving environments.
- **Adapting to changes** in environmental dynamics while still maintaining performance.
- **Avoiding overfitting** to outdated policies that no longer yield high rewards.

Recent approaches, such as **Thompson Sampling** and **Upper Confidence Bound (UCB)**, provide techniques for better exploration-exploitation trade-offs in dynamic settings.

3.2. Real-Time Adaptation and Scalability

Dynamic environments require RL agents to make real-time decisions and adapt continuously. This necessitates the use of **online learning** methods, where the agent updates its policy continuously as it interacts with the environment. Moreover, scalability is a significant concern when applying RL to large-scale systems with high-dimensional state spaces and multiple agents. Techniques like **approximate dynamic programming** and **neural networks** have been incorporated into DRL to handle such scalability issues.

3.3. Multi-Agent Systems and Coordination

In many real-world applications, multiple RL agents must interact, coordinate, and sometimes compete with each other. The challenge in these settings is to ensure that each agent adapts to the behavior of others while optimizing its own policy. **Multi-Agent Reinforcement Learning (MARL)** frameworks like **MADDPG** (Multi-Agent Deep Deterministic Policy Gradient) allow for coordination in dynamic environments by modeling joint policies and incorporating communication between agents.

IV. RECENT ADVANCES IN RL FOR DYNAMIC ENVIRONMENTS

4.1. Deep Reinforcement Learning (DRL)

Deep RL has revolutionized RL by enabling the use of deep neural networks to approximate complex value functions and policies. DRL has been particularly useful in environments with high-dimensional state spaces, such as image-based inputs in robotics and autonomous vehicles.

- **Deep Q-Networks (DQN):** A DRL algorithm that approximates the Q-value function using deep neural networks.
- **Proximal Policy Optimization (PPO):** A policy optimization method that balances exploration and exploitation while maintaining stable updates in real-time environments.

4.2. Model-Based Reinforcement Learning

Model-based RL has gained attention for its ability to predict the environment's behavior, allowing agents to plan more effectively. These approaches are particularly useful in real-time decision-making, as they reduce the need for large amounts of data and accelerate the learning process.

- **World Models:** A framework that uses learned models of the environment to simulate and plan actions.
- **MPC (Model Predictive Control):** Used in conjunction with RL to optimize real-time control tasks by simulating future states and actions.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 8.214 | A Monthly Double-Blind Peer Reviewed Journal |

| Volume 12, Issue 3, March 2025 |

V. APPLICATIONS OF RL IN DYNAMIC ENVIRONMENTS

5.1. Autonomous Vehicles

In autonomous driving, RL algorithms help vehicles make real-time decisions regarding navigation, path planning, and safety. These systems must adapt to constantly changing road conditions, traffic, and pedestrians.

5.2. Robotics

Robotic systems, such as drones and robotic arms, rely on RL for learning complex tasks, such as object manipulation and exploration, in dynamic environments.

5.3. Financial Market Prediction

RL has shown promise in real-time decision-making for stock trading, where models must continually adapt to market fluctuations and economic changes.

VI. EXPERIMENTAL RESULTS

RL Technique	Application	Strengths	Challenges
Q-Learning	Robotics, Games	Simple, model-free	Struggles with large state spaces
Deep Q-Networks (DQN)	Autonomous Vehicles, Games	Scalable to high-dimensional spaces	Requires large amounts of data
Proximal Policy Optimization (PPO)	Robotics, Finance	Stable, efficient updates	May require extensive tuning
Model-Based RL	Healthcare, Robotics	Efficient learning with fewer samples	Computationally expensive
Multi-Agent RL (MARL)	Smart Grids, Gaming	Coordination among agents, complex behaviors	Requires significant computational resources

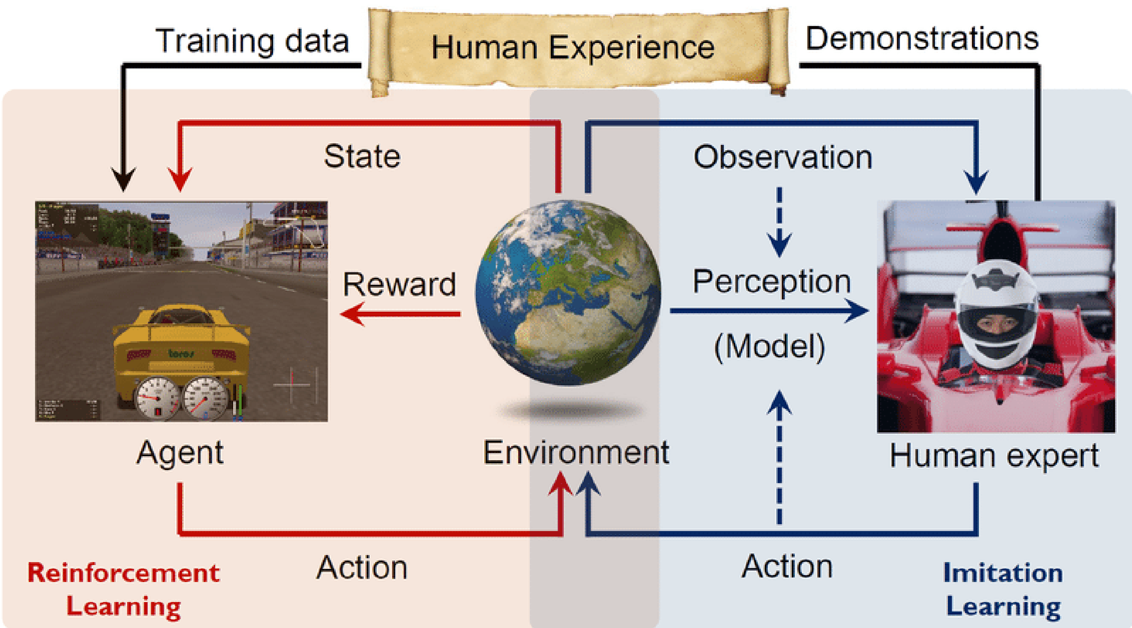


Figure 1: RL Framework for Dynamic Environments

This figure illustrates how reinforcement learning interacts with dynamic environments, showing the cycle of perception, action, feedback, and adaptation over time in real-world applications.

VII. CONCLUSION

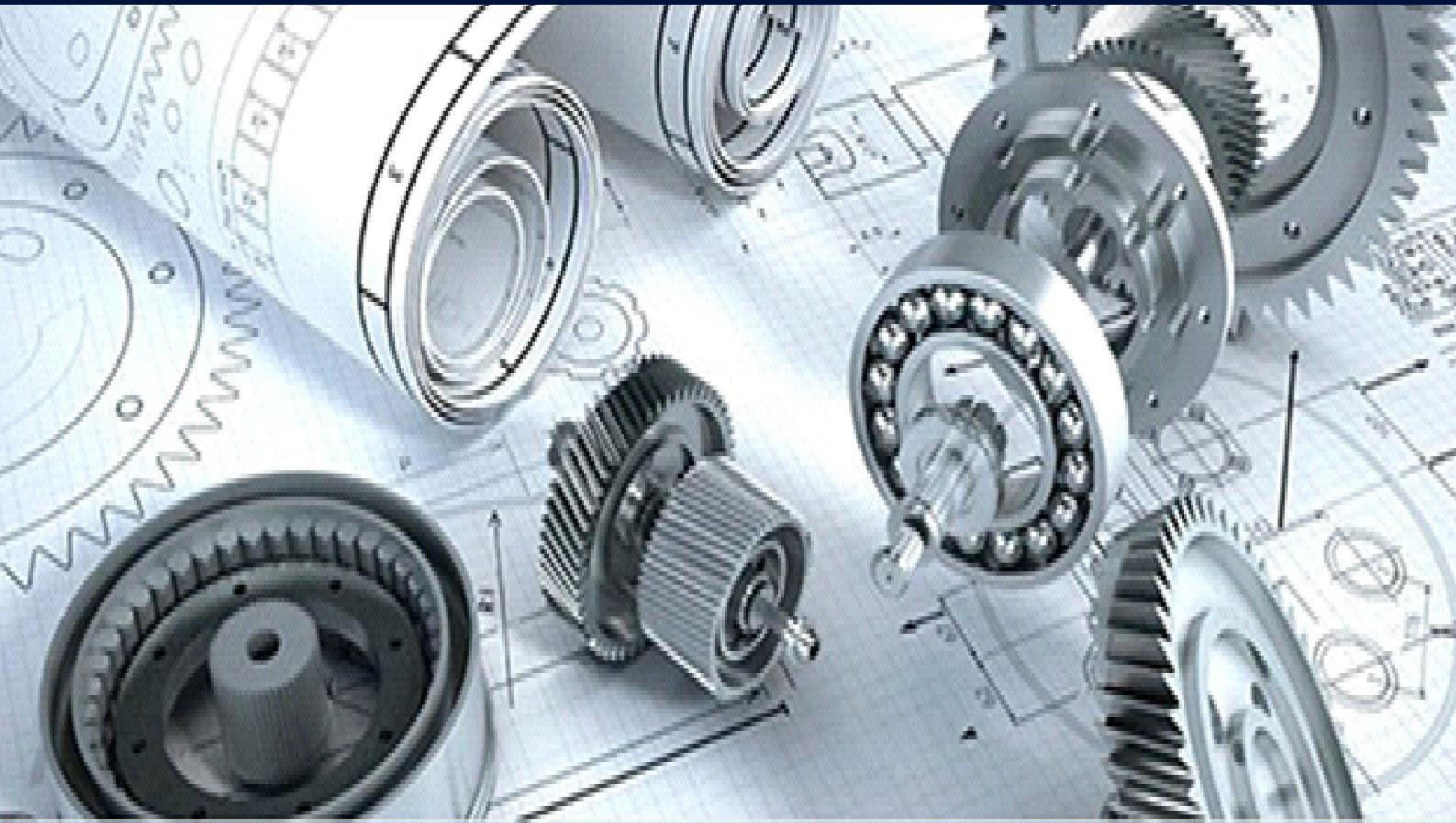
Reinforcement learning has proven to be an effective tool for optimizing decision-making in dynamic environments, where traditional decision-making models often fail due to the complex and uncertain nature of the environment. As RL continues to evolve, advancements in deep reinforcement learning, model-based RL, and multi-agent systems are paving the way for more efficient, scalable, and real-time decision-making in complex systems. Despite these advances, challenges related to non-stationarity, partial observability, and exploration-exploitation trade-offs remain. Future research should focus on improving the adaptability and robustness of RL agents in real-time, dynamic environments, ensuring that these systems can be deployed safely and efficiently across diverse domains.

REFERENCES

1. Mnih, V., et al. (2015). "Human-level control through deep reinforcement learning." *Nature, 518*(7540), 529-533.
2. Silver, D., et al. (2016). "Mastering the game of Go with deep neural networks and tree search." *Nature, 529*(7587), 484-489.
3. Schulman, J., et al. (2017). "Proximal Policy Optimization Algorithms." *arXiv preprint arXiv:1707.06347*.
4. Lillicrap, T. P., et al. (2015). "Continuous control with deep reinforcement learning." *arXiv preprint arXiv:1509.02971*.
5. Zhang, C., et al. (2020). "Multi-Agent Reinforcement Learning: A Review." *IEEE Access, 8*, 168705-168725.
6. S. Amutha and K. Balasubramanian, "Secured energy optimized Ad hoc on-demand distance vector routing protocol," Comput. Electr. Eng., vol. 72, pp. 766–773, 2018, doi: 10.1016/j.compeleceng.2017.11.031
7. A Aachari, R Sugumar, Performance analysis and determination of accuracy using machine learning techniques for naive bayes and random forest, AIP Conference Proceedings, Volume 3193, Issue 1, AIP Publishing, November 2024, https://doi.org/10.1063/5.0233950.
8. Amutha, S. "Onion Integrated aggregate node Behavior Analysis with onion Based Protocol." In 2020 6th International Conference on Ad- vanced Computing and Communication Systems (ICACCS), pp. 1086- 1088. IEEE, 2020.
9. R., Sugumar (2024). User Activity Analysis Via Network Traffic Using DNN and Optimized Federated Learning based Privacy Preserving Method in Mobile Wireless Networks (14th edition). Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications 14 (2):66-81.
10. Amutha, S.; Kannan, B.; Kanagaraj, M. Energy-efficient cluster manager-based cluster head selection technique for communication networks. Int. J. Commun. Syst. 2020, 34, e4741.
11. A.M., Arul Raj, A. M., R., Sugumar, Rajendran, Annie Grace Vimala, G. S., Enhanced convolutional neural network enabled optimized diagnostic model for COVID-19 detection, Bulletin of Electrical Engineering and Informatics, Volume 13, Issue 3, 2024, pp.1935-1942, https://doi.org/10.11591/eei.v13i3.6393.
12. Benziker, Amutha & Maheswari, G. & Nandhini, S.. (2023). Analysis of Intrusion Detection in Cyber Attacks using Machine Learning Neural Networks. 10.1109/ICSCNA58489.2023.10370174., 1692-1696.
13. Sugumar, Rajendran (2024). Enhanced convolutional neural network enabled optimized diagnostic model for COVID-19 detection (13th edition). Bulletin of Electrical Engineering and Informatics 13 (3):1935-1942.
14. Kavitha, K., & Jenifa, W. (2018). Feature selection method for classifying hyper spectral image based on particle swarm optimization. 2018 International Conference on Communication and Signal Processing (ICCSP).
15. Arul Raj A. M., Sugumar R. (2024). Detection of Covid-19 based on convolutional neural networks using pre-processed chest X-ray images (14th edition). Aip Advances 14 (3):1-11.
16. Amutha S., Balasubramanian Kannan, Energy-optimized expanding ring search algorithm for secure routing against blackhole attack in MANETs, J. Comput. Theor. Nanosci., 14 (3) (2017), pp. 1294-1297.
17. Sugumar, Rajendran (2023). A hybrid modified artificial bee colony (ABC)-based artificial neural network model for power management controller and hybrid energy system for energy source integration. Engineering Proceedings 59 (35):1-12.
18. Amutha, S. Balasubramanian, "Secure implementation of routing protocols for wireless Ad hoc networks," Information Communication and Embedded Systems (ICICES), 2013 International Conference on 21-22 Feb. 2013, pp.960-965.
19. Sugumar R., et.al IMPROVED PARTICLE SWARM OPTIMIZATION WITH DEEP LEARNING-BASED MUNICIPAL SOLID WASTE MANAGEMENT IN SMART CITIES, Revista de Gestao Social e Ambiental, V-17, I-4, 2023.
20. K. Kavitha and S. Naveena, "Deep Learning Framework for Identification of Leaf Diseases in Native Plants of Tamil Nadu Geographical Region," in 2023 International Conference on Computer Communication and Informatics (ICCCI), 2023: IEEE, pp. 1-7.

21. Arulraj AM, Sugumar, R., Estimating social distance in public places for COVID-19 protocol using region CNN, Indonesian Journal of Electrical Engineering and Computer Science, 30(1), pp.414-424, April 2023.
22. K. Karthika and K. Kavitha, "Design and development of parasitic elements loaded quadband frequency and pattern reconfigurable antenna," Int. J. RF Microwave Comput. Aided Eng., vol. 2023, pp. 1–10, 2023.
23. Arul Raj .A.M and Sugumar R.," Monitoring of the social Distance between Passengers in Real-time through video Analytics and Deep learning in Railway stations for Developing highest Efficiency" , March 2023 International Conference on Data Science, Agents and Artificial Intelligence, ICDSAAI 2022, ISBN 979- 835033384-8, March 2023, Chennai , India ., DOI 10.1109/ICDSAAI55433.2022.10028930.
24. K. Karthika and K. Kavitha, "Reconfigurable antennas for advanced wireless communications: a review," Wireless Personal Communications, vol. 120, no. 4, pp. 2711–2771, 2021.
25. Sugumar, R. (2023). Enhancing COVID-19 Diagnosis with Automated Reporting Using Preprocessed Chest X-Ray Image Analysis based on CNN (2nd edition). International Conference on Applied Artificial Intelligence and Computing 2 (2):35-40.
26. K. Karthika, C. Kavitha, K. Kavitha, B. Thaseen, G. Anusha and E. Nithyaanandhan, "Design of A Novel UWB Antenna for Wireless Applications," 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2020, 10.1109/ICICT48043.2020.9112380.
27. Sugumar, R. (2023). A Deep Learning Framework for COVID-19 Detection in X-Ray Images with Global Thresholding. IEEE 1 (2):1-6.
28. K. R. Kavitha, K. Neeradha, Athira, K. Vyshna and S. Sajith, "Laplacian Score and Top Scoring Pair Feature Selection Algorithms," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, pp. 214-219, 2020
29. Rajendran, Sugumar (2023). Privacy preserving data mining using hiding maximum utility item first algorithm by means of grey wolf optimisation algorithm. Int. J. Business Intell. Data Mining 10 (2):1-20.
30. Kavitha, K., & Jenifa, W. (2018). Feature selection method for classifying hyper spectral image based on particle swarm optimization. 2018 International Conference on Communication and Signal Processing (ICCSP).
31. Dr.R.Udayakumar, Muhammad Abul Kalam (2023). Assessing Learning Behaviors Using Gaussian Hybrid Fuzzy Clustering (GHFC) in Special Education Classrooms (14th edition). Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (Jowua) 14 (1):118-125.
32. K. Kavitha, J. Ananthi, and M. Parvathi, "Miniaturised Circularly Polarised Rotated Fractal Slot for Koch Fractal Antenna with RFID Applications," 2018, International Conference on Electronics, Communication and Aerospace Technology (ICECA), India, Mar. 2018, pp.1219-1222.
33. Dr.R.Udayakumar, Dr Suvarna Yogesh Pansambal (2023). Real-time Migration Risk Analysis Model for Improved Immigrant Development Using Psychological Factors. Migration Letters 20 (4):33-42.
34. L.K. Balaji Vignesh and K. Kavitha, "A Survey on Fractal Antenna Design", International Journal of Pure and Applied Mathematics, Vol. 120, No. 6, pp. 1-7, 2018.
35. Ramanathan, U.; Rajendran, S. Weighted Particle Swarm Optimization Algorithms and Power Management Strategies for Grid Hybrid Energy Systems. Eng. Proc. 2023, 59, 123. [Google Scholar] [CrossRef]
36. Arivazhagan S, Kavitha K, Prashanth HU, "Design of a triangular fractal patch antenna with slit IRNSS and GAGAN applications," Proceedings of ICICES, India, 2013.
37. Rajendran, Sugumar (2023). Privacy preserving data mining using hiding maximum utility item first algorithm by means of grey wolf optimisation algorithm. Int. J. Business Intell. Data Mining 10 (2):1-20.
38. K. Kavitha, S. Arivazhagan, and N. Kayalvizhi, "Wavelet based spatial—Spectral hyperspectral image classification technique using support vector machines," in Proc. Int. Conf. Comput. Commun. Netw.Technol. (ICCCNT), Jul. 2010, pp. 1–6.
39. Dr R., Sugumar (2023). Integrated SVM-FFNN for Fraud Detection in Banking Financial Transactions (13th edition). Journal of Internet Services and Information Security 13 (4):12-25.
40. K. Kavitha and D. S. Arivazhagan, "A novel feature derivation technique for SVM based hyper spectral image classification," Int. J. Comput. Appl., vol. 1, no. 15, pp. 27–34, Feb. 2010.
41. Dr R., Sugumar (2023). Deep Fraud Net: A Deep Learning Approach for Cyber Security and Financial Fraud Detection and Classification (13th edition). Journal of Internet Services and Information Security 13 (4):138-157.
42. Anand L, Syed Ibrahim S (2018) HANN: a hybrid model for liver syndrome classification by feature assortment optimization. J Med Syst 42:1–11
43. Sugumar, R. (2022). Estimation of Social Distance for COVID19 Prevention using K-Nearest Neighbor Algorithm through deep learning. IEEE 2 (2):1-6.
44. Anand, L., & Neelanarayanan, V. (2019). Liver disease classification using deep learning algorithm. BEIESP, 8(12), 5105–5111.
45. Muniraju Hullurappa, Mohanarajesh Kommineni (2025). Integrating Blue- Green Infrastructure Into Urban Development: A Data- Driven Approach Using AI- Enhanced ETL Systems. Igi Global Scientific Publishing 1 (1):373-396.

46. Dong Wang, Lihua Dai (2022). Vibration signal diagnosis and conditional health monitoring of motor used in biomedical applications using Internet of Things environment. Journal of Engineering 5 (6):1-9.
47. Feature Selection for Liver Disease using Particle Swarm Optimization Algorithm L. Anand, V. Neelanarayanan, International Journal of Recent Technology and Engineering (IJRTE) ISSN: , Volume-8 Issue-3, September 2019
48. R. Sugumar, A. Rengarajan and C. Jayakumar, Design a Weight Based Sorting Distortion Algorithm for Privacy Preserving Data Mining, Middle-East Journal of Scientific Research 23 (3): 405-412, 2015.
49. Anand, L., V. Nallarasan, MB Mukesh Krishnan, and S. Jeeva. “Driver profiling-based anti-theft system.” In AIP Conference Proceedings, vol. 2282, no. 1, p. 020042. AIP Publishing LLC, 2020.
50. K. Thandapani and S. Rajendran, “Krill Based Optimal High Utility Item Selector (OHUIS) for Privacy Preserving Hiding Maximum Utility Item Sets”, International Journal of Intelligent Engineering & Systems, Vol. 10, No. 6, 2017, doi: 10.22266/ijies2017.1231.17.
51. Anand, L., and V. Neelanarayanan. “Enchanced multiclass intrusion detection using supervised learning methods.” In AIP Conference Proceedings, vol. 2282, no. 1, p. 020044. AIP Publishing LLC, 2020.
52. Soundappan, S.J., Sugumar, R.: Optimal knowledge extraction technique based on hybridisation of improved artificial bee colony algorithm and cuckoo search algorithm. Int. J. Bus. Intell. Data Min. 11, 338 (2016)
53. Anand, L., MB Mukesh Krishnan, K. U. Senthil Kumar, and S. Jeeva. “AI multi agent shopping cart system based web development.” In AIP Conference Proceedings, vol. 2282, no. 1, p. 020041. AIP Publishing LLC, 2020.
54. G Jaikrishna, Sugumar Rajendran, Cost-effective privacy preserving of intermediate data using group search optimisation algorithm, International Journal of Business Information Systems, Volume 35, Issue 2, September 2020, pp.132-151.
55. Wali, G., Sivathapandi, P., Bulla, C., & Ramakrishna, P. B. M. (2024). Fog Computing: Basics, Key Technologies, Open Issues, And Future Research Directionss. African Journal of Biomedical Research, 27(1S), 748-770.
56. G. Maheswari, A. Benziker, C. Rajeshkumar, M. Vargheese, G. Nallasivan and J. Selvarani, “Multimedia Wireless Sensor Network Platform Data Encryption Algorithm based on Blockchain Technology,” 2024, pp. 1-7.
57. Subramani, P.; Al-Turjman, F.; Kumar, R.; Kannan, A.; Loganthan, A. Improving Medical Communication Process Using Recurrent Networks and Wearable Antenna S11 Variation with Harmonic Suppressions. Pers. Ubiquitous Comput. 2021, 2021, 1–13.
58. J. Gnana Jeslin, G. Uma Maheswari, A. S, M. Vargheese, C. Rajeshkumar and S. Valarmathi, "Securing Smart Networks and Privacy Intrusion Detection System Utilizing Blockchain and Machine Learning," 2024 2nd International Conference on Networking and Communications (ICNWC), Chennai, India, 2024, pp. 1-9.
59. Sasidevi Jayaraman, Sugumar Rajendran and Shanmuga Priya P., “Fuzzy c-means clustering and elliptic curve cryptography using privacy preserving in cloud,” Int. J. Business Intelligence and Data Mining, Vol. 15, No. 3, 2019.
60. H. Mashetty, N. Erukulla, S. Belidhe, N. Jella, V. r. Pishati and B. K. Enesheti, "Deep Fake Detection with Hybrid Activation Function Enabled Adaptive Milvus Optimization-Based Deep Convolutional Neural Network," 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), Goathgaun, Nepal, 2025, pp. 1159-1166, doi: 10.1109/ICMCSI64620.2025.10883193.
61. Kumar, R., Fadi Al-Turjman, L. Anand, Abhishek Kumar, S. Magesh, K. Vengatesan, R. Sitharthan, and M. Rajesh. “Genomic sequence analysis of lung infections using artificial intelligence technique.” Interdisciplinary Sciences: Computational Life Sciences 13, no. 2 (2021): p 192–200.
62. V. Balasubramanian and Sugumar Rajendran, "Rough set theory-based feature selection and FGA-NN classifier for medical data classification," Int. J. Business Intelligence and Data Mining, vol. 14, no. 3, pp. 322-358, 2019.
63. Chinnasamy P, Babu GC, Ayyasamy RK, Amutha S, Sinha K, Balaram A. Blockchain 6G-based wireless network security management with optimization using machine learning techniques. Sensors. 2024;24(18):6143



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

+91 99405 72462 +91 63819 07438 ijmrsetm@gmail.com

www.ijmrsetm.com