Reinforcement Learning for Adaptive Cloud Resource Allocation

Gaurav Malave (2022BCD0017)
Hemanth Sai Manikanta (2022BCD0008)
Eedara Sri Rahul (2022BCD0004)
Pradipth Gautam S (2022BCS0235)

Guided By,
Dr. Divya Sindhu Lekha

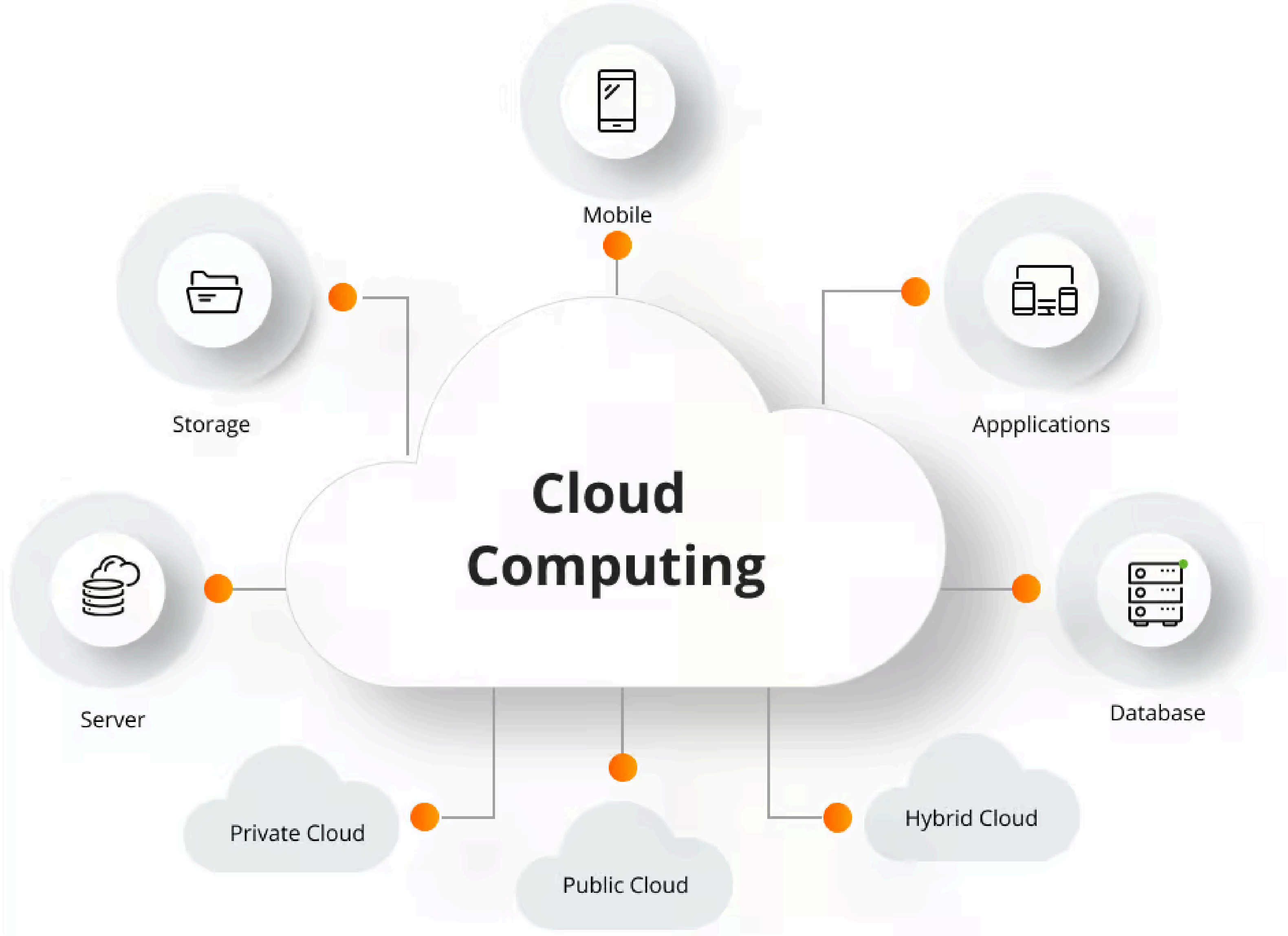
Introduction

Brief Introduction to Topic

- **Cloud Evolution:** Modern clouds face diverse, dynamic workloads.
- **Resource Allocation Challenge:** Static methods fail in real-time, heterogeneous environments.
- **RL Solution:** Agents learn optimal allocation via interaction.

Significance of the Area

- **Market:** $1294.9B in 2025, growing 15.7% CAGR.
- **Efficiency:** 30–40% resource waste due to poor allocation.
- **Performance:** Direct impact on SLAs & user experience.
- **Scalability:** IoT surge (75B devices by 2025).
- **Sustainability:** Data centers = 1.5% global electricity; optimization cuts impact.



Source: Markets & Markets, “Cloud Computing Market,” 2025

Literature Review

Paper Title	Year	Authors	Model Used	Key Novelty	Models compared	Performance Gain
Edge Computing and Cloud Collaborative Resource Scheduling Optimization	2025	Wang & Yang	Double DQN with Dueling Architecture	Joint edge-cloud optimization with migration cost modeling	DQN,PPO,A3C, DDPG	18% processing time reduction, 12% resource utilization improvement
Optimizing task offloading and resource allocation in edge-cloud networks	2023	Ullah et al.	DDQN-EC	Batch processing approach with MDP formulation	DQN, DDPG, heuristic approaches	Superior task acceptance ratio, reduced rejection
Deep reinforcement learning for collaborative edge computing	2020	Li et al.	DQN	Vehicular edge computing focus	DQN vs heuristic baseline	Improved energy efficiency and latency

Literature Review

Paper Title	Year	Authors	Model Used	Key Novelty	Models compared	Performance Gain
Deep reinforcement learning based resource allocation strategy	2022	Xu et al.	DRL-based	Cloud-edge computing integration	Compared with heuristic algorithms (greedy, random, round-robin)	Enhanced resource utilization
Resource allocation based on deep reinforcement learning in IoT	2020	Xiong et al.	DDQN	IoT edge computing specialization	Compared with DQN, heuristic baselines	Optimized computation offloading
Deep reinforcement learning for online computation offloading	2019	Huang et al.	DROO Algorithm	Wireless-powered MEC networks	Compared with greedy offloading, local-only, random allocation	Energy efficiency improvement

Research Evolution Trends

Indian Institute of Information Technology Kottayam

2019

Focus on basic DQN implementations for edge computing

Huang, L., Bi, S., & Zhang, Y.-J.A. (2019).

2020

Optimized computation offloading

Li, M., Gao, J., Zhao, L., & Shen, X. (2020).

2022

Introduction of DDQN and advanced architectures

Xu, J., Xu, Z., & Shi, B. (2022).

2023

Integration of edge-cloud collaborative approaches

Ullah, I., Lim, H.-K., Seok, Y.-J., & Han, Y.-H. (2023)

2025

Enhanced resource utilization

Wang, Y., & Yang, X. (2025)

PRESENT

Enhanced multi-objective optimization

Motivation

Research Gaps

- **Learning efficiency** – models need many iterations.
- **Training overhead** – high cost, slow deployment.
- **Convergence stability** – unstable under dynamic workloads.
- **Fault tolerance** – failures rarely considered.

Why It Matters

- Poor allocation → higher costs & degraded performance.
- Need for smarter, adaptive methods to handle IoT growth & dynamic workloads.
- Energy savings = cost savings + sustainability (green computing).

Our Approach

- Implement a Reinforcement Learning–based allocator for a testbed.
- Use a dynamic reward function balancing performance, cost, and energy.
- Evaluate with realistic workload traces (e.g., from Google cluster dataset).
- Focus on practical feasibility, not just theoretical improvement.

Motivation

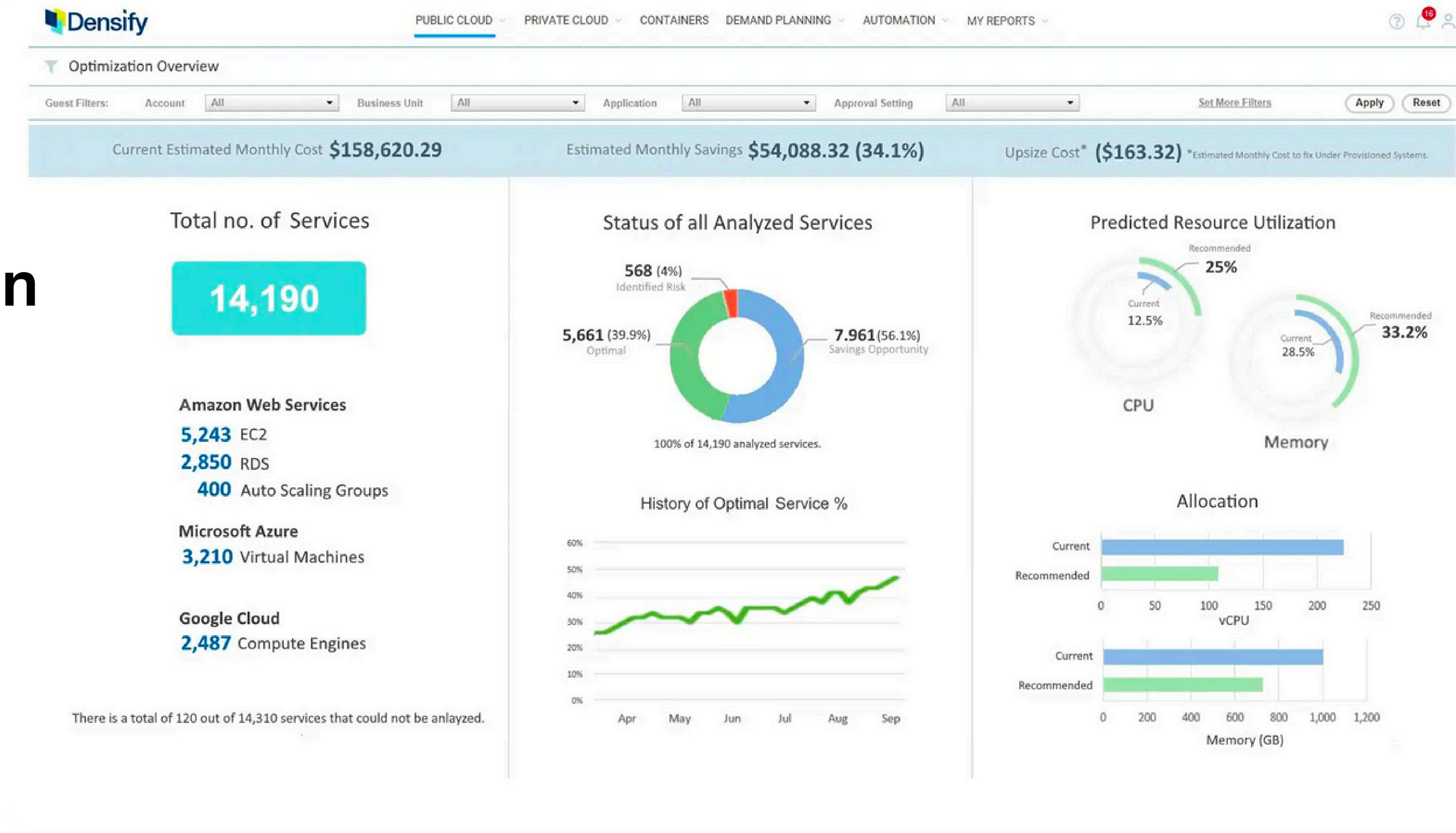
Metrics for Evaluation

Objective 1: Improve Performance
- Task completion time (avg, p95)
- Throughput (tasks/sec)

Objective 2: Increase Resource Utilization
- Resource utilization (%)
- Task acceptance ratio vs. rejection ratio

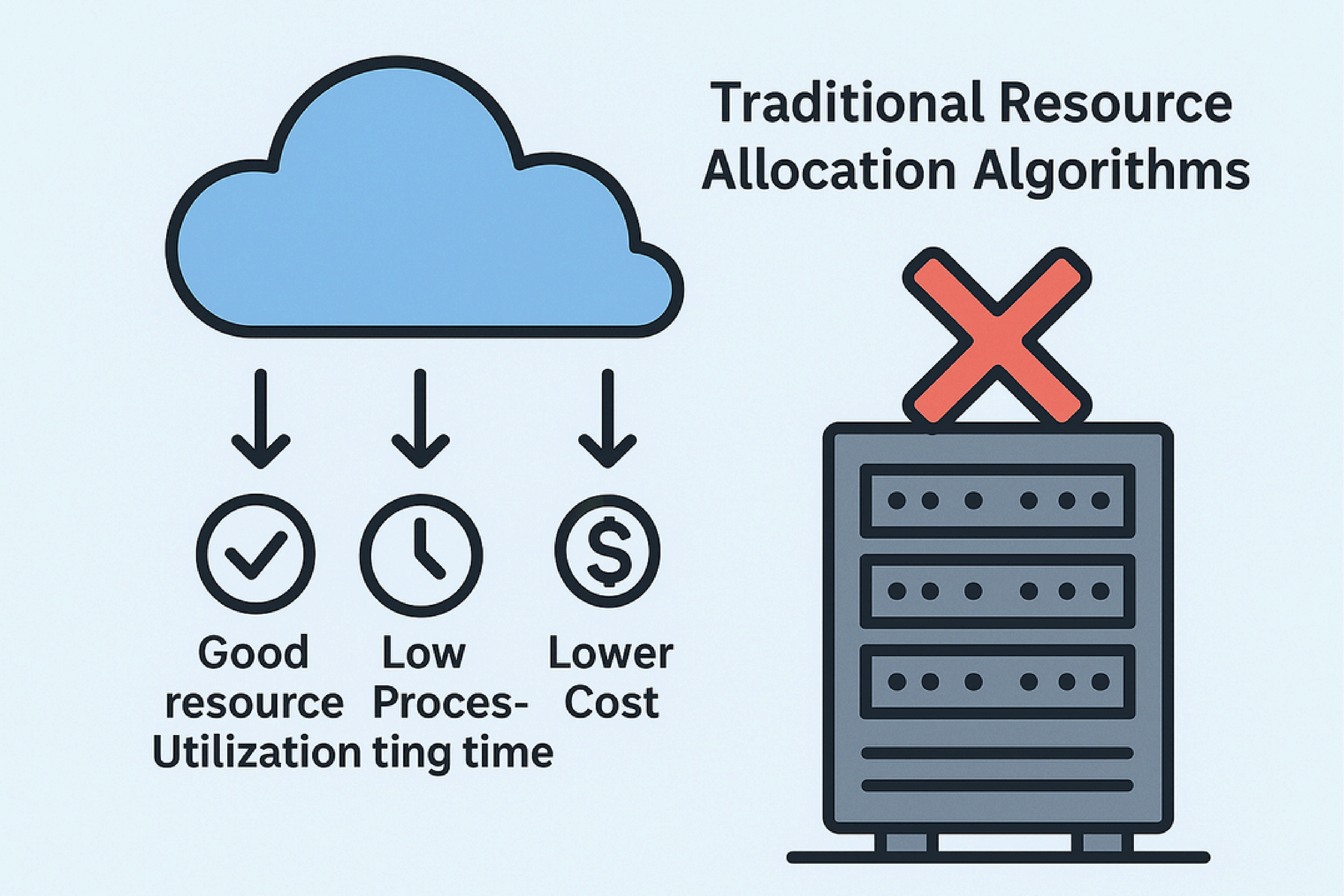
Objective 3: Ensure Practical Feasibility
- Convergence speed (iterations to stabilize)
- Training overhead (time/compute cost)

Objective 4: Adaptability & Robustness
- Recovery time under workload spikes
- Stability of convergence under dynamic loads



Problem Statement

Primary Challenge: Develop an intelligent, adaptive resource allocation system for cloud computing environments that dynamically optimizes resource distribution across heterogeneous infrastructure while satisfying multiple competing objectives.



Problem Statement

Specific Problems Addressed:

1. **Dynamic Workload Management**
   - Handle unpredictable task arrivals.
   - Adapt to varying computational requirements.
   - Manage time-critical applications with strict latency constraints.

2. **Multi-Resource Optimization**
   - Balance CPU allocation across cloud nodes.
   - Optimize bandwidth utilization with variable network conditions.
   - Minimize task migration overhead while maintaining system performance.

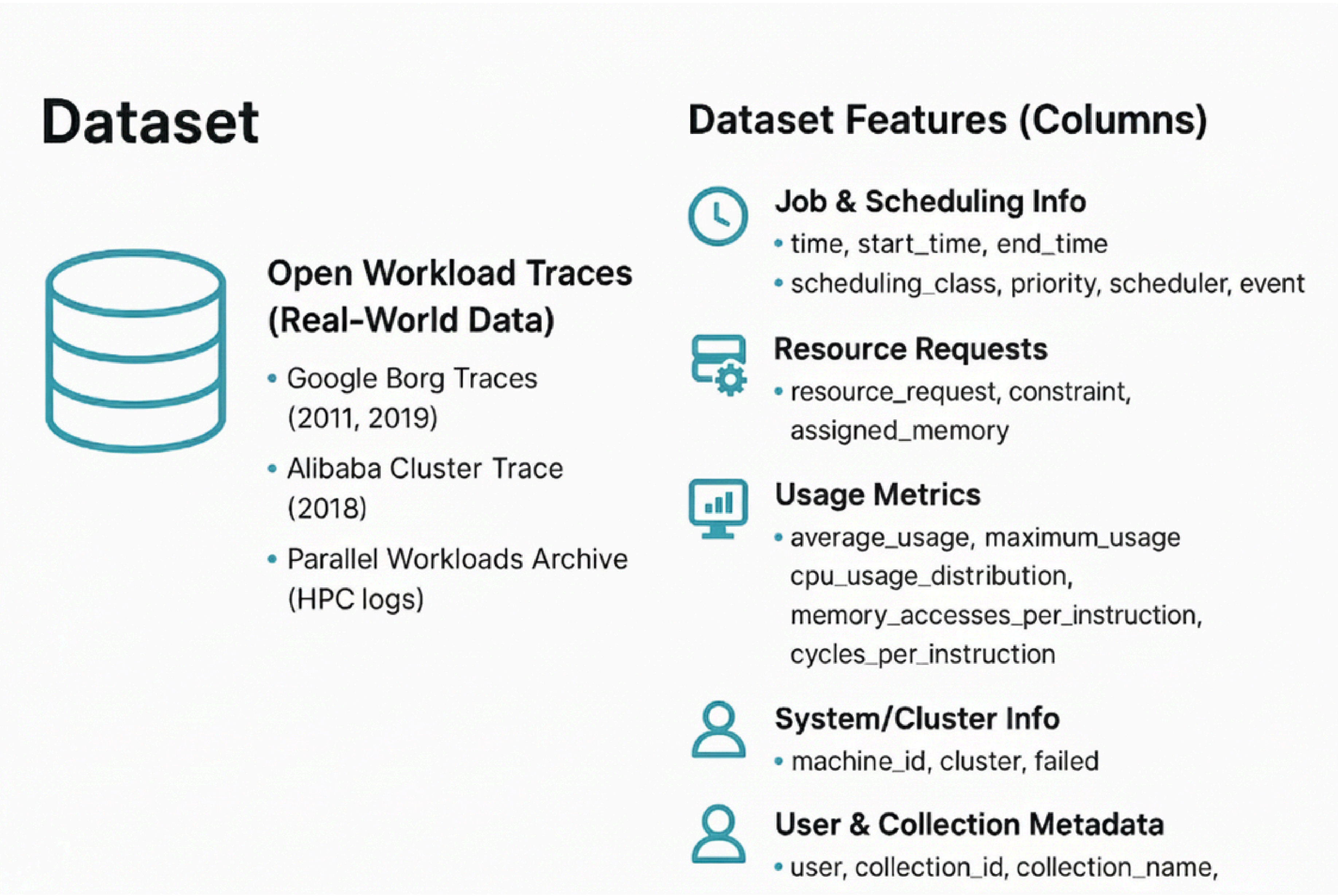
3. **Multi-Objective Goals**
   - Maximize resource utilization efficiency.
   - Minimize average processing time.
   - Control operational costs through intelligent resource allocation.

Dataset Information

Source: Open Workload Traces (Real-World Data)
- Google Borg Traces (2011, 2019)
- Alibaba Cluster Trace (2018)
- Parallel Workloads Archive (HPC logs)

Workload Details:
- Job CPU & memory usage patterns
- Job arrival rates, runtimes, and dependencies

Why These?
- Realistic & heterogeneous workloads
- Enables fair comparison with existing schedulers
- Scalable for simulation



Proposed Architecture

Indian Institute of Information Technology Kottayam

1 Input Layer

- Workload traces
- Google cluster data
- Synthetic bursts
- Available resources
- VM specifications

→

2 State Representation

- Queue length
- Resource availability
- Task size encoding
- Compact vectors
- Normalized features

✓ Training overhead reduction

→

3 DDQN Architecture

- Double DQN
- Dueling streams
- Value function V(s)
- Advantage A(s,a)
- Bias reduction

✓ Convergence stability & efficiency

→

4 Reward Function

- Multi-objective
- Completion time penalty
- Utilization reward
- Training cost penalty
- Dynamic weighting

→

5 Action Selection

- ε-greedy policy
- Task placement
- VM assignment
- Load balancing
- Exploration decay

→

6 Execution & Feedback

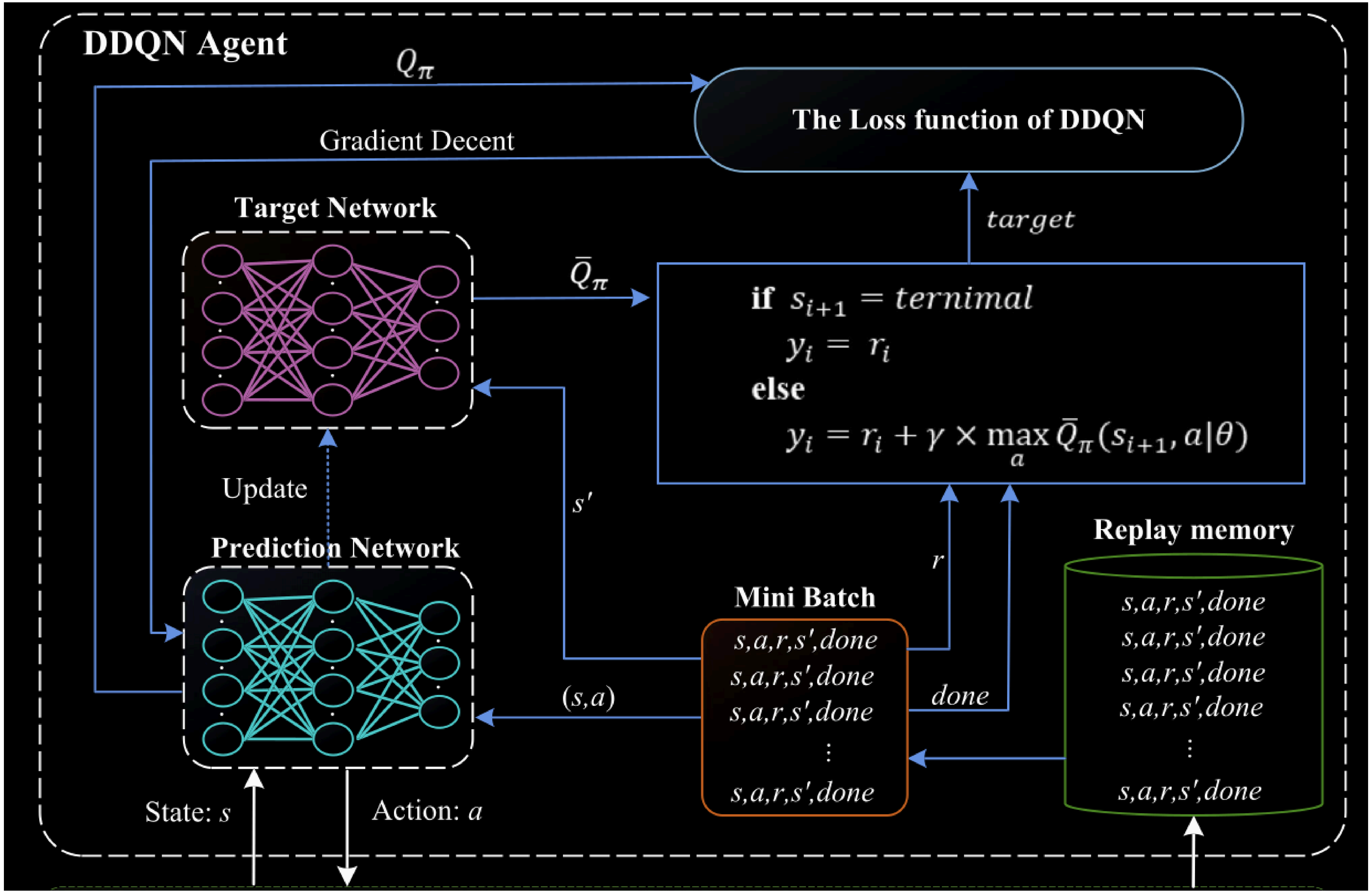
- Run allocation
- Monitor performance
- Collect metrics
- Completion times
- Utilization rates

→

7 Learning Loop

- Experience replay
- Q-network update
- Target sync
- Batch learning
- Policy improvement

Base Architecture



📍 State (S)

Current situation/configuration of the environment

⚡ Action (A)

Choices available to the agent in each state

🤖 Agent

The learner/decision maker that takes actions

🎁 Reward (R)

Immediate feedback after taking an action

🎯 Goal

Maximize cumulative reward over time

Conclusion

Indian Institute of Information Technology Kottayam

Project Summary

- **Objective:** Develop intelligent cloud resource allocation using Deep Reinforcement Learning
- **Approach:** Enhanced DDQN with dueling architecture and prioritized experience replay
- **Innovation:** Multi-objective optimization with adaptive reward functions and migration cost awareness

Performance Improvements

Improvement in resource utilization efficiency

Technical Advances

Novel adaptive reward function design

Practical Impact

Enhanced system adaptability to dynamic workloads

Timeline

Phase 1 Outcome : Build theoretical and intuitive understanding of deep learning algorithms and planning of project phases and implementation details

Indian Institute of Information Technology Kottayam

Phase 1 (Months 1-2)

Phase 2 (Months 3-4)

Phase 3 (Months 5-6)

Phase 4 (Months 7-8)

Foundation and Literature Analysis

- Comprehensive literature review and gap analysis
- Problem formulation and mathematical modeling
- Environment setup and baseline algorithm implementation
- Initial DQN implementation and testing

Algorithm Development

- Double DQN architecture implementation
- Dueling network integration and testing
- Prioritized experience replay implementation
- Reward function optimization and tuning

Experimentation and Evaluation

- Dataset validation
- Baseline comparison experiments
- Performance evaluation across different scenarios
- Statistical analysis and result validation

Analysis and Documentation

- Results analysis and interpretation
- Paper writing and technical documentation
- Presentation preparation and peer review
- Final revisions and project completion

Phase 2 Outcome: Build a simulation environment, develop and test various algorithms

References

Primary Research Papers

1. [1] Y. Wang and X. Yang, "Research on Edge Computing and Cloud Collaborative Resource Scheduling Optimization Based on Deep Reinforcement Learning," Microsoft Redmond Technical Report, 2025.
2. [2] I. Ullah, H.-K. Lim, Y.-J. Seok, and Y.-H. Han, "Optimizing task offloading and resource allocation in edge-cloud networks: a DRL approach," Journal of Cloud Computing: Advances, Systems and Applications, vol. 12, no. 112, 2023, doi: 10.1186/s13677-023-00461-3.
3. [3] Ullah, H.-K. Lim, Y.-J. Seok, and Y.-H. Han, "Optimizing task offloading and resource allocation in edge-cloud networks: A DRL approach," Journal of Cloud Computing: Advances, Systems and Applications, vol. 12, no. 112, 2023, doi: 10.1186/s13677-023-00461-3.

Supporting Literature

1. Chen, Z., Zhang, L., Wang, X., & Wang, K. (2023). "Cloud–edge collaboration task scheduling in cloud manufacturing: An attention-based deep reinforcement learning approach." Computers & Industrial Engineering, 177, 109053.
2. Xu, J., Xu, Z., & Shi, B. (2022). "Deep reinforcement learning based resource allocation strategy in cloud-edge computing system." Frontiers in Bioengineering and Biotechnology, 10, 908056.
3. Li, M., Gao, J., Zhao, L., & Shen, X. (2020). "Deep reinforcement learning for collaborative edge computing in vehicular networks." IEEE Transactions on Cognitive Communications and Networking, 6(4), 1122-1135.
4. Huang, L., Bi, S., & Zhang, Y.-J.A. (2019). "Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks." IEEE Transactions on Mobile Computing, 19(11), 2581-2593.

References

Supporting Literature

1. Xiong, X., Zheng, K., Lei, L., & Hou, L. (2020). "Resource allocation based on deep reinforcement learning in IoT edge computing." IEEE Journal on Selected Areas in Communications, 38(6), 1133-1146.
2. [1] Y. Wang and X. Yang, "Research on Edge Computing and Cloud Collaborative Resource Scheduling Optimization Based on Deep Reinforcement Learning," Microsoft Redmond Technical Report, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405844024067288
3. [2] A. D. M., "How DDPG (Deep Deterministic Policy Gradient) Algorithms Works in Reinforcement Learning," Medium, 2023. [Online]. Available: https://medium.com/@amaresh.dm/how-ddpg-deep-deterministic-policy-gradient-algorithms-works-in-reinforcement-learning-117e6a932e68
4. [3] OpenAI Spinning Up, "DDPG," [Online]. Available: https://spinningup.openai.com/en/latest/algorithms/ddpg.html
5. [4] D. K., "PPO Algorithm," Medium, 2023. [Online]. Available: https://medium.com/@danushidk507/ppo-algorithm-3b33195de14a
6. [5] Hugging Face, "Deep Reinforcement Learning with PPO," [Online]. Available: https://huggingface.co/blog/deep-rl-ppo

Foundational References

1. Mnih, V., et al. (2015). "Human-level control through deep reinforcement learning." Nature, 518(7540), 529-533.
2. Van Hasselt, H., Guez, A., & Silver, D. (2016). "Deep reinforcement learning with double q-learning." Proceedings of the AAAI Conference on Artificial Intelligence, 30(1).
3. Wang, Z., et al. (2016). "Dueling network architectures for deep reinforcement learning." International Conference on Machine Learning, PMLR, 1995-2003.

Acknowledgement

Indian Institute of Information Technology Kottayam

Project Team:

- Gaurav Nanasaheb Malave
- Hemanth Sai Manikanta
- Eedara Sri Rahul
- Pradipth Gautham S

Mentorship & Guidance:

- Dr Divya Sindhu Lekha