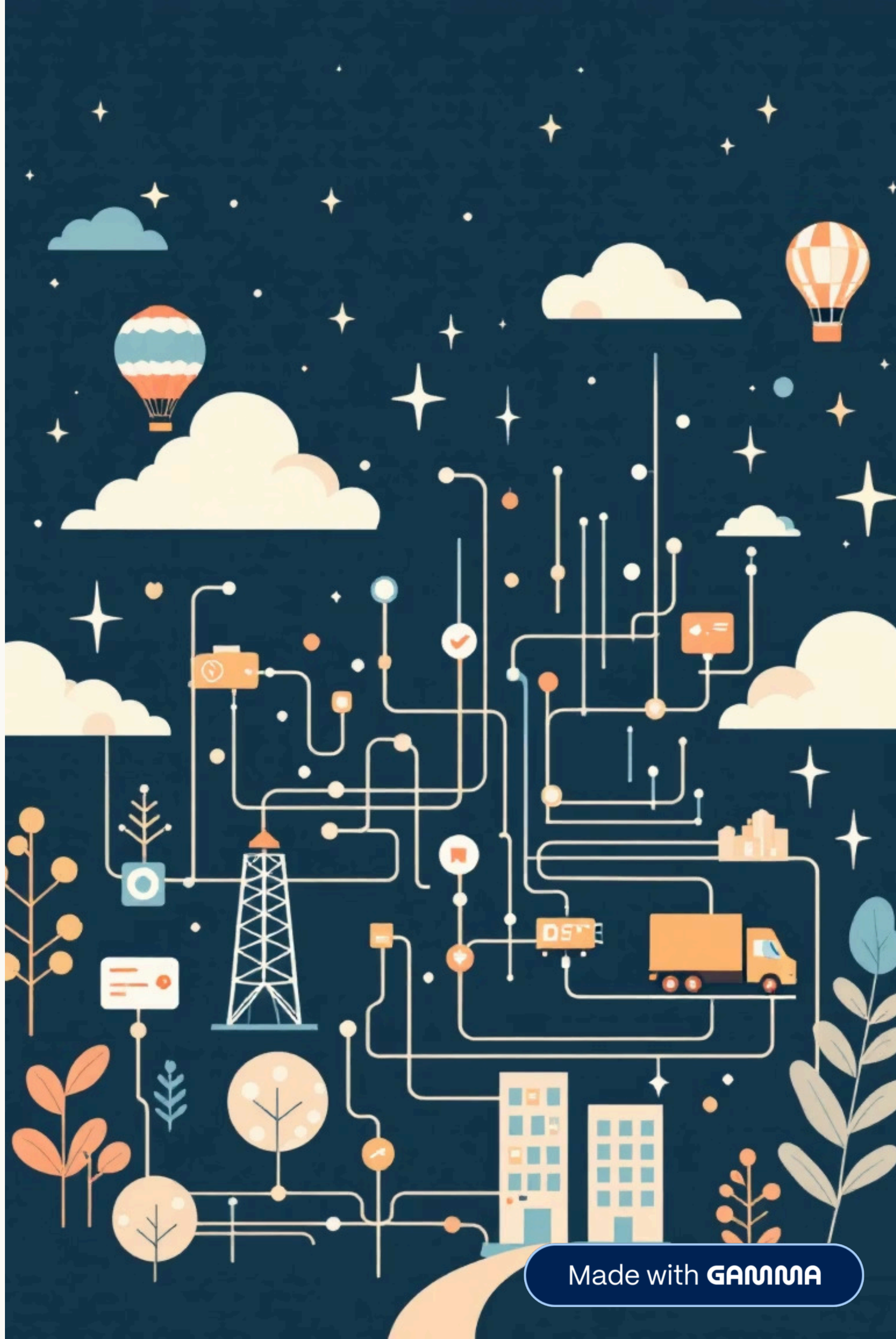
AdaptiveChain: Multi-Agent RL for Supply Chain Optimisation

Discovering the Limits of Coordination

Sravan Kumar Kurapati | INFO 7375, Northeastern University | December 2025



The £3.2 Trillion Problem: Supply Chain Disruptions

Global Economic Impact

Supply chain disruptions impose an annual cost of £3.2 trillion (approximately $4 trillion USD) on the global economy.

Widespread Corporate Vulnerability

94% of Fortune 1000 companies reported significant impacts from the COVID-19 pandemic, highlighting systemic fragility.

Limitations of Traditional Policies

Conventional inventory management policies often prove inadequate during unforeseen disruptive events, leading to substantial losses and inefficiencies.

Reinforcement Learning Agents

Our proposed solution leverages sophisticated Reinforcement Learning (RL) agents capable of developing adaptive ordering policies.

Multi-Agent Coordination

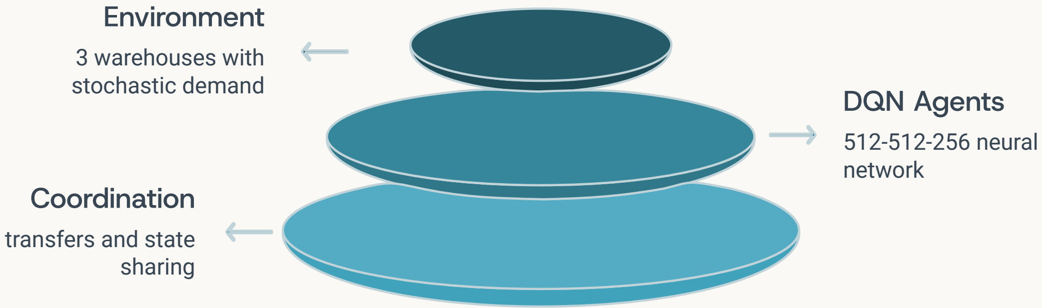
Investigating the efficacy of coordinated multi-agent systems to manage complex supply chain dynamics.

Comprehensive Evaluation

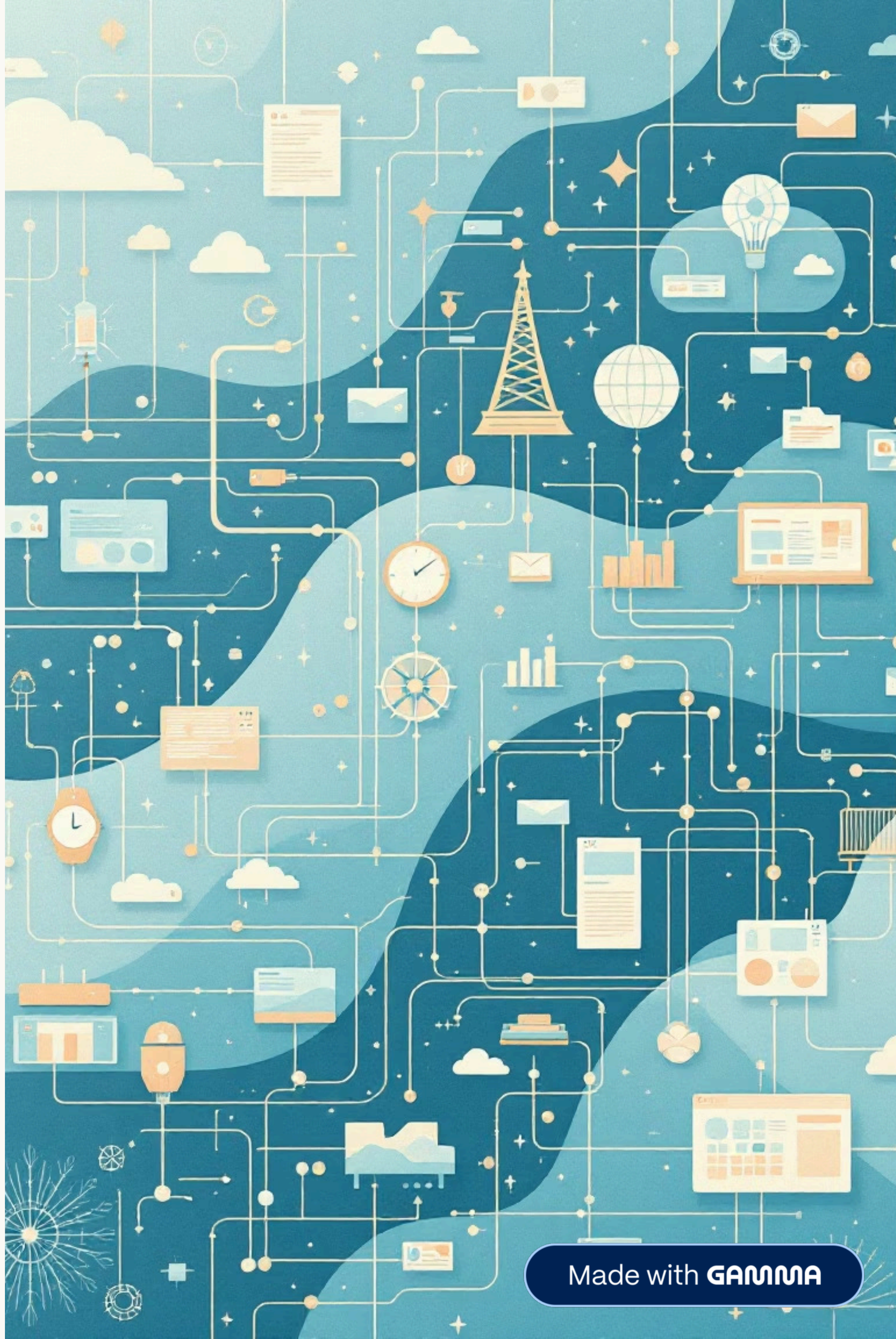
We rigorously tested 8 distinct approaches across 5 diverse supply chain scenarios to ascertain optimal strategies.

Three-Layer Architecture for Supply Chain Simulation

Our simulation framework is structured into three distinct layers, enabling a modular and comprehensive approach to supply chain optimisation.



This architecture facilitates the dynamic interaction between environmental states, intelligent agents, and coordination mechanisms, reflecting real-world supply chain complexities.



Eight Approaches Compared: Initial Findings

We evaluated a range of policies, from classical operational research methods to advanced reinforcement learning strategies, under a unified cost metric.

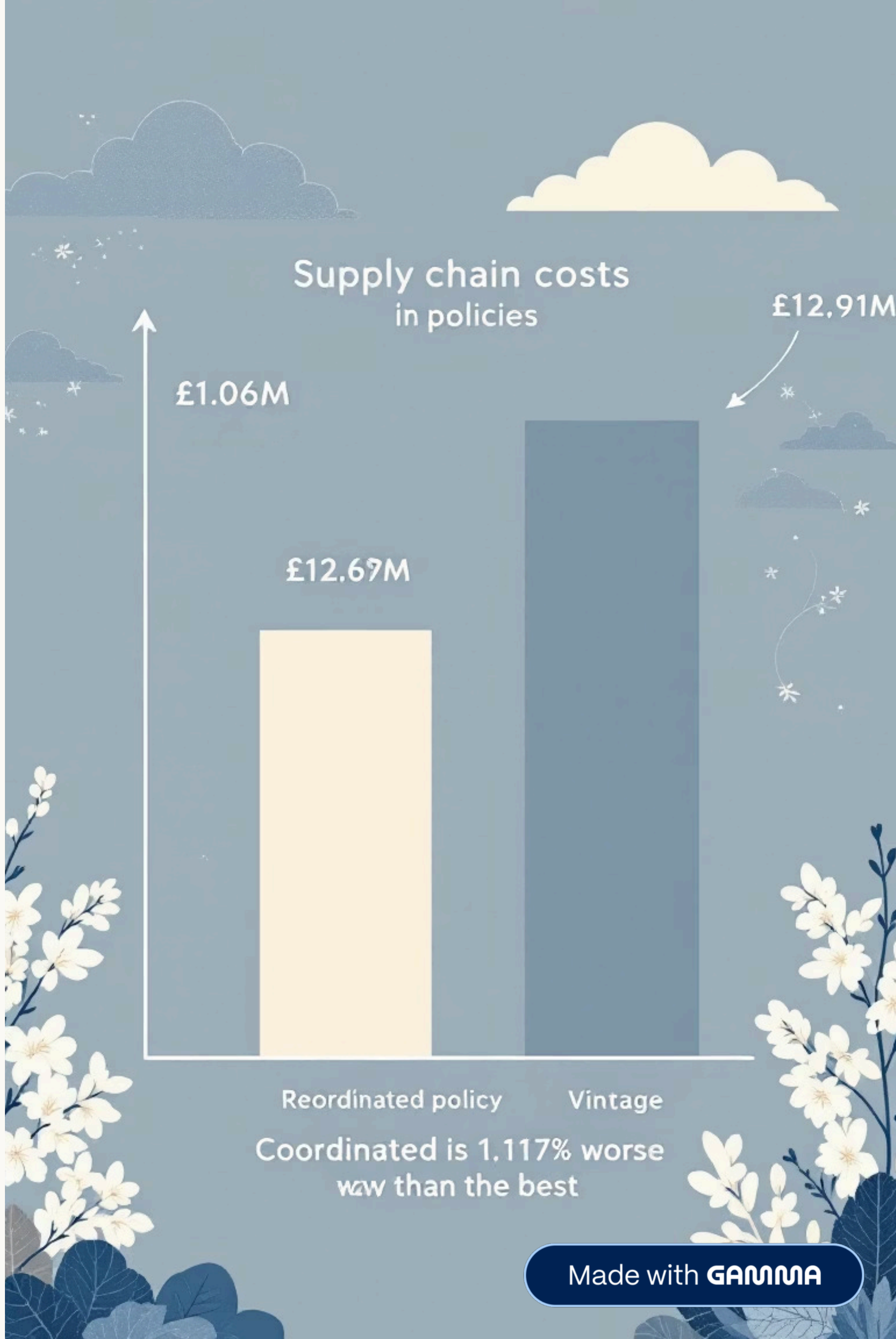
Reorder Point (1 Warehouse)	£1.06M	🥇 BEST PERFORMER
Economic Order Quantity (1 Warehouse)	£1.94M	✅ GOOD BASELINE
DQN (1 Warehouse)	£2.27M	⚠️ PARTIALLY LEARNED
Random (1 Warehouse)	£5.30M	❌ INEFFECTIVE BASELINE
Independent Agents (3 Warehouses)	£5.55M	⚠️ MODERATE PERFORMANCE
Coordinated Agents (3 Warehouses)	£12.91M	❌ CATASTROPHIC FAILURE

Notably, the simplest classical approach, Reorder Point, significantly outperformed all advanced multi-agent RL policies in initial tests.

Performance Comparison: A Stark Reality

The comprehensive performance comparison reveals a dramatic disparity between optimal classical methods and our coordinated multi-agent approach.

> **Coordinated Agents performed 1,117% worse than the best policy.**
>
> This highlights a critical challenge in achieving effective coordination within complex RL systems.



The Coordination Paradox: Why Multi-Agent RL Failed

132.8%
WORSE

Root Causes of Catastrophic Failure:

- **Excessive Transfers**

  The coordinated agents initiated 319 transfers per episode, a 6.3x increase compared to 51 for independent agents.

- **High Transfer Overhead**

  Each transfer incurred an average cost of £130K (approximately $163K USD) per episode, severely impacting overall profitability.

- **Failed Convergence**

  Despite 833 training episodes, the coordinated policy never achieved stable convergence, indicative of a fundamental learning instability.

- **Over-Reactive Policy**

  Agents developed an overly aggressive strategy, attempting to rebalance inventory at virtually every perceived imbalance, regardless of cost implications.

This analysis reveals that the coordination mechanism, while conceptually sound, was severely mismanaged by the learning algorithm.

Ablation Study: Isolating the Transfer Mechanism

To understand the source of the coordination failure, we conducted an ablation study focusing on the transfer feature itself.

Key Insight: Transfer Mechanism is Viable, But Misused

The transfer mechanism, when appropriately utilised, demonstrated the potential to save 10.3% in costs. However, the RL agent over-used it 326 times per episode.

This suggests the problem lies not with the feature itself, but with the learned policy's inability to apply it judiciously.

Learning Curves: Convergence and Divergence

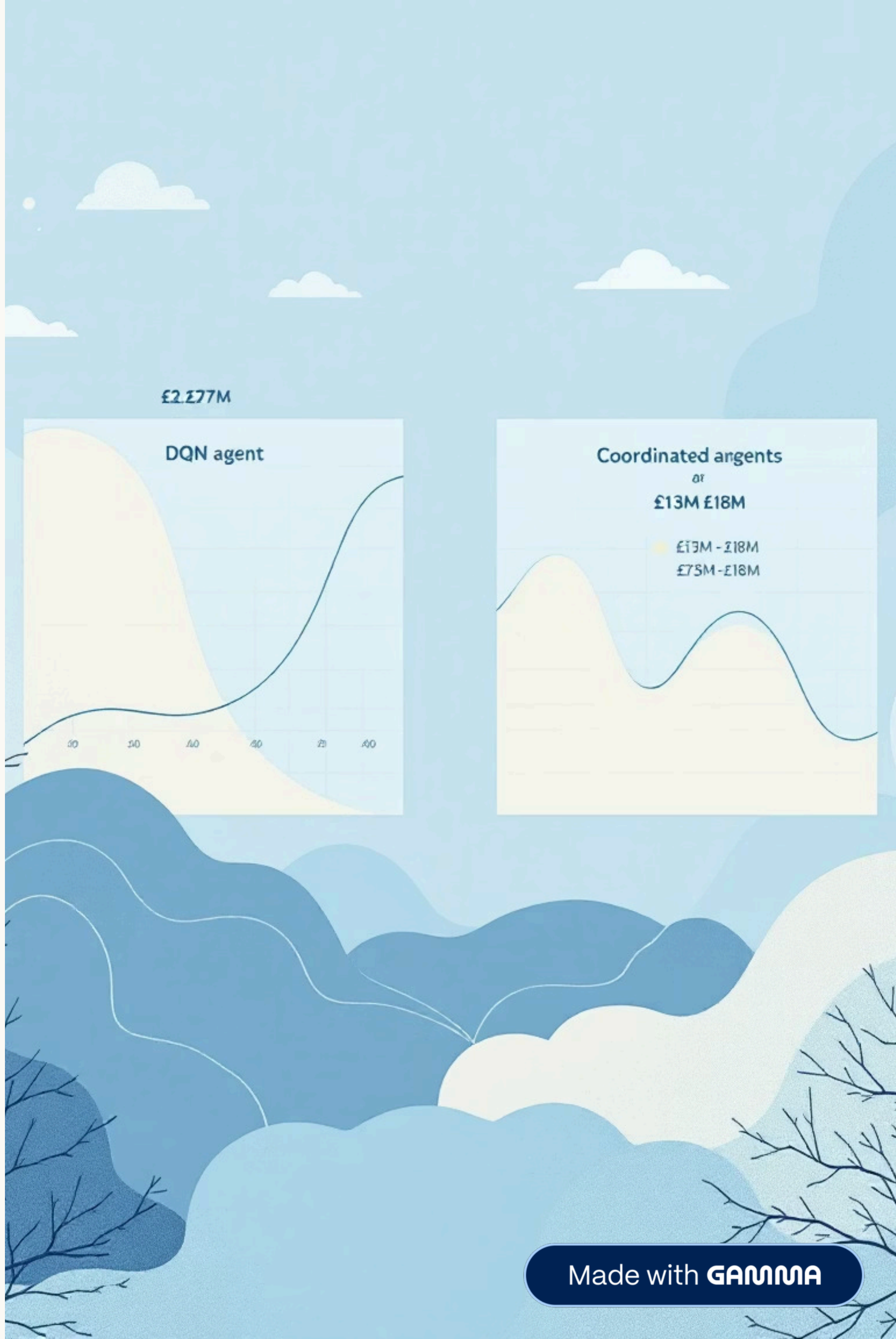
Visualising the learning process provides crucial insights into the stability and effectiveness of different agent policies over time.

DQN Convergence

The single-warehouse DQN agent achieved stable convergence within approximately 400 episodes, reaching a cost of £2.27M.

Coordinated Divergence

In stark contrast, the coordinated multi-agent system never stabilised, exhibiting wild oscillations between £13M and £18M, indicating a fundamental lack of learning.



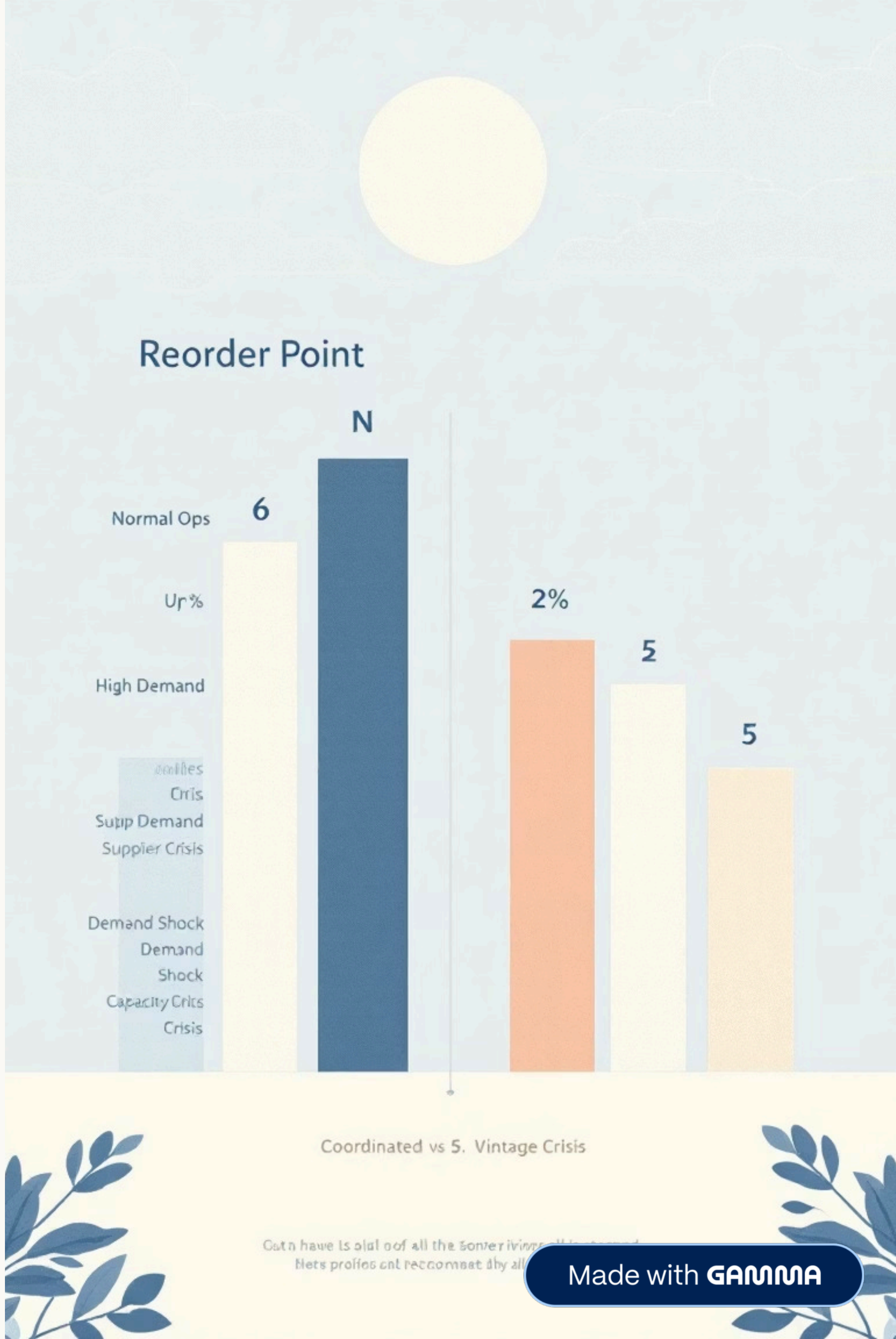
Disruption Testing: Robustness Across Scenarios

To assess the resilience of each approach, we tested them against five distinct disruption scenarios.

Systematic Validation

Across all five critical scenarios—Normal Operations, High Demand, Supplier Crisis, Demand Shock, and Capacity Crisis—the classical Reorder Point strategy consistently outperformed all other approaches.

Conversely, the Coordinated multi-agent policy consistently performed the worst, demonstrating that its failure is systematic and not merely a random anomaly.



Key Takeaways & Live Demonstration

Critical Insights from AdaptiveChain:

1 Classical Methods Prevail

Empirical evidence suggests that classical operational research methods remain highly effective, achieving optimal costs of £1.06M.

2 RL's Optimisation Gap

Reinforcement Learning agents can learn adaptive policies, but did not reach global optimality in this context. They performed 57% better than random, but 114% worse than the optimal classical approach.

3 Coordination's Detriment

Unchecked coordination in multi-agent RL can severely degrade performance, resulting in costs 133% higher than independent agents.

4 Empirical Validation is Paramount

Rigorous testing and comparative analysis against established benchmarks are essential for validating the efficacy of advanced RL solutions.

Interactive Dashboard Features:

- Real-time agent simulation
- Performance comparison tools
- Learning curve visualisation
- Scenario testing functionalities
- Comprehensive statistical analysis

GitHub: https://github.com/sravankumarkurapati/INFO_7375/tree/main...

Resources: 18-page report, 22 code files, 13 vis...