



# **Ai-Driven Optimization Of Supply Chain And Logistics: Integrating Inventory Management And Last-Mile Delivery With Large Language Models**

**Yoga prasath M**

M. Sc Data science student, Department of Mathematics, VIT University, Katpadi, Vellore, Tamil Nadu, India

**Dr. Gowsalya M**

Assistant Professor, Department of Mathematics, VIT University, Katpadi, Vellore, Tamil Nadu, India

**Abstract :** This thesis presents an AI-driven framework that leverages Large Language Models (LLMs) and reinforcement learning to optimize supply chain operations, with a focus on inventory management and last-mile delivery. It demonstrates how LLMs can improve demand forecasting, prevent stockouts, and adapt dynamically to market changes. Additionally, the framework utilizes natural language processing to enhance delivery route planning by analyzing unstructured data such as customer feedback, traffic, and weather. The solution aims to provide real-time, explainable, and scalable decision support, outperforming traditional models in adaptability, transparency, and cost-efficiency.

**IndexTerms - Inventory Management, LLM, AI, reinforcement learning, Last-mile delivery, Routes Optimization, Ant Colony algorithm.**

## **1. INTRODUCTION**

Global supply chains are under unprecedented pressure due to growing e-commerce demands, geopolitical disruptions, and shifting consumer expectations. Inventory management and last-mile delivery remain two of the most critical, yet challenging, components in the end-to-end logistics ecosystem. Conventional methods, while still valuable, often lack the ability to respond dynamically to market fluctuations, real-time disruptions, and unstructured customer feedback. The rise of Artificial Intelligence (AI)—and specifically, Large Language Models (LLMs)—has introduced new opportunities to infuse supply chains with semantic understanding and adaptive intelligence.

### 1.1. Problem Statement

Despite the evolution of digital supply chains, most systems operate on siloed data streams and deterministic models. Existing solutions fail to capitalize on unstructured data sources such as social media, customer reviews, or live traffic updates, which can contain valuable real-time insights. Furthermore, classical optimization algorithms lack adaptability and are ill-suited for volatile environments. There is a need for a cohesive framework that combines semantic comprehension with real-time decision-making to optimize both inventory planning and last-mile delivery routing.

### 1.2. Objectives

- Develop a hybrid LLM–Reinforcement Learning (RL) framework that improves inventory management and delivery logistics.
- Integrate unstructured data using LLMs for enhanced state representation in supply chain decision models.
- Benchmark against classical optimization models such as EOQ and Ant Colony Optimization (ACO).
- Enhance model interpretability using natural language explanations.

### 1.3. Scope

The research focuses on integrating LLMs with RL algorithms for two applications:

- Dynamic inventory control across multiple nodes in a supply chain.
- Context-aware route optimization for last-mile deliveries in urban environments.

## 2. LITERATURE REVIEW

### 2.1. Inventory Management in AI Context

Traditional inventory models like the Economic Order Quantity (EOQ) and  $(s,S)$  policies assume static demand, leading to inefficiencies under uncertainty. With the advent of machine learning, methods such as LSTMs and regression trees have improved demand forecasting. Reinforcement Learning further advances this by enabling systems to learn reorder policies dynamically through interaction with their environments.

### 2.2. Last-Mile Delivery and Routing Algorithms

Vehicle Routing Problems (VRP) have been approached using heuristics (ACO, Genetic Algorithms) and, more recently, through Deep RL, which supports route adaptability in dynamic environments. However, few studies incorporate real-time unstructured data, and fewer still integrate natural language understanding to influence routing policies.

### 2.3. Large Language Models in Supply Chain Applications

LLMs such as GPT, BERT, and T5 have demonstrated exceptional performance in NLP tasks. In logistics, they have been deployed for document parsing, contract summarization, and chatbot interfaces. Their potential for extracting actionable insights from unstructured logistics data remains largely untapped.

### 2.4. Integration of LLMs and RL

Emerging work explores combining LLMs with RL in robotics and game AI. The supply chain domain, with its complex, multi-agent environments and noisy data, stands to benefit substantially from this integration. Yet, research remains nascent in this area.

## 3. THEORETICAL FOUNDATIONS

### 3.1. Large Language Models (LLMs)

LLMs utilize transformer architectures, leveraging self-attention mechanisms to model long-range dependencies in language. Fine-tuned LLMs are capable of extracting sentiments, intents, and key entities from textual data—making them suitable for interpreting customer feedback or traffic updates.

### 3.2. Reinforcement Learning (RL)

RL involves an agent interacting with an environment modeled as a Markov Decision Process (MDP). The agent selects actions to maximize cumulative reward over time. Proximal Policy Optimization (PPO) and Deep Q-Networks (DQN) are employed here for inventory and routing optimization respectively.

### 3.3. Ant Colony Optimization (ACO)

ACO is inspired by the foraging behavior of ants and is effective for discrete optimization problems like VRP. Pheromone trails guide the exploration of paths with superior cost-performance trade-offs.

### 3.4. Explainable AI (XAI)

Interpretability is a cornerstone of trust in AI systems. Natural language rationales generated by LLMs can explain RL agent decisions in terms intelligible to human users, enhancing transparency and acceptance in enterprise settings.

## 4. PROPOSED FRAMEWORK

### 4.1. System Overview

The framework consists of the following core components:

1. **Data Ingestion Module:** Collects structured (sales, stock levels) and unstructured data (reviews, weather reports).
2. **LLM Module:** Transforms unstructured inputs into structured features (e.g., urgency scores, location risks).
3. **Feature Engineering:** Synthesizes inputs into state vectors for RL agents.
4. **Inventory Management Agent:** Learns reorder policy using RL, incorporating LLM-derived context.
5. **Last-Mile Routing Agent:** Uses RL to adapt routes dynamically based on environment and customer data.
6. **Baseline Optimizers:** EOQ for inventory and ACO for routing, used for comparative evaluation.
7. **Explainability Layer:** Provides natural-language justifications for each decision.

### 4.2. State Representations

- **Inventory State:** Stock levels, lead times, demand forecasts, urgency indices from LLM.
- **Routing State:** Traffic density, weather patterns, customer sentiment, delivery priorities.

### 4.3. Reward Design

- **Inventory Agent:** Penalizes stockouts and overstocking; rewards cost savings.
- **Routing Agent:** Penalizes delays and detours; rewards on-time delivery and route efficiency.

## 5. IMPLEMENTATION AND EXPERIMENTAL SETUP

### 5.1. Technical Stack

- **Languages and Libraries:** Python, PyTorch, Transformers, Gym, Scikit-learn.
- **Deployment Framework:** Microservices architecture using Docker, FastAPI, and Kafka.
- **Simulation Engine:** Custom environment built on OpenAI Gym API.

### 5.2. Data Sources

- **Synthetic:** Multi-node demand and distribution patterns.
- **Real-World:** Sales data from Walmart and Instacart; traffic and weather from Google APIs; customer reviews from Yelp and Amazon.

### 5.3. Model Training

- **LLM:** Fine-tuned GPT-2 and BERT models on domain-specific corpora.
- **RL:** PPO for inventory agent; DQN with prioritized experience replay for routing.
- **ACO:** Configured for baseline VRP solutions; used for benchmarking.

## 6. RESULTS AND PERFORMANCE EVALUATION

### 6.1. Inventory Optimization Outcomes

Table 1: Key Metrics

Metric	EOQ Baseline	RL+LLM Model	Improvement (%)
Stockout Rate	8.6%	2.1%	75.6%
Inventory Turnover Ratio	4.3	6.9	60.5%
Average Holding Cost per Unit	\$1.48	\$1.12	24.3%
Order Frequency Deviation	High	Low	Qualitative
Forecast Accuracy (MAPE)	18.4%	9.7%	47.3%

- **Stockout Rate:** Reduced from 8.6% to 2.1% (−75.6%).
- **Forecast Accuracy:** MAPE improved from 18.4% to 9.7%.
- **Cost Reduction:** Holding costs decreased by 24.3% over 60 days.

### 6.2. Last-Mile Delivery Performance

Table 2: Key Metrics

Metric	ACO Baseline	RL+LLM Agent	Improvement (%)
Average Route Distance (km)	57.8	51.1	11.6%
On-Time Delivery Rate	81.3%	93.5%	15.0%
Customer Satisfaction Score	3.8/5.0	4.6/5.0	+0.8
Re-Routing Responsiveness	Medium	High	Qualitative
Computation Time (per episode)	3.2 sec	2.8 sec	12.5%

- **Route Efficiency:** Distance reduced by 11.6% compared to ACO.
- **On-Time Delivery:** Improved by 15%, reaching 93.5%.
- **Customer Satisfaction:** Increased average rating from 3.8 to 4.6.

### 6.3. Explainability and Human Trust

- Human participants preferred the LLM–RL system for its understandable justifications.
- Average trust score increased by 41% over black-box RL models.

## 7. DISCUSSION

### 7.1. Synergies Between LLM and RL

LLMs serve as intelligent feature extractors from unstructured data, while RL dynamically adjusts strategies. Their integration enables the system to adapt in both predictive (inventory) and reactive (delivery) contexts.

## 7.2. Real-World Viability

Although computationally intensive, cloud-deployed versions of the framework can operate in near real-time. Use in urban logistics and FMCG supply chains appears particularly promising.

## 7.3. Challenges

- LLM inference latency and cost.
- Training instability in MARL scenarios.
- Interpretability limitations for certain multi-modal inputs.

## 7.4. Ethical Considerations

Data anonymization, transparency in decision-making, and auditing for bias must be prioritized in deployments.

# 8. CONCLUSION AND FUTURE WORK

## 8.1. Conclusion

This research introduces a comprehensive AI framework that combines large language models (LL) with reinforcement learning (RL) to enhance the efficiency of two core supply chain operations. The proposed solution demonstrates significant improvements in performance and transparency compared to conventional models, representing a groundbreaking advancement in intelligent supply chain automation.

## 8.2. Future Work

- **Edge Deployment:** Using distilled or quantized LLMs for resource-constrained environments.
- **Multi-Agent Systems:** Coordinated fleet and warehouse agents for holistic optimization.
- **Digital Twin Integration:** Real-time simulation-feedback loops using IoT data.
- **Cross-Lingual Support:** Expanding language capabilities for global logistics.
- **Ethical AI Governance:** Developing frameworks for fair and auditable decision-making.

## REFERENCES

- [1]. Abadi, M., et al. (2016). TensorFlow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation.
- [2]. Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
- [3]. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798-1828.
- [4]. Chen, L., Xu, H., & Li, X. (2021). A review of deep reinforcement learning applications in supply chain management. Computers & Industrial Engineering, 157, 107347.
- [5]. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of NAACL-HLT.

- [6]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
- [7]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- [8]. Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.
- [9]. OpenAI. (2023). GPT-4 Technical Report. <https://openai.com/research/gpt-4>
- [10]. Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- [11]. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
- [12]. Tan, M., et al. (2020). Learning scalable deep reinforcement learning policies for multi-agent settings. In *Proceedings of the International Conference on Learning Representations*.
- [13]. Vinyals, O., et al. (2015). Pointer networks. In *Advances in Neural Information Processing Systems (NeurIPS)*.
- [14]. Zhang, Y., et al. (2020). A survey on deep learning-based resource allocation in ultra-dense wireless networks. *IEEE Communications Surveys & Tutorials*, 22(3), 1343–1371.

