OPTIMIZING AI AND MACHINE LEARNING TECHNOLOGIES TO IMPROVE LAST-MILE DELIVERY EFFICIENCY IN QUICK-COMMERCE

By

SRINIVASAN K. P., MBA

Supervised by

SAGAR BANSAL, DBA.

DISSERTATION

Presented to the Swiss School of Business and Management, Geneva

In Partial Fulfilment of the Requirements for the Degree

DOCTOR OF BUSINESS ADMINISTRATION

SWISS SCHOOL OF BUSINESS AND MANAGEMENT GENEVA

APRIL 2026

OPTIMIZING AI AND MACHINE LEARNING TECHNOLOGIES TO IMPROVE LAST-MILE DELIVERY EFFICIENCY IN QUICK-COMMERCE

By

SRINIVASAN K. P., MBA

Supervised by

SAGAR BANSAL, DBA.

APPROVED BY:

Apostolos Dasilas

Dissertation Chair: Apostolos Dasilas, PhD

RECEIVED/APPROVED BY:

Renee Goldstein Osmic

Admissions Director

Dedication

To my beloved wife, Pavithra, and my children, Kesav Pranav and Kesav Sai. Your unconditional love, patience, and care have been my greatest strength throughout this journey. Your faith in me has given me the courage to persevere, even in the most challenging moments. This achievement is not mine alone, it belongs to you as much as it does to me. You are my constant source of inspiration, and I dedicate this work with gratitude and love.

Acknowledgments

This research would not have been possible without the meticulous guidance and support of my research supervisor, Dr. Sagar Bansal. His mentorship not only provided academic direction but also shaped the analytical lens through which I approached the complex world of the Quick-Commerce ecosystem. His dedication, clarity of thought, continuous gentle push, and constant encouragement helped me remain focused and resilient throughout this journey. I am deeply grateful for his belief in the work and in me.

I would also like to thank my wife, Mrs. Pavithra, for her invaluable support and editorial assistance. Her thoughtful editorial suggestions ensured that this thesis communicates clearly and effectively to both academic and industry readers. Her patience, care, and quiet strength kept me balanced, and her unwavering belief in me has been my anchor.

My heartfelt gratitude also extends to my children, Kesav Pranav and Kesav Sai, whose love and innocence brought me joy during this demanding journey. Despite me spending their weekend hours for this thesis, they reminded me every day of the bigger purpose beyond academic pursuits.

I am sincerely thankful to the various Quick-Commerce startups, policy stakeholders, subject matter experts, fleet drivers and domain advisors who shared their experiences and insights so generously. Their real-world perspectives were invaluable in grounding this research in practical relevance and bridging theory with practice.

Lastly, I am thankful to everyone, particularly friends, colleagues, mentors, and institutions, who played a role, however small, in supporting this academic journey. This study stands on the foundation of your generosity, belief, and trust.

ABSTRACT

OPTIMIZING AI AND MACHINE LEARNING TECHNOLOGIES TO IMPROVE LAST-MILE DELIVERY EFFICIENCY IN QUICK-COMMERCE

by

SRINIVASAN K. P.

APRIL 2026

Dissertation Chair: Apostolos Dasilas, PhD

This study explores the optimization of Artificial Intelligence and Machine Learning (AI/ML) technologies to enhance efficiency in last-mile delivery within Quick-Commerce industries. Consumer expectations for fast and reliable delivery are increasing; therefore, it is essential to assess the effectiveness of existing AI/ML technologies. This thesis addresses the key objectives such as assessing the performance of current algorithms, identifying biases in the existing tech stacks, and designing a hybrid AI/ML technology framework tailoring to the unique needs of last-mile delivery in Quick-Commerce. To explore algorithmic strengths and weaknesses, this study deploys a mixed-methods research approach which combines qualitative insights and quantitative analysis. The findings from this study specify the opportunities to optimize and address the biases in AI/ML algorithms that can affect performance and efficiency. The proposed hybrid model addresses the challenges within Quick-Commerce and lays the foundation for future innovation in last-mile delivery. This thesis bridges theoretical trends with practical applications and provides insights for industry professionals to optimize last-mile delivery processes and contributes to the academic discussion on AI/ML optimization in last-mile delivery of Quick-Commerce.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Background and Context

The renaissance of Quick-Commerce has significantly transformed the global retail and logistics sector, anchored by its lightning-fast delivery of 10 to 90 minutes. This innovation in consumer experience has transformed consumer expectations, placing unmatched demands on supply chains, and challenging the traditional logistics framework. Last-mile delivery is the most critical and high-cost segment of the logistics value chain. It must evolve and adapt to meet the profound challenges for speed, performance, and efficiency.

The rapid advancement of Quick-Commerce underpins a robust and effective last-mile delivery approach that can withstand an array of complexities. This segment of the logistics value chain is associated with challenges, including volatile order placement, dynamic traffic patterns, and geographic constraints. Also, last-mile delivery accounts for over 53% of total logistics costs, which reiterates the need for strategic improvements. Features such as unpredictable consumer demand, urban traffic congestion, and growing environmental concerns further embroil last-mile operations, hence demanding advanced automated solutions for optimal performance and consumer delight.

The importance of last-mile logistics deliveries has increased drastically which leads to increased competition among the E-Commerce vendors for “faster delivery.” These vendors aspire to target ambitious service level agreements (SLA) by setting up highly optimized supply chains comprising micro-fulfillment centers and dark warehouses. This delivery network enables seamless “Order-To-Cash” cycle by executing faster order processing and dispatch, facilitating quicker turnaround times for consumer orders (Al Rahib, 2024).

The key Quick-Commerce categories which are considered for this study are as illustrated in figure 1.1a.

Figure 1.1a: Key Quick-Commerce Categories (adapted from Al Rahib, 2024)

Particularly, the Quick-Commerce categories like Consumer-Packaged Goods (CPG), Food and Beverage are adopting hyperlocal delivery approaches, offering essential products with quick turnaround. This innovative approach is feasible by harnessing the power of location intelligence. By leveraging spatial data, companies can effectively troubleshoot demand hotspots, optimize warehouses positions, and streamline last-mile dispatch, all of which yield to faster deliveries. Similarly, AI powered "Route Optimization" plays a key role.

As Quick-Commerce vendors expedite their focus on last-mile logistics, supply chain strategies driven by location intelligence and AI analytics are gaining traction and becoming transformative factors in the industry by 2025-26. This study focuses on these dynamics, investigating the intersection of AI/ML technologies and logistics parameters within the Quick-Commerce perspective. When optimizing the last-mile delivery ecosystem, several key metrics must be evaluated to ensure efficiency, usefulness, and continuous improvement. The high impact, key metrics of last-mile delivery in Quick-Commerce are displayed in Figure 1.1b.

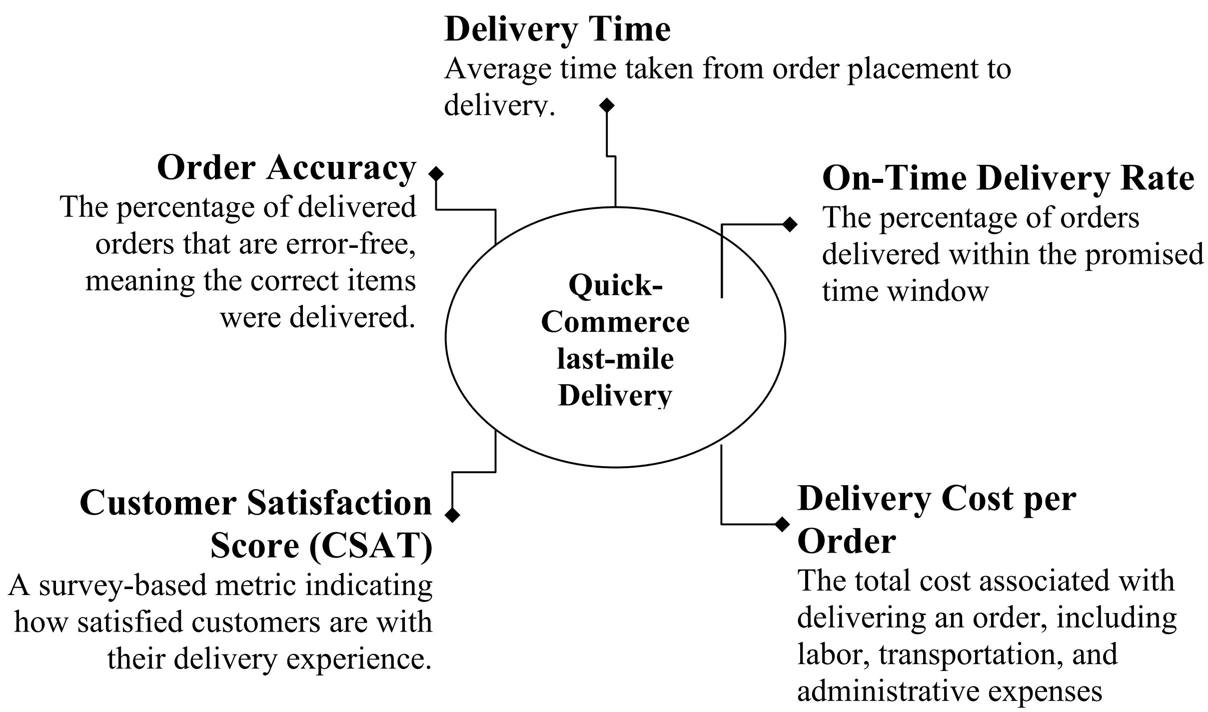

Figure 1.1b: Priority Metrics in last-mile Delivery (original work)

To realize the above-mentioned metrics and address complexities, businesses are increasingly focused on AI/ML. Last-mile in logistics value-chain contributes more than 53% of total shipping costs, making it a significant area of focus for optimization (Sharma, 2024). Quick-Commerce platforms are characterized by fast-track delivery within 10-90 minutes but face challenges, including volatile demand spikes and urban congestion. AI/ML algorithms have demonstrated strong capabilities across various aspects of last-mile value chain. Dynamic vehicle routes often leverage AI/ML models such as Genetic

Algorithms (GA) and Ant Colony Optimization (ACO) for real-time planning, although scalability remains a hurdle in highly dynamic environments (Ferreira, 2025). Delivery time prediction models, like Neural Networks (NN) and Regression-based methods (RM), provide accurate delivery-window estimation, yet their dependency on static datasets often diminish adaptability to volatile conditions. Reinforcement Learning (RL) techniques are applied to optimize resource allocation and delivery workforce deployment; however, they commonly require elaborate training data, which limits effectiveness in new or less penetrated regions (Droomer, 2020)

Bias in algorithms, which often arise from training datasets are recognized as critical bottlenecks. For example, geographic biases may downgrade certain regions, leading to delay in delivery and customer displeasure. Mishra (2024) highlights the disruptive role of technology innovations and process excellence in addressing key challenges in last-mile delivery, reinforcing AI and machine learning solutions that enhance both cost efficiency and customer satisfaction. The study further highlights the importance of scalable technology stacks that only deal with operational constraints but also promote fairness across the delivery track. The author, Mishra investigates the mindset of four specific stakeholders in Quick-Commerce landscape: online buyers who require on-time delivery, E-Commerce platform providers who streamline digital purchases, Warehouse and Logistics experts responsible for executing a seamless supply chain, and last-mile delivery partners who prefer adhering to service-level agreements. Based on this multi-stakeholder assessment, the scholar proposes a structured framework summarizing the technologies appropriate for each stakeholder level, as demonstrated in Figure 1.1c.

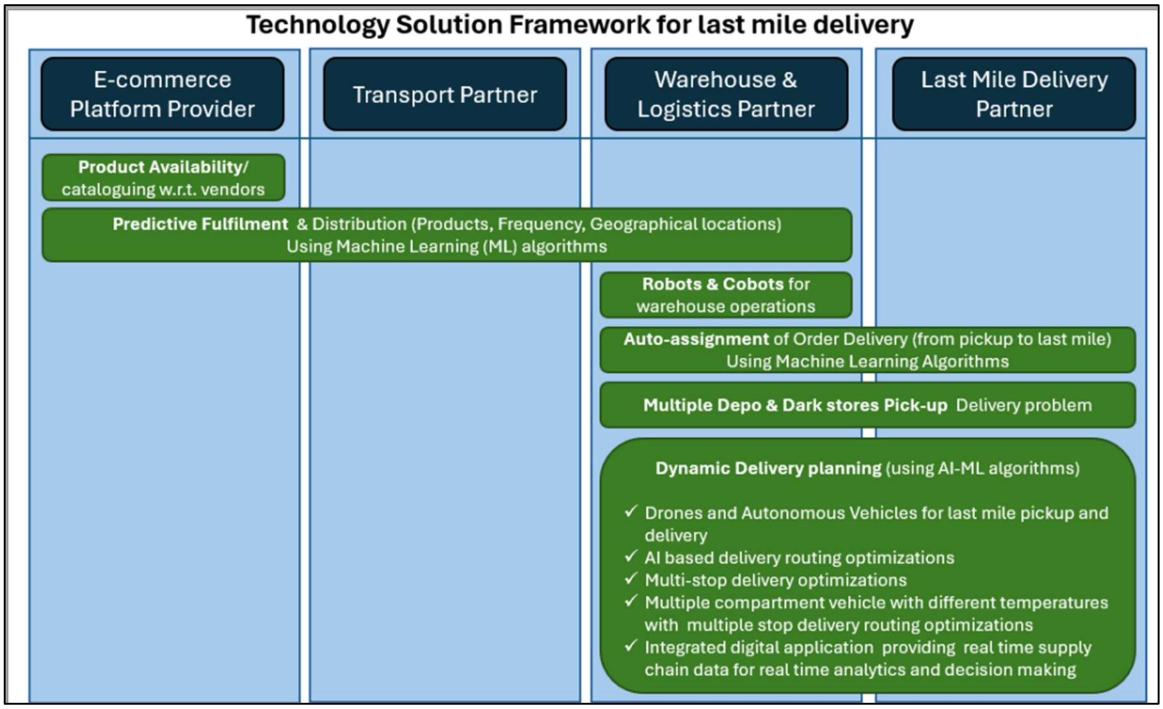

Figure 1.1c: Technology Stack for last-mile delivery (adapted from Mishra, 2024)

Despite these advancements, the literature (as we shall see in detail in Chapter 2: Literature Review) discloses a major gap in the context of AI/ML technology stack integration. In the domain of optimizing last-mile delivery for Quick-Commerce by leveraging AI/ML technologies, several research gaps persist that demand further examination:

Lack of Extensive Comparative Studies: Current research often screens AI/ML algorithms separately, under varying conditions, which restricts the ability to make direct comparisons. There is a requirement for research that implement and analyse multiple algorithms under the same conditions to understand their relative strengths and weakness in last-mile delivery situations.

Limited Quest of Hybrid Approaches: Many literatures explore on single-algorithm solutions like Neural Networks, Regression or Reinforcement Learning, which do not leverage the holistic benefits provided by combining different AI/ML algorithms. Navigating hybrid approaches, which integrate various AI/ML models and methodologies, could deliver precise and robust solutions by leveraging on the unique features of each algorithm or method.

Detect and Mitigate Bias: There is a pivotal research gap to identify and mitigate biases within last-mile delivery algorithms. Biases result in unfair and faulty delivery practices, impacting both operational integrity and consumer delight. Extensive studies are needed to develop methodologies for detecting such biases and creating strategies to tackle them.

Lack of Integrated Hybrid Tech Stacks: In-spite of AI/ML advancements in last-mile delivery, there is a shortage of research on building hybrid technology stacks that incorporate multiple algorithms and templates. These hybrid stacks could enhance system performance, scalability, and flexibility in volatile delivery value chain. Resolving these gaps has the potential to significantly advance the field by providing integral, scalable, and ethical AI/ML applications in Quick-Commerce sector. Focused research in these areas can lead to more refined and seamless rollouts that align with the increasing industry demands. The current study is designed to address these gaps through four key contributions. Beginning, it provides a thorough performance analysis of existing AI/ML algorithms implemented in last-mile delivery. Second, it identifies and investigates algorithmic bias and its operational impacts within deliver system. Third, it evolves and validates an upgraded hybrid technology stack to optimize vehicle routing and delivery. Finally, it sets up a strategic recommendation to support the selection and roll-out of AI/ML technologies in Quick-Commerce. This structured research methodology assures a rigorous approach to evaluate, optimize, and recommend AI/ML solutions for last-mile delivery, with a precise target on operational efficiency and fairness across stakeholders.

1.2 Research Problem and Rationale

The accelerated growth of Quick-Commerce has transformed consumer expectations, demanding spontaneous delivery services that rely on organized last-mile logistics. In-spite of AI/ML technologies providing significant guarantee for streamlining these delivery processes, vital challenges restrict their full potential and operational success.

Primarily, the capability of current AI/ML algorithms is often compromised under real-world conditions. These methodologies must cope with various delivery loads, complex urban traffic patterns, and geographically diverse routes, yet many existing solutions fail to deliver consistent performance across the dynamic ecosystems. Therefore, the versatility and scalability of AI/ML models are crucial to confirming that delivery services meet the evolving volatile demands of Quick-Commerce in a reliable and efficient way. Moreover, the presence of biases within algorithmic processes remains a crucial issue. Demographic disparities and unequal prioritization of delivery zones can result in inequitable service distribution, negatively impacting customer service and leading to reputational damage for service providers. Mitigating these biases is crucial to achieving equitable and seamless delivery services, promoting fairness and consistency across all customer engagements (Sooroshian, 2024).

Figure 1.2 illustrates the major operational challenges of Quick-Commerce which mandates the adoption of AI/ML solution stack:

Figure 1.2: Operational Challenges (adapted from Sooroshian, 2024)

The evolution of Quick-Commerce has introduced a range of operational complexities that significantly hamper the health index of last-mile delivery services. These challenges arise from several key factors that must be addressed to optimize the operational excellence of Quick-Commerce. To start with, expanding logistics silos result in disconnected processes that restricts seamless engagement across different segments of supply chain. This lack of integration affects inventory management and forecasting accuracy, leading to operational wastes and stock variances. Poor inventory management not only influences the ability to meet consumer demand but also escalates costs associated with overstock or stockouts. Next, the disorganized fleet management of delivery vehicles aggravates operational inefficiencies, contributing to increased fuel consumption and higher running costs. Without reliable scrutiny and optimization of fleet utilization, organisations face barriers in managing competitive delivery timelines. In addition, the lack of vision for customer-centric and scalability goals poses a barrier to long-term vision. Businesses that fail to embed these practices risk losing market share.

The inadequacy to optimize asset usage led to resource neglect and lost opportunities for improving service delivery. Moreover, oscillating market demand complicates planning and resource allocation, resulting in fluctuations that impact logistics operations. These integrated issues result in higher consumer acquisition costs, as organisations must incur significant costs to attract and retain consumers in a competitive digital era. The lack of real-time insights further worsens decision-making processes, preventing businesses from responding faster to market changes and operational dilemmas. Addressing these versatile problems is essential for optimising the last-mile delivery services in Quick-Commerce. This study aims to explore innovative AI/ML solutions that can streamline workflows, reduce cost, and promote sustainable practices which leads to improved performance and consumer experience.

1.3 Research Objectives

This study is tailored to address critical elements of optimizing last-mile delivery in Quick-Commerce through the diligence of AI/ML technologies. The following objectives guide the research and development of solutions targeted at enhancing operational efficiency and customer satisfaction:

To evaluate the performance of existing AI/ML algorithms: This objective aims to conduct a comprehensive evaluation of various AI/ML models currently in use for last-mile delivery metrics. By assessing their effectiveness against key metrics such as delivery time, accuracy, cost, and responsiveness, this objective focuses to recognize strengths and weaknesses in existing solutions. This evaluation will help establish a baseline for arriving at how well these models perform under practical scenarios and provide perceptions into areas that require refinement.

To identify and quantify biases in current Algorithmic models: Along with performance metrics, it is vital to analyse biases that may exist within the algorithms used in last-mile delivery. This purpose seeks to disclose any unequal preference of delivery zones or demographic disparities that may affect the fairness and effectiveness of service delivery. By measuring these biases, the study focuses to specify significant concerns that need managing to ensure unbiased service across all consumer segments.

To design and validate a hybrid AI/ML algorithm: Evolving on the findings from the previous objectives, this study focuses to create a holistic hybrid algorithm that integrates multiple AI/ML frameworks. This hybrid approach will harness the integral strengths of various models to cater particularly to the dynamic needs of Quick-Commerce. The design will aim on optimizing models to handle complexities such as demand variation, delivery load variation, and urban traffic conditions. Evaluating this new model will involve stringent testing to ensure its reliability and efficacy in progressing with last-mile delivery performance.

To provide a strategic technology stack for selecting and deploying model: The final objective prioritizes the practical implementation of the study. This objective seeks to develop a strategic framework to guide business models in selecting and deploying relevant AI/ML models that align with their operational constraints and goal settings. By offering a complete tech stack, firms will be better positioned to make informed decisions regarding technology integration, hence refining their overall last-mile delivery performance.

By continuing these objectives, this study aims to deliver valuable and innovative insights for Quick-Commerce, eventually promoting more efficient and integral last-mile delivery ecosystem.

1.4 Research Questions

The following research questions are furnished to convey the gaps in integrating AI/ML technology into last-mile delivery of Quick-Commerce and to achieve the expectations of this research:

How successful are current AI/ML methods in optimizing key metrics for last-mile delivery: This question focuses on evaluating the performance of existing AI/ML techniques in improving the critical delivery metrics. What specific biases are available in the existing stacks used for last-mile delivery, and how do these biases affect service integrity and consumer trust? This question investigates to identify biases in models and their effect on fairness and customer journey. How can the integration of multiple AI/ML models in a hybrid approach optimize the performance and integrity of last-mile delivery solutions? This question explores the significant benefits of combining various AI/ML models for seamless delivery. What are the constraints and complexities faced by current AI/ML models when operating under practical conditions such as uncertain delivery loads and unpredictable urban traffic layouts? This question is about the practical resistances that existing algorithms come upon in real-world applications.

How can geo-intelligence be successfully leveraged within AI/ML techniques in last-mile delivery: This question explores how geospatial data can optimize routing decisions and enhance operational efficiency. What strategies can be developed to alleviate identified biases in AI/ML models to ensure integral delivery services across different demographic and geographic segments? This question aims to identify solutions to rectify biases and enhance fairness in delivery models. How can firms evaluate operational constraints when choosing and implementing AI/ML models for last-mile delivery, and what criteria should assist their decisions? This question focuses on planning the critical attributes that firms must consider for nominating the successful algorithm. What metrics should be categorized when assessing the performance of hybrid AI/ML models in comparison to conventional single-algorithm techniques in the context of Quick-Commerce. This question focuses on choosing the key performance indicators (KPI) for assessing multiple algorithmic techniques. What role do real-time data insights play in optimizing the decision-making capabilities for AI/ML models in last-mile delivery operations? This question explores how early access to data can optimize the effectiveness of delivery models. How consumer aspirations influence the design and enhancement of AI/ML technology stacks for last-mile delivery, and how can companies align their deliverables accordingly? This question explores the relationship between consumer preferences and the implementation of delivery models.

1.5 Significance of the Study

This study is relevant as it focuses on rationalizing last-mile delivery in Quick-Commerce by optimizing AI/ML models. It attends to operational roadblocks such as bias, lack of accuracy, and scalability, ensuring fair and successful dispatch service. By integrating smart hybrid ideas, the study aims to refine consumer journey and operational efficiency. Finally, it shares strategic insights that can transform last-mile processes and drive industry best practices.

1.5.1 Academic Contributions

This study adds to the academic field by providing a meticulous review of AI/ML algorithms specific to last-mile delivery for Quick-Commerce. It clarifies biases within current models and suggests smart hybrid approaches to optimize algorithmic efficiency and fairness.

Fostering theoretical understanding: This research addresses the gap in literature and elaborates on the theory of last-mile in Quick-Commerce by consistently examining the confluence of AI/ML stack. Also, the study emphasizes innovation by suggesting new methodologies for integrating AI/ML models, leveraging location intelligence, and enhancing operations. It focuses on transforming last-mile logistics, benchmarking new industry standards, and inspiring innovation in algorithms.

Creation of Hybrid Model: The research introduces a unique hybrid AI/ML model that synchronizes multiple techniques to improve delivery excellence. This model serves as a practical reference guide for businesses to deploy successful models based on market pull and operational challenges.

Inspiring further research: By spotting current gaps and biases in the existing AI/ML models, this study sets the stage for further academic exploration with the latest trends like Generative AI, Agentic AI, and Quantum computing. It motivates additional research into upgrading algorithms, enhancing scalability, and designing sustainable delivery solutions.

1.5.2 Practical Contributions

This research provides various critical practical contributions that can benefit business stakeholders in the Quick-Commerce sector:

Streamlined delivery performance: By enhancing AI/ML techniques, the study offers actionable insights that ensure businesses refine their last-mile delivery workflows. This approach ensures reduced delivery period, seamless order accuracy, and overall operational excellence, hence improving the consumer journey.

Bias counter measures: The identification and measurement of biases within existing algorithms have practical challenges for achieving unbiased services. The study offers strategies to control these biases, assisting firms to develop fair and ethical practices that improve consumer trust and satisfaction.

Operational blueprint for Hybrid Algorithms: The suggested hybrid AI/ML model provides a feasible methodology for firms looking to integrate various machine learning techniques. This approach helps organizations in arriving at informed decisions regarding technology adoption, aligning algorithms with specific company goals and operational challenges.

Recommendations for strategic decision-making: By offering a strategic technology stack for choosing and implementing algorithms, the study enforces firms with key criteria to evaluate and execute AI/ML solutions seamlessly. This guidance supports firms in maximizing their resource utilization and assuring scalability in their delivery operations.

Real-Time decision-making capabilities: The research underscores the importance of real-time data insights in improving operational goals. It assists firms to build real-time analytics into their decision-making processes, enabling them to react faster to variations in demand and operational constraints.

Sustainability factors: With growing priorities on sustainable practices in logistics, this study investigates how optimized AI/ML techniques can contribute to mitigating the environmental impact of last-mile delivery. Organizations can adopt greener practices by leveraging efficient routing and resource management strategies, aligning with corporate social responsibility goals.

Learning and development programs: The results of this research can apprise training programs for logistics and data science professionals. By understanding the nuances of AI/ML features in last-mile delivery, firms can cultivate a skilled manpower equipped to execute and handle sophisticated techniques.

Overall, the practical contributions of this study are designed to delegate firms in Quick-Commerce to refine their operational efficiency, address biases, and prepare for future innovations in last-mile operations. Overall, our research assists to bridge the gap between theoretical investigation and practical relevance. The holistic approach for embedding AI model into last-mile delivery of Quick-Commerce workflows delivers dynamic insights to assist academics, industry, and logistics ecosystem at large and streamlines a digital future for Quick-Commerce industry.

1.6 Structure of the Thesis

This thesis is structured into six chapters as below; It offers a holistic analysis of optimizing AI/ML technologies in last-mile delivery of Quick-Commerce:

Chapter 1: Introduction gives an overview of our research, including its origin, problem statement, objectives, research queries, and importance. It launches the foundation for addressing the gap between As-Is and To-Be AI/ML stack and summarizes the scope and purpose of the research.

Chapter 2: Literature Review probes the current literature on Quick-Commerce, AI/ML model evolution and last-mile delivery metrics. It recognizes existing trends, challenges, and gaps in optimizing AI/ML technologies into last-mile delivery of logistics value chain. It emphasizes the needs for strategic AI recommendations and facilitates the way for devising the proposed stack.

Chapter 3: Research Methodology investigates the mixed-methods approach for research and the methods of data collection, including qualitative interviews, quantitative surveys are explained along with the analysis techniques to generate results. The ethical guidelines and limitations of the study are also discussed.

Chapter 4: Results recognize the key AI/ML models from the findings and have expressed in the new model. Case studies and benchmarking results that substantiate the model and illustrate its practical relevance are considered.

Chapter 5: Discussion clarifies the context of the research objectives and questions, discusses what the results signify for businesses and academics and underscores their potential to advance last-mile delivery efficiency. The challenges and opportunities of adopting AI/ML algorithm are also included.

Chapter 6: Conclusion and Recommendations summarize the results of the study. It examines the research contributions to theoretical understanding and practical applications. It provides a set of actionable suggestions for firms pursuing the adoption of AI/ML stack in the last-mile delivery of Quick-Commerce. Finally, future research directions are described to further the effort of maximizing operational efficiency and consumer journey through AI/ML and an optimized approach to AI/ML integrated logistics methods in Quick-Commerce.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This study acts in accordance with Mishra (2024) for conducting Multivocal Literature Reviews (MLRs) in Quick-Commerce, combining grey literature (e.g., blogs, videos, and white papers) alongside traditional scholarly sources. In their study, the authors outline the interests of MLRs, including how this bridges the gap between academic research and industry best practices. Figure 2.1 highlights the literature review workflow which followed iteratively to achieve a holistic set of discoveries.

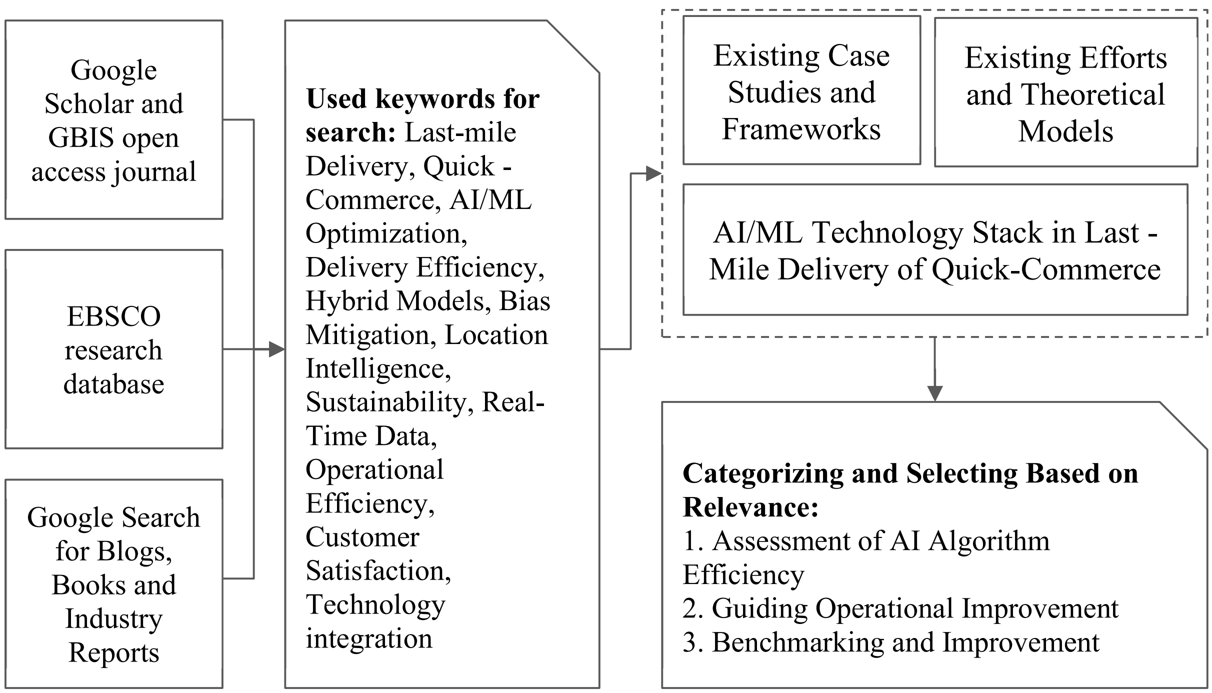

Figure 2.1: Literature review workflow and discovery (original work)

A systematic approach is chosen to analyze the impact of AI/ML technology on the Quick-Commerce (last-mile value chain). The primary data sources include scholarly databases such as Google Scholar, GBIS open access journal and EBSCO research database and broader searches encompassing blogs, books, and industry reports. For the literature review, relevant keywords are used such as last-mile Delivery, Quick-Commerce, AI/ML Optimization, Delivery Efficiency, Hybrid Models, Bias Mitigation, Location Intelligence, and Customer Satisfaction as mentioned in Figure 2.1a. This methodical approach enables the identification of relevant case studies, theoretical models, frameworks, and specific AI metrics critical to last-mile delivery optimization. Findings are categorized based on their relevance in three key areas: evaluating algorithm efficiency, assisting operational improvements, and identifying scalability potential with key benchmarks. This strategic categorization supports a balanced perspective on the role of AI/ML in improving last-mile delivery and addressing operational constraints. This systematic approach helps to review and occupy a balanced perspective, emphasize the role of AI/ML in last-mile delivery of Quick-Commerce, and accelerate the potential to address technology gaps in the digital era.

In this review, the learnings are presented in a structured and phased approach. The consultation starts with basic research that provides the need for AI and ML optimization in last-mile delivery. It then proceeds to an assessment of algorithms and their impact on key efficiency metrics. Discoveries in last-mile logistics enabled by AI and ML techniques are investigated with focus to their operational impacts. Ultimately, a gap analysis highlights areas where additional research is required. This methodology ensures strong thematic continuity and ensures a seamless explanation of both the opportunities and challenges associated with refining last-mile delivery in Quick-Commerce.

The literature review contemplates current trends and practical scenarios related to last-mile logistics in Quick-Commerce, with each study assessed in terms of its cognitive mastery, research methodology, critical observations, challenges, and future research attention. The conceptual coverage spans the main challenges of Quick-Commerce, the last leg of distribution, and the contribution of AI/ML in managing with operational biases. It additionally includes insights into last-mile business metrics, hybrid optimization models, and vehicle routing rationalization techniques configured to fast delivery ecosystems. The examination also includes studies on predictive fulfillment in both Quick-Commerce and broader e-commerce frameworks, automation driven warehouse optimization, and AI enabled routing strategies. Moreover, it investigates the use of genetic algorithms for rider efficiency, real-time AI/ML analytics for supply chain visibility, solutions handling persistent last-mile challenges, and emerging features of agentic AI to enhance delivery efficacy.

2.2 Challenges of Quick-Commerce

Quick-Commerce emerges as a disruptive force in retail, providing speedy delivery of consumer items within 10-15 minutes of online purchase (Gupta, 2024). This unique phase of retail digital transformation depends on a visionary logistics firms centered around urban "dark stores" and last-mile delivery by cyclists or motorists (Dhiman et al., 2024). While Quick-Commerce provides convenience, it encounters significant challenges. These comprise legal hurdles, with potential legislation in countries that could restrict its growth (Dhiman et al., 2024). Environmental concerns are also paramount, as frequent deliveries result in increased traffic congestion and greenhouse gas emissions (Son & Kwon, 2023). Furthermore, Quick-Commerce leads to exorbitant waste due to the requirement for freshness and safe delivery (Son & Kwon, 2023). To handle these challenges, governments need to draft effective implementation framework for Quick-Commerce policies, concentrating on analysis, solution design, and identification of benefits (Younus et al., 2023).

The findings from Gupta (2024) examine the complex aspects of Quick-Commerce in the burgeoning market of India. The study strives to deliver an elaborate overview of the

Quick-Commerce ecosystem, its development, progress, key stakeholders, and the challenges encountered by retailers, grocers, and e-commerce platforms. These findings strive to collate information on the issues faced by grocers and the growth of Quick-Commerce in India. Gupta (2024) confirms that the emergence of Quick-Commerce impacts the traditional retail economy in India, driven by the speedy development of application-based platforms that provide expedited grocery delivery. These organizations target 10-minute fulfillment from strategically located dark stores and depend heavily on technology to achieve operational excellence. Nevertheless, they encounter risks in consistently meeting consumer expectations for on-time delivery and experience. This literature also explains that FMCG organizations are exponentially investing in Quick-Commerce channels, predicting that a significant volume of their future e-commerce revenues will happen from these platforms as consumer behavior shifts towards immediate, convenience driven buying patterns. However, Gupta (2024) admits key limitations in the study. The dependence on secondary data is sourced from websites may lead to inaccuracies or bias, and the overall scope of the topic combined with time constraints prevents in-depth exploration in few areas of literature.

2.3 Last-Mile Delivery in Quick-Commerce

Last-leg distribution in digital commerce gains considerable attention due to its impact on last-mile delivery costs and consumer happiness. Recent study focuses on smart solutions to improve last-mile efficiency, which includes AI/ML powered techniques by leveraging robots, drones, and autonomous vehicles (Sorooshian et al., 2022). A suggested system integrates social media analytics and deep learning techniques for traffic forecasting in urban areas (Laynes-Fiascunari et al., 2023). Primary elements impacting last-mile costs include failed deliveries, customer density, and workflow automation (Mangiaracina et al., 2019). Practical solutions to examine these constraints comprise parcel lockers, crowdsourcing logistics, and policies for dynamic pricing (Mangiaracina et al., 2019). The department of last-mile research grows rapidly, with most publications happening in the last five years (Olsson et al., 2019). A proposed model for last-mile literature includes five interrelated elements: last-mile logistics, distribution, fulfillment, transport, and delivery (Olsson et al., 2019).

The paper, "Impact and challenges of intelligent technology on modern last-mile delivery" by Sooroshian (2024) examines the potential impacts of AI/ML powered techniques, on refining last-mile delivery, also problems that impedes the benefits of modern last-mile delivery. The approach used in this literature is a narrative literature review. A holistic finding of the Scopus database is handled by the authors for literature on last-mile delivery and AI/ML techniques, focusing on the past 5 years. Due to the absence of a significant portion of prior literature, the authors chose to extract the existing research through a narrative review instead of a systematic one.

While investigating technological progress that shape last-mile delivery, Sooroshian (2024) focuses on the transformative role of AI in optimizing efficiency and controlling operational cost across delivery ecosystems. The literature highlights that AI-flavored systems, comprising route optimization techniques and autonomous delivery tools like drones, self-driving vehicles, and robots, can effectively rationalize work-force requirements which improve speed and productivity. These skills strengthen AI/ML as a transformative force for the future of quick fulfillment frameworks adopted in Quick-Commerce. Despite these opportunities, Sooroshian (2024) contemplates that the full value of AI-enabled last-mile delivery is yet to be realized due to significant challenges in execution. Rules-related constraints involving ethical considerations, transparency, accountability, and technology integrity remain highly noticeable. The study further examines that cross-border inconsistency in legal and compliance structures leads to inefficiencies in data processing and restricts the AI scalability. Adoption is also slowed by firms' resistance to technological change and a dearth of in-house AI/ML subject matter expertise. Additional barriers comprise security clauses and increase in costs associated with autonomous systems, along with less robust infrastructure and facility readiness. Hence, operational constraints and legal fragmentation continue to restrict the transformation of AI/ML based delivery ecosystems.

The author prioritizes that last-mile delivery (LMD) is crucial to supply chains, and e-commerce platform providers must advance with industry trends. Logistics workflows for LMD are challenging, inefficient, and expensive, and many firms are aiming to achieve consumer confidence by delivering faster, more efficient, and accurate delivery. Emerging technologies are introduced and leveraged to satisfy client requirements while also gaining a competitive advantage. As per the references in this article, AI shapes LMD's potential and challenges. Consequently, the aim of this literature evaluation is to assess, identify, and optimize the finest LMD model.

This indicates that AI/ML driven technologies improve last-mile logistics and firms' operations by improving efficiency and reducing spending. Companies outperform their competitors whose growth strategies do not take advantage of AI based emerging technologies. While insightful AI can assist optimizing delivery routes, it is not always flexible to adjust with the dynamic and volatile last-mile pattern. With the use of AI/ML, LMD can improve delivery routing, identify patterns in past data, and create algorithms using real-time data. Drones, robots, and autonomous vehicles employ AI to maximize performance and reduce labor spending. Amazon and DHL are orchestrating the industry's adoption of AI driven technologies, such as autonomous vehicles, robots, and route planning applications, even though a matured AI based model for last-mile logistics is still in primary stages.

Moreover, big data science and machine learning are considered as data-driven decision-making technologies within the AI domain, capable of improving business decisions and facilitating effective predictions and similar outcomes. AI's emerging techniques are creating new markets for Go-To market strategy, prompting numerous LMDs to embrace innovation and initiate intelligent future. Despite this, latest LMDs, revolving on AI-driven techniques, are feasible and beneficial. However, current constraints impede their advancement. Overall, the deployment of AI models offers substantial benefits and opportunities for modern LMD.

2.4 Biases in AI/ML for Last-Mile Delivery

Recent research focuses on various biases in AI/ML models used for last-mile delivery and Quick-Commerce. Data skew is a primary concern, with imbalanced datasets ending up in unfair AI models (Kelkar et al., 2021). The last-mile gap in AI execution comprises roadblocks in data governance and the requirement for "data awareness" practices (Coiera, 2024). Machine learning features are utilized to last-mile delivery optimization along with traditional techniques (Giuffrida et al., 2022). To better comprehend and reduce bias, scholars create frameworks for producing dummy datasets with various forms of biases, allowing for experiments on various scenarios and their impacts on performance and fairness assessments (Baumann et al., 2023). These findings underscore the importance of evaluating bias throughout the entire modelling pipeline to avoid negative results in AI/ML models.

Kelkar (2024) aims on reducing bias in machine learning features by suggesting a structured model designed to detect, reduce, and finally avoid bias throughout the development lifecycle of the model. The author shares an exhaustive tutorial that explains the various modes of bias that can impact decision-making systems and summarizes practical ways to inspect and mitigate their outcome. The findings involve an automated algorithm capable of bias correction of native data sets through techniques such as statistical parity, and varied impact analysis, ensuring fair representation in model outputs. To explain class imbalance issues, Kelkar applies the SMOTE-Tomek resampling method, which upgrades the quality of training data and optimizes the robustness of predictive outcomes. A significant contribution of this study is the introduction of a decentralized data architecture that focuses on web scraping and homomorphic encryption, which allows safe access to real world data models while maintaining privacy and integrity. This study reinstates the requirement for fairness centred data engineering approaches in AI powered last-mile delivery methods. Businesses leverage machine learning techniques for crucial decision-making workflows. Despite this, inequity inherent in the firm's model may lead to critical issues such as erratic hiring decisions and adverse effects on the company's reputation. Companies frequently unfamiliar with the expertise or resources to execute a method that aligns with social norms, culminating in the implementation of error prone models. The author proposes a framework to help enterprises in controlling bias in their models. This approach starts with a holistic user guide to clarify critical terms and provide recommended practices for developers to evaluate and mitigate bias. Kelkar (2021) experiments the discussion on bias mitigation in AI by leveraging an automated approach that explains fairness issues throughout the data and method pipeline. The research applies statistical uniformity and disparate impact features to find and correct bias within raw datasets, while the SMOTE-Tomek resampling method is provided to handle class imbalance and optimize the inclusivity of minority groups in model training. To further alleviate discriminatory model behaviour, the author inherits a "Reject Option

Classification” algorithm, which reduces biased outcomes during estimation. Kelkar reiterates the significance of exploring datasets with incremental records representing underprivileged and minority populations, asserting that such inclusion improves sampling accuracy and assist the implementation of fairness in AI methods.

Regardless of its contributions, the research admits various bottlenecks. Implementing Generative Adversarial Networks (GANs) confirms challenge due to their intricacy, although this delivers a promising aspect for future research. The author also highlights that while the SMOTE+ENN technique may outperform SMOTE-Tomek in handling imbalanced data collections, it is kept out of this work due to its aggressive down sampling behavior and the additional effort required for tuning. In addition, the practical relevance of the suggested approach remains to be assessed at scale within industry ecosystems.

The author proposes future work aimed at improving fairness techniques using GAN based approaches, optimizing the SMOTE+ENN resampling method for better performance in bias reduction, and orchestrating large scale deployments to evaluate the practical feasibility of the recommended model in operation AI systems. In this research, Kelkar (2021) articulates the structure and benefits of the suggested algorithmic model designed to reduce bias in AI/ML systems such as a comprehensive user guide, an automated bias correction approach, SMOTE-Tomek, and a Reject Option Classification algorithm. Despite this, the prospect of AI transforming nearly all industries, blended with the crucial risk of prejudice against certain demographic groups, delivers an irresistible urge for seamless progression.

2.5 Evaluate Last-Mile Delivery Metrics Using AI/ML Algorithms

Recent study emphasizes the increasing relevance of optimizing last-mile delivery (LMD) operations in e-commerce and logistics. Various criteria are deployed to evaluate LMD success, comprising delivery time, cost, and client happiness (Verbytskyi, 2023). AI/ML models are increasingly applied to improve these business metrics. Techniques such as reinforcement learning, metaheuristics, and predictive analytics are deployed to enhance delivery routes, forecast needs, and improve efficiency (Verbytskyi, 2023; Chandramouli, 2023). The inclusion of sophisticated technology such as route optimization models and real-time tracking systems can noticeably reduce delivery times and expenses while enhancing consumer happiness (Chandramouli, 2023). Furthermore, sustainable practices are being explored to reduce environmental impacts and potentially lower costs (Chandramouli, 2023). However, the execution of AI-powered technologies in LMD faces challenges that need to be handled (Sorooshian et al., 2022). Overall, these studies highlight the potential of AI/ML in transforming LMD operations (Giuffrida et al., 2022). Verbytskyi (2024) examines machine learning (ML) models for delivery route optimization to improve metrics like delivery time, cost, and consumer happiness. Refinement of delivery routes consume more time and need a lot of work on the ground. To overcome these problems, they are using ML models quite often for better and faster optimization. This draft addresses the latest trends to use ML models to improve delivery routes and the major problems that delivery companies must manage. It highlights metaheuristic methods, reinforcement learning, and machine learning, as well as their merits and demerits. When designing a delivery route refinement system, attributes like the volume of trucks, their capacity, delivery windows, road networks, and consumer demand are considered. There is various optimization goals revealed, such as reducing delivery time, lowering transportation costs, and increasing consumer satisfaction. Lastly, the research explains about areas for future research, such as multi-agent systems, swarm intelligence, and hybrid algorithms. This study reviews everything which are needed to understand about using ML algorithms to improve delivery routes. It can assist personas who work in or study the extensive logistics business.

The primary research gaps and challenges mentioned in this study are a lack of findings that improve last-mile delivery routes while considering several objectives such as cost, time window limits, and environmental impact. ML techniques are infrequently used in practical applications due to scattered data or the challenges in receiving correct triggers from existing models. Insufficient attention is paid to sustainability issues such as energy usage throughout the logistics process, which can have a significant impact on overall efficiency and performance metrics related with last-mile delivery optimization workflows.

Machine Learning (ML) is the most efficient and state-of-the-art method for optimizing delivery routes. ML methods gain momentum for improving delivery channels due to their ability to overcome the challenges of conservative methods such as real-time traffic monitoring. ML algorithms can improve performance by leveraging historical travel data and deploying predictive models that recognize trends withing the data cluster. These models can eventually suggest more efficient routes for future deliveries, considering factors such as traffic conditions, and weather forecasts. ML highlights the introduction of models and algorithms that provide predictions without explicit coding. It engages procedures such as supervise learning, which utilizes labelled training datasets to design model, and unsupervised learning, which leverages unlabelled datasets. ML algorithms can improve distribution routes by forecasting accidents and assessing the interaction of high-frequency risk attributes across the operation. ML methods for refining distribution routes generally entail data, including customer locations, delivery times, and distances among consumers. They may also need information regarding the state of the items being distributed (e.g., weight or dimension), traffic conditions in specific locations, and road limits due to construction activities, and more. This data is consumed by ML models to identify ideal distribution routes that mitigate costs while improving efficiency and consumer happiness. Numerous state-of-the-art algorithms exist that can facilitate transport optimization. Algorithms for ML include Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Genetic Algorithms (GA).

Integrated models combining conventional ML techniques are proposed for versions of vehicle routing problems, including simulated case studies. They function by combining the advantages of various models to achieve optimal results than any individual model could provide independently. One approach may integrate genetic algorithm with particle swarm optimization and large neighbourhood search (LNS). This will enable faster convergence to an optimal solution while also assuring resilience. Hence, future study will focus on multi-agent systems, swarm intelligence, and hybrid models.

Giuffrida (2022) investigates optimization and ML methods in last-mile delivery, even though does not explicitly evaluate operational metrics or compare AI/ML methods. The intention of this study is to submit a holistic overview of the latest trends in literature concerning last-mile delivery strategies, serving as a baseline for researchers aiming to trouble shoot unique methodologies in the field of last-mile transport optimization. A bibliometric study and critical review are executed to clarify the key issues examined, the methods employed, and the case studies submitted. The assessment results in enabling the categorization of studies into established optimization models, ML methodologies, and hybrid models. The key research gaps and limitations in the existing literature are assessed to highlight unresolved challenges and provide suggestions for future investigations.

Giuffrida (2022) underscores a thriving scholarly interest in optimizing last-mile delivery, specifying a significant improvement in related research in recent years. The findings observe that supervised ML techniques are frequently being applied to solve operational constraints such as demand forecasting and anomaly detection in distribution operations. Moreover, the review emphasizes that conservative operational research methods continue to control much literature, particularly in the context of vehicle routing channels, return logistics, demand planning, and the management of multi-model distribution fleets. These studies collectively propose a significant shift toward data-driven models in the improvement of last-mile transportation. Nonetheless, the findings confirm certain hurdles that prevent the breadth of its penetration. It does not claim to be a comprehensive review of all urban logistics research but rather provides a baseline representation of the current state of the ground. The scope is limited to available publications and indexed up to June 2021, leaving out the latest trends and best practices. In addition, the author specifies that the review provides limited focus to the role of AI and ML applications, an area that progresses and requires deeper analysis in future research. According to Giuffrida (2022), future research should aim at transforming last-mile delivery models to be effective, automating logistics operations with the help of AI/ML, and coming up with new unsupervised learning techniques.

2.6 Hybrid Model Optimization Techniques

The latest research probes into visionary approaches to upgrading last-mile delivery operations. ML methods such as deep reinforcement learning and constrained clustering models are suggested to refine truck-drone routing mechanisms and delivery location clustering models (Arishi et al., 2022). Banyai et al., 2018) identifies that smart scheduling using black hole optimization and Industry 4.0 systems could enhance first-mile and last-mile delivery efficiency. Customized cost matrices and a blend of accurate and estimate algorithms are designed to handle route sequencing constraints, substantiating adaptability across different demographic contexts (Hernandez et al., 2024). Multi-criteria optimization models encompassing automated smart lockers, capillary distribution networks, and crowd shipping are deployed to refine last-mile delivery operations, considering factors such as cost efficiency, performance optimization, and resource utilization (Sawik, 2024). These varied approaches provide potential for creating hybrid models that could improve last-mile delivery parameters.

The author, Arishi (2022) suggests an integrated ML model blending with constrained clustering and deep reinforcement learning which can effectively solve the last-mile delivery issue using truck-drone methods. His research methodology included a two-stage integrated ML model, a constrained clustering method in the early stage to classify delivery location as per user-defined constraints, followed by a deep reinforcement learning model in the upcoming stage to optimize routing with the clusters. The study provides a hybrid ML method that includes a constrained clustering algorithm with a deep reinforcement learning framework to refine last-mile delivery using truck-mounted drones. The constrained clustering component classifies delivery locations based on predefined operational conditions, while the deep reinforcement learning model determines the fine-tuned route across these clusters. The findings confirm that this hybrid approach can generate high-quality routing designs in real-time, reinforcing its potential usage in dynamic and volatile distribution ecosystems.

As e-commerce progresses and consumer demand spikes, supply chains are striving to achieve efficient delivery operations. The last-mile delivery signifies one of the most powerful distribution challenges, specifically mentioning the final segment of the logistics process in which a parcel is transported from a warehouse to the end consumer. Despite the significant spike in online orders and direct consumer deliveries, reliance on truck-based delivery does not change for decades. The inclusion of drones in the last-mile delivery can explain various operational issues and yield competitive benefits. Traditional algorithms for truck-drone systems provide continuous calculations and are limited to solving only minor problems. This study focuses to distinguish an effective solution for the last-mile delivery challenges. Arishi (2022) introduced an integrated ML strategy to explain the challenge of distribution channels to different locations leveraging the latest vehicle embedded with many drones. A constrained clustering method is introduced to cluster delivery locations according to user-defined challenges in the beginning phase. The centre of each limited cluster functions as a launching site for truck parking and drone dispatch for final deliveries. A deep reinforcement learning method is introduced in the second phase to establish an optimal route among all clusters. Meticulous experiments are performed to analyse the proposed methodology. The solutions derived are blended with conventional model. The studies indicate that this approach can provide high-quality solution instantaneously. Banyai (2018) suggests an integrated model and a black hole optimization technique to refine first mile and last-mile (FMLM) distribution operations. Supply chain management increasingly includes Industry 4.0 technology to optimize availability, flexibility, sustainability, and efficiency. In integrated logistics networks, operations are handled from suppliers through third-party logistics providers to multiple clients. Different delivery approaches exist based on time and cost considerations. In recent years, a diverse set of clients expresses acceptance to incur an additional cost for standard or expedited delivery. This practice provided the importance of the efficient design and management of FMLM distribution solutions. Service ideas based on cyber-physical systems augment the enhancement of FMLM delivery productivity within the big data ecosystem. The design and operational issues can be characterized as hard optimization challenges. These issues can be explained with latest models and techniques based in heuristic and metaheuristic algorithms. This research provides an exhaustive supply model for FMLM distribution. This study presents a mathematical model for the real-time intelligent scheduling for FMLM delivery following a comprehensive literature analysis. The hybrid model includes the allocation of first-mile, and last-mile delivery works to available resources and the optimization of operational costs, while accounting for restrictions such as capacity, time window, and availability. A methodology based on black hole optimization (BHO) explaining a multi-objective supply chain model is gradually introduced. The enhanced algorithm's sensitivity is validated using benchmark modules. Quantitative outputs across various datasets highlights the importance of the proposed model and recommend the application of Industry 4.0 discoveries in FMLM distribution network.

Banyai (2018) introduced a mathematical model designed to assist the real to near to real-time scheduling of FMLM deliveries. The method explains both the allocation of distribution activities to existing resources and the rationalization of operational expenses, while encompassing key challenges such as capacity, delivery timelines, and resource availability. To solve this multi-objective supply chain problem, a black hole optimization (BHO) algorithm is applied, and its sensitivity is evaluated using benchmark functions. The finding illustrates that the hybrid approach refines the efficiency of FMLM task allocation and cost management, and the quantitative evaluation ensures the feasibility of Industry 4.0 frameworks in advancing real-time logistics excellence within the FMLM ecosystem. The key challenges include the complexity of the optimization problems explained in the paper, which may involve refined algorithms for resolution. The proposed BHO algorithm is assessed on benchmark functions rather than practical dataset, and the abstract lacks particulars regarding the datasets of quantitative outcomes deployed to demonstrate the strength of the proposed algorithm.

2.7 AI/ML Algorithms for Vehicle Routing Optimization (VRO)

Machine Learning (ML) algorithms are rapidly applied to enhance delivery truck routing in Quick-Commerce. These solutions offer considerable improvements compared to traditional ways, handling issues such as reducing delivery time, decreasing expenses, and enhancing customer satisfaction (Verbytskyi, 2023). ML approaches display encouraging results, with a study reporting up to 25% distance savings and 14% decrease in total delivery time (Boumpa et al., 2022). To enhance algorithmic performance and problem formulations, researchers are examining various ML models, such as hybrid algorithms, reinforcement learning, and metaheuristics (Bai et al., 2021). Provided the importance of last-mile logistics in today's e-commerce landscape, it is highly advantageous to combine ML with analytics and IoT technologies (Giuffrida et al., 2022). Vehicle routing optimization in Quick-Commerce scenarios can be further optimized with the use of multi-agent systems, swarm intelligence, and more sophisticated hybrid methods.

According to the author, Bai (2021), ML can enhance the modelling and optimization techniques of vehicle routing issues. The Vehicle Routing Problem (VRP) is a highly explored integrative optimization issue, for which different models and algorithms are deployed. To handle the complexities, uncertainties, and dynamics built-in real-world VRP applications, ML techniques are blended with analytical models to upgrade problem formulations and algorithmic efficiency across different problem-solving contexts. However, the relevant articles are separated across numerous standard research domains distinguished by diverse and occasionally confusing terminology. This research provides a complete evaluation of hybrid methods that integrate analytical models with ML methodologies to tackle vehicle routing challenges. This study examines the evolving research areas regarding ML assisted VRP modelling and optimization. Bai (2012) confirmed that machine learning may improve vehicle routing problem modelling and enhance model performance for both online and offline VRP optimizations.

The research methodology of this research aims on three major approaches to incorporate machine learning with VRP optimization. Primarily, machine learning assisted VRP modelling is deployed to enhance the estimation of probability distributions for uncertain attributes, improving the predictive accuracy of routing algorithms. Next, ML guided offline optimization harnesses machine learning to assist the solution process for offline VRP scenarios, applying techniques such as adaptive neighbourhood search and trainable constructive methods. Ultimately, ML assisted online refinement applies machine learning to enhance the performance of models in real-time dynamic VRP ecosystems, ensuring responsive and systematic decision-making in last-mile delivery operations. As per the author, despite the increasing visibility of inculcating machine learning in VRP research, the literature community continues to encounter several problems. Primarily, the VRP is typically integrated within a more exhaustive complicated system that demands a high degree of reliability across all probable scenarios to avoid the system from getting into detrimental states. In addition, a high degree of understandability of the designated algorithms is imperative. Regardless of existing research initiatives in explainable AI and verification techniques, further theoretical investigation is important to establish a robust base for future studies and extensive applications.

Finally, there is not a successful VRP research platform with ML support for testing and training. The relevant libraries and knowledge repositories are widely spread within the two distinct study groups of operations research and machine learning, respectively.

The two research communities must engage in designing an innovative platform with ML and optimization libraries and tools to assist the expansion of ML assisted VRP research. Normally, there would be opportunities for interdisciplinary study as well. Along with that, much of the latest ML guided VRP research requires considerable quantities of data that are typically not readily available. The trained methodologies provide inconsistent abstraction across numerous instances, scenarios, and domains. The lack in data and model abstract is mainly derivable to the trial-and-error model inherent in conventional deep learning. The key information relevant to the objectives and limitations is not completely leveraged. Innovative strategies are crucial to improve the integration of current analytical models grounded in mathematical frameworks with data driven neural network systems. Ultimately, the leverage of additional real-world features is vital to overcome the challenges of current VRP studies. It is imperative to harmonize the adoption cost with the efficiency of the solution models. A more robust VRP simulation program is vital for practitioners to curate customized VRP ecosystems that provide high-quality performance at a minimal cost. Bai (2021) highlights the necessity for the research community to provide trained VRP base models and libraries to further restrict training cost.

2.8 Predictive Fulfilment in Quick-Commerce

Recent study envisions machine learning features in Quick-Commerce and E-Commerce value chains. Akulov et al. (2023) suggests using a modified K-means model to enhance distribution centre locations, significantly decreasing transportation costs and improving customer experience. In Quick-Commerce, Dhiman et al., (2024) examines the relevance of “dark stores” for urban deliveries, underscoring the important role of logistics in achieving tight delivery time windows. For example, Al Rahib et al. (2024) identifies that support vector machines are accurate in forecasting consumer turnover and on-time delivery, proving the importance of ML models in this area. Droomer & Bekker (2020) examines the feasibility of decreasing paper-based marketing and personalized promotions by leveraging ML to forecast when the targeted consumers will do repeated purchases. In the dynamic e-commerce, these research shows how ML can refine targeted marketing, logistics, and consumer behaviour forecast.

Al Rahib (2024) clarified in his literature that ML models may skilfully forecast consumer attrition and timely product delivery on last-mile. Consumer attrition exhibits a significant challenge for e-commerce firms, resulting in revenue decline and fading consumer loyalty. In latest years, ML approach surface as powerful tools for predicting and reducing consumer attrition. This research is to examine the application of ML models for forecasting consumer attrition, annual cost, and timely product delivery in e-commerce. The study initially conducted an experimental assessment of the literature regarding consumer attrition in ML. Study highlights that customer turnover is effectively forecasted leveraging several ML models, such as Support Vector Machines (SVM), random forest, and decision tress across multiple industries. This study executed an experimental investigation of customer churn in e-commerce leveraging ML models to fill the current gaps in the literature. The data are eventually refined and analysed using ML models for predictive reasonings. The study's insights convey that ML models are competent in forecasting customer attrition and promising timely goods delivery in e-commerce. The most robust model, SVM, achieved an accuracy of 83.45% in forecasting customer attrition and 68.42% in anticipating goods on-time delivery performance. Al Rahib (2024) deploys a structured ML approach to examine various aspects of consumer pattern and delivery performance. The research leverages three datasets, considering customer turnover, annual consumer cost, and on-time delivery, which are refined to manage missing values and transform explicit features into numerical illustrations. The data are then split into training and testing subsets to accommodate model design and validation. For predictive modelling, decision trees, SVM, random forests, and XGBoost technique are harnessed to forecast customer turnover and on-time delivery, while linear regression is utilized to predict annual consumer cost. Model performance is evaluated using standard classification metrics, including accuracy, precision recall, and F1 score. Moreover, the research executes a detailed assessment of predictions on the test datasets, examining outputs across various attributes to measure the resilience and practical feasibility of the models. Some of the important observations are ML techniques, particularly Support Vector Machines (SVM), are promising in forecasting customer attrition, annual expenditure, and on-time delivery. Buying frequency, product categories, and average session duration are all key features of customer attrition, annual cost, and on-time delivery performance. The study provides insights into customer behaviour in the e-commerce sector and recommends ways to improve customer experience and increase in revenue.

The author identifies that leveraging ML methods to predict e-commerce attributes like customer annual spending, attrition, and on-time delivery can assist firms to improve customer retention and boost their revenue. This approach signifies that we can leverage regression and classification models to assess customer data efficiently. The study showed that attributes like how often customers purchase, the nature of items, and how long they spend on the website are primary indicators of customer attrition, annual spending, and timely product delivery.

2.9 Predictive Fulfilment in E-Commerce Using AI

ML models exhibit promising outcomes in predicting digital shopping behaviour and enhancing c-commerce operations. Random forest and Stochastic Gradient Descent models are effective in predicting user buy intentions, with Random Forest claiming an F1-score of 0.90 (Kumar et al., 2023). The Support Vector Machine (SVM) proves considerable accuracy in forecasting customer attrition (83.45%) and timely delivery (68.42%) in the e-commerce sector (Rahib et al., 2024). Predictive models guide firms in interpreting items and retail attributes that affect sales. Through the assessment of extensive information, ML algorithms can observe patterns and correlations in consumer behaviour, facilitating data-driven decision-making (Kumar et al., 2023). The application of these models can provide substantial ideas for organizations to enhance consumer journey, reduce turn over, and augment revenue. ML models provide powerful templates for predicting and analysing numerous aspects of digital shopping behaviour.

Pandian (2019) clarified in his article the development of a predictive model leveraging ML to anticipate item sales at Big Mart digital marketplace. Everyone explores methods to buy items at a competent price or to promote them economically. The project "Predictive analysis for big market sales using ML models" aims to develop a predictive model to ascertain the sales of individual items at a specific retail outlet. Big Mart will leverage this technique to analyse the attributes of items and stores that substantially contribute to incremental sales. Author experimented few methods like Linear regression analysis, Random Forest Regression, Polynomial regression, and Date preprocessing, by managing the missed values to arrive at the ML model. With the least Root Mean Square Error (RMSE) when compared to other models like linear regression and polynomial regression, the Random Forest methodology outwitted the others in terms of sales prediction. When compared to linear and polynomial regression techniques, the Random

Forest and Ridge regression models exhibited superior prediction performance in terms of accuracy, Mean Absolute Error (MAE), and RMSE. Kumar (2023) examines the application of ML techniques, particularly Stochastic Gradient Descent and Random Forest, to forecast digital user intent for item buys in his literature.

ML in e-commerce, distinguished by the execution of self-learning models that autonomously enhance their performance through experience, is a paramount trend in retail digital transformation. ML models can be efficiently trained by evaluating large data collections, identifying repeating patterns, correlations, and anomalies within the data, and designing mathematical models that resonate these patterns. These methods are optimized when algorithms transform progressively into larger body of data, yielding decent insights on certain ecommerce related events and the linkages among the basic factors. An instrument that is highly recommended in examining latest issues, forecasting future events, and accommodating data-driven decisions. This research evaluates the application of ML models to forecast user intent to buy an item on a particular store's website. This experiment leveraged a data collection on online shoppers' buying intentions from the ML repository. This study included classification-based ML models, such as Stochastic Gradient Descent (SGD) and Random Forest. The SGD model is leveraged for the first time to predict online user intent. The outcomes highlight that the Random Forest had the greatest F1-Score of 0.90, exceeding the stochastic Gradient Descent approach.

Sarkar (2023) proposes a machine learning approach to enhance e-commerce pricing and forecast digital transactions to ease Quick-Commerce. The study presented a resilient approach for forecasting online consumer buy through dynamic pricing, leveraging a blend of statistical and ML methods. The methodology focused on improving pricing models on e-commerce platforms, prioritizing the choosing of the most relevant purchase price over purely the least option. The inclusion of K-means clustering permits customer segmentation, considering unique consumer segments with buy patterns, while the dynamic pricing concept, reinforced by regression equations for each client cluster, efficiently decides consumer choices, can optimize the framework further.

2.10 AI driven robotics for warehouse optimization

Warehouse and distribution operations are being transformed by hyper automation and artificial intelligence (AI), particularly in the scope of Quick-Commerce. As per, Borboni et al. (2023), due to collaborative robots (cobots), industrial workflows are experiencing an increase in safety and productivity. According to Pandian (2019), smart warehouse infrastructure automates storage and retrieval by leveraging AI, the Internet of Things (IoT), and cloud computing. This encourages for real-time inventory management. There are many benefits of including robotics and automation into supply chains, including lower prices, better efficiency, and safety (Banur et al., 2024). As per Banur et al. (2024), the use of AI/ML is on the rise in supply chain management, and it has the advantage to optimize decision-making. Robotics and AI are transforming logistics automation, which is enhancing market position and reducing errors (Ferreira & Reis, 2023). A potential industry transformative development could be social robots that can accompany placidly with people, made possible by the prospects presented by human-robot integration in logistics (Ferreira & Reis, 2023).

In his journal, Borboni (2023) explained how AI/ML are feasible for collaborative robots to play a significant role in Industrial infrastructure. A collaborative robot, or cobot, let people interact closely with it by conversing with it directly, so they don't want to utilize screens. When cobots engage inside of warehouse walls, they narrow the gap that exists between industrial robots and workforce. Cobots can handle various tasks, including communication robots in public places and logistics robots that transfer items inside the distribution centre. Industrial robots that assist automating tasks that are not health-centric, like helping workforce handle heavy items or working on assembly lines. Applications for humanoid robot that are designed with dependability and safety ensures workforce trusting of delivering tasks. These features also enhance employee performance and working terms. Machine learning is transforming cobots from science fiction into practical life scenarios. They can adapt to new scenarios, reduce cost, and enrich consumer experience. In this paper from Borboni (2023), the latest research articles are systematically validated to identify what part of AI already plays in cobots for industrial relevant and how it might prosper in the future.

A distinct and well-defined scheme for systematic reviews supports the understanding and evaluation of the employed approached. This research employed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach. The PRISMA model underscores the evolution of information through numerous stages in a systematic literature review, including the total count of studies that are identified, excluded, and included, along with the reason. The databases like Web of Science, IEEEXplore, PubMed, ScienceDirect, SpringerLink, Scopus, and Research Gate are examined for the literature search on collaborative robots leveraging the keywords like Human-Robot Interaction (HRI), cobots, AI in robots, collaborative learning, reinforcement learning, deep learning in robotics, and industrial robots. This structured review covered peer-reviewed academic journals, conference papers, and reviews published that include qualitative, quantitative, or both types. Borboni (2023) submits an assessment of how AI and robotics are transforming industrial operations by narrowing the gap between human capabilities and technological solutions, empowering firms to explain operational challenges while adjusting to dynamic customer expectations. The study evaluates that the integration of AI considerably optimizes cobot performance, with machine learning and deep learning techniques improving capabilities in functions such as visual perception, motion control planning, and adaptive learning, all while managing reliability and operational safety. Despite these improvements, Borboni (2023) highlights limitations and areas for future research. Current deep learning models are time consuming and be short of the flexibility required for cobots operating in complex, real-time situations. In addition, existing models do not assist effective knowledge sharing among networked cobots, which could expedite learning and enhance operational excellence. The application of reinforcement learning is challenged in continuous engagement with real-world production ecosystems, as well as high computational costs. These challenges explain the need for continued study to optimize the adaptability, scalability, and practical application of AI-driven engagement robots.

The futuristic study focuses on improving stable and safe ecosystems for cobots to synchronize with humans, following quality standards and requirements. It explains advanced algorithms for cobots to operate in low-volume scenarios, including logistics planning and collision prevention frameworks that can manage volatile environments. This study improves deep learning models for cobots to be effective and adaptive in complex settings. This assessment by Borbani (2023) scrutinized the results of the specific journals, including accuracy, safety concerns, time delays, and robotic capabilities. This structured study finally provided suggestions and future scope for the alliance between collaborative robots and ever-evolving domain of AI. Woschank (2020) also examines future directions for AI, ML, and Deep learning in last-mile logistics. His study identifies numerous critical areas where the use of AI, ML, and Deep Learning (DL) models in smart logistics showed assurance, including predictive maintenance, decision support systems, refined planning and scheduling, and logistics operations. Woschank (2020) executed a conceptual framework to assist future research and guide experts in executing AI, ML, and DL methodologies in their logistics operations.

2.11 Auto-Assignment of Last-Mile Delivery Using Machine Learning

Recent study investigates machine learning and optimization models for enhancing last-mile delivery efficacy in e-commerce. Research examines different models, particularly smart distribution task using regression models and the Hungarian algorithm to enhance assignment allocation (Vummadi et al., 2024), and delivery location/zone tagging approaches for work-force allocation and route optimization (Younus et al., 2023). Machine learning and traditional optimization models are leveraged to handle difficulties in last-mile value chain, insisting low cost, delivery time enhancement, and consumer delight (Giuffrida et al., 2022). Study on sustainable operations in last-mile function explores the benefit of ML predictions and real-time data to influence consumer decisions and optimize distribution channels (Orsic et al., 2022). These studies highlight the feasibility of integrating ML models with optimization systems to address the nuances of last-mile delivery, offering assured options for refining efficiency and sustainability in the accelerating e-commerce segment.

Orsic (2023) explained in his research that machine learning may enhance last-mile chain for sustainable delivery operations, including dynamic route planning and the configuration of end customer's delivery window. Last-mile chain is considered one of the less efficient, expensive, and environmentally detrimental segments of the supply chain, particularly affecting sustainable delivery excellence. The dominant e-commerce business model, known as Attended Home Delivery (AHD), is the expensive and resource-intensive, specifically when a narrow delivery slot is aligned with the consumer, hence squeezing potential flexibility options. In contrast, last-mile logistics is progressing as decisions must be considered immediately. This article explains the suggested solution for sustainability opportunities in “Attended Home Delivery”, leveraging a holistic approach to achieve sustainable deliveries through ML forecasts derived from real-time data, numerous dynamic route planning factors, tracing logistics events, transport capacities, and other important data attributes. The suggested strategy focuses to motivate consumers to choose a sustainable delivery time slot, using machine learning to predict accurate time slots influenced by real-time data, so arriving sustainability advantages. In addition, seamless fleet utilization, reduced congestion, and reduced home delivery failures are achieved. In the preparatory phase, new sustainable routes are included based on expected traffic or other factors. Stretching time slots from 2 to 4 hours provides a journey distance refinement of approximately 5.5% and a cost reduction of 11%, if merely 20% of clients agree to the extended delivery time windows.

The literature insufficiently manages the most crucial aspect of last-mile logistics, “Time Window (TW)”. Current study highlights a dearth in the capacity to include real-time data into the decision-making approach. Consumers in AHD can order at any time and from any premise, with different freight dimensions, making it difficult to compute feasible time slots while the consumer awaits the offer, as all decisions must be computed in real-time. Last-mile delivery efficiency is dependent upon numerous factors, such as consumer density, depot capacity and location, shipment size distribution, uniformness, traffic conditions, and congestion. Hence, the ideal solution is to deploy ML to precisely predict the options of new consumers regarding their orders, particularly timing and location for delivery execution. The essence of last-mile logistics optimization is primarily aimed on cost reduction rather than environmental effect reduction.

The responsible AHD model is a newly established methodology focused at refining sustainability in AHD for large retail organisations and logistics service providers, leveraging real-time data and machine learning. Primarily, the focus is to refine the ADH delivery process through routine methods by streamlining routing with superior models, selecting a more suitable resources for optimized efficiency and reduced environmental effect, along with other available solutions. The key development is applicable to TW enhancement through sustainable processes, which simultaneously provide reduced cost. Ultimately, the key innovation is in optimizing the environmental impact of AHD. As a result, low cost can be realized, and renewable practices draw in new customer base. The transformation to prioritizing sustainability in business when choosing TW, instead of focusing only on cost efficacy, is a basic aspect of the innovative strategy. The simulations of AHD deliveries executed for model validation exposed that providing TWs to consumers is a critical phase in identifying potential TW while simultaneously achieving refined sustainability objectives. To recognize feasible delivery time slots, all delivery constraints must be included into machine learning, which also incorporates the retailer's business model, as all subsequent decisions depend on it. When the author attempted to deploy tighter time windows, such as one hour, to provide a feasible time slot, all indicators for economically and sustainably efficient delivery declined. Using a ML Prediction that accounts for various sustainable delivery windows, this article examines the effects of an AHD model that is environmentally friendly. To achieve accurate prediction outputs, the relevant model is trained using past data compiled from track and trace systems during execution delivery slot. With the inclusion of real-time data on traffic loads, weather, freight capabilities, and delivery capacities, as well as additional elements particular to each location, continuous reskilling model is recommended.

Orsic (2023) suggests a method that uses machine learning, specifically supervised learning frameworks, to anticipate feasible and sustainable delivery time windows for consumers. The article highlights the continuous reskilling of these models using real-time data from the delivery process as well as external sources such as traffic conditions and weather patterns, confirming that predictions remain accurate and responsive to dynamic operational contexts. The study narrates a three-step delivery process including time slot prediction, enhanced planning, and execution assisted by real-time feedback loops, specifying a practical approach to refining both efficiency and reliability in last-mile delivery. This article provides a new "Sustainable AHD framework" that deploys machine learning to predict accurate, sustainable time slots for AHD in real-time, with the aim of encouraging consumers to pick sustainable delivery options. The method integrates real-time data from many sources to refine the planning and execution of AHD deliveries for sustainability. The framework's goal is to prompt consumers to select sustainable delivery slots, as research proposes that consumers are prepared to withstand longer delivery slots for more sustainable deliveries. The architecture proposed by Orsic (2023) optimizes forecasts by leveraging real-time data from the planning and execution phases of delivery, as well as other sources, to deliver individualized, accurate, and sustainable TW choices for AHD in real-time environments.

2.12 AI-Driven Optimization in Quick-Commerce Delivery Routing

Recent study in vehicle routing optimization for Quick-Commerce evaluates AI/ML methods to streamline last-mile delivery efficiency. Studies examine the application of ML models to refine vehicle routing problem (VRP) modeling and optimization (Bai et al., 2021). Authors create tools including multiple ML models to investigate multi objective routing issues, resulting in decent improvements in distribution distance and time (Boumpa et al., 2022; T soukas et al., 2023). These models experiment with average savings of up to 25% in distance and 14% in total distribution time. Comparative assessment of various upgrading algorithms, including Large Neighborhood

Search (LNS), Particle Swarm Optimization (PSO), Differential Equation (DE), and Ant Colony Optimization (ACO), highlight that hybrid models, especially LNS, ASO, surpass algorithms in addressing the Vehicle Routing Problem (VRP) with temporal challenges. These enhancements highlight the underscore of the probable of AI/ML models in improving Quick-Commerce distribution workflows.

Boumpa (2022) suggests an efficient solution to the VRP leveraging ML algorithms at many phases of the delivery workflows, attaining up to 25% decrease in distance and 14% reduction in distribution time. This study investigates the distribution of trading items to consumers through a platform that supports multiple orders. The platform aims to offer an outcome to the VRP and its derivatives, including the Capacitated Vehicle Routing Problem (CVRP). The approach incorporates all phases of the ordering procedure till the product is delivered to the consumer. The suggested platform leverages multiple ML algorithms to dynamically recognize distribution locations, allocate the relevant batch of orders to available logistics assets, and enhance routing based on existing traffic situations. Moreover, various elaborate multi-objective issues are also considered into account. Boumpa (2022) managed procedures like K-means clustering to recognize geographic regions for grouping orders, modified neural network to diagnose batch configurations that optimize the number of items included while reducing distance, and visualization of routes leveraging Google Maps.

The suggested platform may identify distribution routes that are up to 25.61% shorter (specific scenarios) than those taken by experienced drivers, able to recommend delivery routes that are 14.22% shorter in total time than routes taken by experienced drivers. The platform's delivery process optimization can assist the organization both financially (lower fuel spending) and in terms of quality metrics (goods delivery time). Currently, the platform analyses a single point of cargo load, but the authors aim to explore with various dynamic load sites in the future. The study is executed over a two-month timeline, which may not cover all shifts in market and consumer journey. In the future, the recommended platform will consider various cargo loading locations and factors in dynamic points at the start of the logistics journey.

2.13 Optimization of the Initial Rider using Genetic Algorithm

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) is leveraged in a range of optimization scenarios. In logistics, it is utilized for perishable product distribution (Nisperuza et al., 2019) and transit network design. For multi-depot pick up and distribution use cases, a two-stage hybrid model including NSGA-II represents superior performance compared to particle swarm optimization (Wang et al., 2021). In recent times, the first proven performance assures for NSGA-II on an optimization problem, particularly the spanning tree problem is established, proving its observed efficacy. These studies prove NSGA-II's relevance in fixing complex multi-objective optimization problems across various domains. The model creates quality solutions, secures population diversity, and examines solution spaces, proving its contribution for researchers by addressing optimization challenges and risks.

Wang (2021) explores a multi-depot pick up and distribution scenario comprising resource sharing and assigns NSGA-II within a two-stage hybrid model to optimize vehicle routing efficiency. Merging Resource Sharing (RS) into the optimization of the multi-depot pickup and delivery problem (MDPDP) considerably reduces logistics operational cost and the relevant transportation assets through the restructuring of logistics network. This study develops mathematical programming model to reduce logistics cost and the number of vehicles, allowing for multiple usage of assets by one or several logistics facilities. A two-stage hybrid model is designed, integrating a k-means clustering algorithm, a Clark-Wright

(CK) algorithm, and a non-dominated sorting genetic algorithm II (NSGA-II). The k-means algorithm is leveraged in the initial design to reassign consumers to logistics infrastructure based on the distance. In the next stage, a joint force of CW and NSGA-II is leveraged to enhance vehicle routes and identify optimal results. The CW algorithm is considered to identify the initial results, hence streamlining the efficiency of identifying the enhanced solution in NSGA-II. Finally, benchmark tests are conducted to evaluate the performance and efficacy of the proposed two-stage hybrid algorithm. Quantitative results demonstrate that the proposed framework overcomes the standard NSGA-II and multi-objective particle swarm optimization algorithm.

2.14 Real-Time AI-Driven Supply Chain Analytics

The transformative effect of AI/ML on supply chain management (SCM) is the growth engine of several studies. Kafka and Akka, two real-time data processing techniques, improvise operational agility and transparency in the supply chain (Singh, 2023). The integration of IoT, Big Data, and AI refines production and logistics planning, with unique ideas like Incessant Data Processing (IDP) exhibiting considerable improvements in prediction accuracy and analysis rate (Zeng & Yi, 2023). Reinforcement learning models, particularly Q-learning, are steadily applied to SCM use cases, particularly in inventory management, though many studies still use artificial data (Rolf et al., 2022). AI and ML are accelerators for value creation in SCM, enhancing efficiency and improving resilience across operations from demand forecasting to logistics and consumer experience (Singh, 2023). The models are shaping the future of SCM, despite challenges in data privacy and workforce impacts.

Singh (2023) recommends an architecture that leverages Kafka and Akka to support real-time data processing and decision-making in SCM, as explained in their article, "Revolutionizing Supply Chain Management: Real-time Data Processing and Concurrency". In the existing business climate, effective SCM is imperative for firms focusing to succeed in the face of changing market dynamics and increasing consumer expectations. This study introduces an innovative method to SCM that leverages advanced techniques, particularly Kafka and Akka, to streamline data integration and decision-making processes. Leveraging Kafka as a distributed event streaming platform and Akka as a toolkit for simultaneous and distributed applications, this system introduces effective communication and collaboration among different nodes in the supply chain network. This paper elucidates the complexities of the suggested architecture, explaining the implementation approach and performance assessment metrics. This paper explains how our result refines supply chain visibility, promotes agility, and drives real-time responsiveness to market changes and consumer journey. Moreover, practical use cases explain the significant features of the approach on enhancing inventory management, streamlining order fulfilment efficiency, and improving last-mile journey in logistics. In addition, researchers investigate the challenges faced during implementation and post Go-Live, providing insights into feasible mitigation strategies. This study explains significant avenues for future research, identifying emerging trends and opportunities in SCM enabled by Kafka and Akka methods.

The key conclusions are that the suggest model, which leverages Kafka and Akka, has the potential to transform SCM by optimizing data integration and decision-making procedures. It also optimizes inventory management, order fulfilment efficiency, and logistics optimization. It improves supply chain visibility, promotes operational excellence, and allows real-time responsiveness to market volatility and consumer expectations. Future research, according to the practitioners, strive to explore new developments and opportunities in the field of SCM facilitated by Kafka and Akka technologies.

2.15 Solutions for the Last-Mile of Delivery in Quick-Commerce

Last-mile delivery (LMD) in Quick-Commerce gains significant focus due to its influence on logistics cost and customer satisfaction. Latest study focuses on smart ideas to enhance LMD efficacy, including AI-powered technologies like drones, robots, and autonomous vehicles (Sooroshian et al., 2022). A recommended framework consolidates social media analytics and deep learning techniques for traffic forecasting in urban LMD (Laynes-Fiascunari et al., 2023). Key factors impacting LMD costs include failed deliveries, customer segmentation, and workflow automation (Mangiaracina et al., 2019). Practical options to handle these challenges include parcel lockers, crowd sourced logistics, and dynamic pricing principles (Mangiaracina et al., 2019). The domain of LMD study progresses well, with most publications developing in the past five years (Olsson et al., 2019). A suggested model for last-mile delivery study consists of five interconnected components: last-mile logistics, order fulfilment, distribution, transport, and end-delivery (Olsson et al., 2019).

Mangiarcina (2019) details unique strategies to improve last-mile shipment effectiveness in digital retailing for consumers, including parcel lockers, crowdsourced logistics, and dynamic pricing. This research has two objectives: primarily, to assess and categorize scientific journals engaged on new solutions that streamline the efficiency of last-mile shipping on business-to-consumer (B2C) e-commerce; and finally, to explain avenues for future research in this domain. The last-mile delivery process provides particular focus for refinement due to its considerable impact on overall logistics cost and the economic sustainability of a B2C venture. The study assures that the key factors influencing cost include the chance of unsuccessful deliveries, client cluster in delivery regions, and the level of process automation. Smart last-mile delivery options that may influence the drivers including parcel lockers, crowdsourced logistics, consumer proximity mapping, and dynamic pricing strategies. Finally, few gaps and opportunities for more research efforts are appreciated. This review shares valuable findings for both practitioners and scholars. This study investigates and categorizes literature on creative and efficiency focused last-mile delivery models, proposing areas for future research efforts. From a management lens, it provides a holistic overview of the critical elements influencing last-mile delivery cost and possible new solutions that can be leveraged to optimize operational efficiency.

From a financial perspective, the cost of last-mile logistics includes three primary elements: average transportation cost, driver cost, and opportunity cost, which are impacted by various attributes. This study examines innovative last-mile delivery results that can affect these cost drivers, including parcel lockers, crowdsourcing logistics, drones, and dynamic pricing. Author investigates gaps in the existing literature, including the lack of holistic frameworks measuring all innovative solutions and the requirement for more research on certain unexplored topics.

This research comprehensively handles two central questions aligned with the research topics. The first question investigates the primary attributes impacting the last-mile shipping cost. An exhaustive analysis of these elements and their deliverables is imperative to interpreting how innovative ideas can effectively mitigate distribution costs. The second question evaluates which unique last-mile shipping techniques could positively affect these cost drivers, facilitating a measurable decrease in overall distribution cost. By identifying those cost drivers, it becomes feasible to assess the effectiveness of innovative solutions and evaluate their operational and economic outcomes.

The study highlights various needs and significant directions for future research. High-level research that provides a holistic overview of smart last-mile results while assessing their relevant cost drivers are relatively insufficient. Most articles engross on a limited set of results, often highlighting optimization principles or comparative benefits over traditional distribution methods. Certain groundbreaking delivery methods, such as underground or in-car deliveries, remain underexplored, mostly due to the practical challenges associated with their deployment. Even for full-blown solutions like parcel lockers, the limitations continue. Existing study primarily depends on qualitative assessments, including consumer surveys and intentions studies, underscoring the requirement for stringent, quantitative evaluations to understand their outcome on last-mile cost reduction. Therefore, frugal research is conducted on the "crowdsourcing logistics business model', which pursues the crowd as a source of on-demand workforce, despite its prospective as a successful shipping solution expected to attract urban logistics (Wang et al., 2016). Moreover, other industries may significantly derive decent benefits from the use of dynamic pricing, and future studies might focus on these areas. Consequently, the passion in rolling out the results in industries beyond the grocery sector is increasing among practitioners, as seen by initiatives like Milkman delivery.

2.16 Enhancing Last-mile delivery with Generative AI

By improving performance and gratifying consumers, Generative AI is altering the face of last-mile delivery in Quick-Commerce. AI enabled models, both tangible (robots, drones) and intangible (decision support tools), are enhancing last-mile value chain, streamlining productivity and sustainability (Sooroshian et al., 2022). These benefits address risks like mounting shipment volumes and consumer demands, specifically in estimating delivery times. Generative AI is critical in developing data-driven models that permit refined omnichannel commerce experiences, overcoming conventional AI in unifying retail operations (Singh, 2023). The expanding need for immediate delivery yields last-mile delivery a key enabler of success for e-commerce companies (Sharma, 2024).

Logistics organisations can efficiently embrace AI solutions by sticking to standardized models such as “Crisp-Delivery Model” framework, which assures adequate data preparation and analysis. Although these models provide significant benefits, they also create problems that must be resolved for successful implementation (Sooroshian et al., 2022).

Singh (2023) underlines that Generative AI is increasingly dominating the technological landscape and continuously interrupting both technology and consumer journey sectors. One of the significant milestones of the post-Covid era is the omnichannel commerce experience. The Covid era enforced retail giants to leverage all their resources to share optimal user options and consolidate them, hence creating a unique experience that combines technology enabled digital commerce and fulfilment methods with the experience of store shopping. The seamless experience of omnichannel commerce can be driven by the capabilities of extensive data-driven models and intelligence powered by Generative AI. Conventional AI leverages by shops to date, substantiating growth and refining a cohesive experience for both retailers and their customers.

2.17 Case Studies and Frameworks

The submission of AI, ML, and Deep Learning in innovative logistics is a fast-evolving field, demanding an empirical approach to comprehending its potential. Woschank, Rauch, and Zsifkovits (2020) provide a conceptual framework for addressing future research on AI, ML and Deep learning in smart logistics. Their structured literature review emphasizes the potential challenges around AI, such as the feasibility of AI methods exceeding human capabilities and potentially curbing human survival. They propose that the late adoption of AI in Latin America could be strategic advantage, as it permits for assessing and reducing significant negative impacts. Akbari and Do (2021) review examine the current state of research on ML applications in logistics and SCM. Their structured review and publication analysis of numerous articles exhibits that only a small portion of literature reviews are published on the application of ML in logistics and SCM. Their evaluation was limited to assessing only academic sources from certain databases and to sources including specific keywords in the title or abstract. These reviews highlight the importance of designing robust methods and executing exhaustive case studies to assist the ethical implementation of AI/ML methods in Quick-Commerce. Future research should strive to expand the scope to include a diverse range of sources and research the application of ML to detection problems in logistics and SCM

2.18 Gap Analysis

The study of current literature on AI/ML enhancement in last-mile delivery for Quick-Commerce reveals gaps that underline the need for more research. While AI/ML models exhibit promise in streamlining logistics efficiency, current implementations often be without the refinement to handle real-world complexities in a dynamic distribution ecosystem.

One major gap is the limited evaluation of hybrid AI/ML models. Most investigations aim on single method approaches that do not influence the integral strengths of various technologies. This challenge resists the potential for perfect solutions that can manage varying distribution scenarios and diverse operational demands. Moreover, there is clear deficiency in handling biases withing existing methods. While some study acknowledges the existence of disparities in service distribution, there is a lack of structured approaches to identify, quantify, and reduce these biases. This lapse prevents the ability of service providers to provide fair and unbiased distribution experiences across various customer demographic segments.

Also, the existing literature often fails to provide practical solutions for integrating and launching AI/ML solutions within the operational challenges of Quick-Commerce. The current literature does not deep dive into the entire value chain of Last-mile delivery in Quick-Commerce. The primary value chain attributes like, "End customer and Returns/Reverse Logistics" are missing in AI/ML roadmap. The research handles these gaps by providing innovative hybrid AI/ML models particularly designed for the complexities of Quick-Commerce distribution. The focus is to create frameworks that not only improve operational efficiency but also consolidate fairness and integrity into the distribution process. Moreover, the study suggests strategic platform that assist firms in executing these models effectively, assuring scalability and flexibility in a fast-growing market. The introduction of these advanced solutions is quintessential for multiple reasons. Primarily, by handling algorithmic biases and executing hybrid AI/ML models, logistics providers can attain higher levels of operational efficiency and service quality. Next, systematic bias mitigation guides integral distribution delivery, assuring that consumers across diverse markets receive reliable and fair access to services, which in turn refine overall satisfaction. Ultimately, the acceptance of flexible and scalable integration strategies permits companies to reach effectively to dynamic market conditions while maintaining ideal performance in last-mile delivery. By linking these gaps, the study delivers to the development of integral, advanced and effective logistics solutions for the Quick-Commerce.

2.19 Summary of Findings

The literature review examines various insights relevant to enhancing AI/ML solutions stacks for last-mile delivery in Quick-Commerce. Primarily, while the existing AI/ML algorithms prove significant capacity for optimizing distribution efficiency, many tussle to adapt to dynamic operational factors, highlighting the need for flexible and robust solutions. Secondly, there is growing recognition of the benefits of hybrid AI/ML approaches that consolidate multiple methods, which can handle varied logistical challenges than single-model methods. Thirdly, various research determines algorithmic biases that impact the fairness and integrity of distribution services, reinforcing the importance of strategies to detect and reduce such inequalities. Next, existing literature points to a gap in practical frameworks that consolidate AI/ML methods within the operational challenges of Quick-Commerce, underlining the requirement for customized solution stacks. Ultimately, the effective utilization of real-time data appears as a primary factor for enhancing decision-making and operational excellence, yet its consistent integration into AI/ML approaches remains limited. These insights emphasize both the opportunities and gaps within the existing AI/ML framework for last-mile delivery. There is a keen interest in designing integral, refined, and scalable solutions, beside a need for structured frameworks to assist practical roll out in the Quick-Commerce sector.

2.20 Implications for My Research

This research creates the basis for streamlining the competency of last-mile delivery in Quick-Commerce by enhancing AI/ML platforms. The findings gained will reinforce the need to design advanced models that handle existing challenges in delivery operations. By aiming on the consolidation of hybrid AI/ML models, the study aims key gaps in current solutions, offering a resilient methodology for enhancing delivery efficiency, accuracy, and performance. The result of this study serves as a basis for consecutive inquiries (Giuffrida, 2022) regarding the application of AI/ML models in collaborative Quick-Commerce and provides a holistic platform suggestion for professionals.

This chapter begins with the overview of last-mile logistics value chain initiating from E-commerce service provider, Transport partner, Warehouse and logistics partner, and last-mile delivery partner. It probes into the leverage of AI/ML models across various scenarios and responds with summary, methodologies, findings, challenges and future research. The chapter explores the application of theories such as Vehicle Routing Optimization (VRO), the SMOTE-Tomek algorithm (Kelkar et al., 2021), Particles Swarm Optimization (PSO) (Verbytskyi, 2024), Ants colony optimization (ACO) (Verbytskyi, 2024), Genetic algorithms (GA), Large Neighbourhood Search (LNS), Blackhole Optimization (BHO) (Banyai, 2018), Random Forest, Stochastic Gradient Descent algorithm (Al Rahib, 2024), K-Means Clustering, Non-dominated Sorting Genetic Algorithm (NSGA) (Wang et al., 2021), and Kafka. This chapter meticulously assesses the impact of AI/ML model strategies on E-commerce product availability, robotics and cobotics in warehousing, predictive order fulfilment, multiple pickups, and dynamic distribution channels, while also handling issues related to biases in AI/ML adoptions.

In addition, current inquiries into the AI/ML last-mile platform are evaluated, highlighting its success in achieving last-mile goals such as performance, cost and customer satisfaction (Verbytskyi, 2023), beside the application of AI/ML models. This chapter suggests future research directions, such as Generative AI, Agentic AI, and Generative adversarial networks (GAN), recommending for the inclusion of ML and AI within the last-mile value chain, thus embracing the audience in the continuous advancement of the foundational platform for last-mile in collaborative Quick-Commerce.

The study leverages findings on current logistics constraints and AI/ML trends to suggest solutions that are both effective and scalable. These ideas will enable logistics vendors to enhance resource allocation, steer urban complexities, and revert proactively to volatile market demands. By consolidating real-time data and location intelligence, the research will allow for dynamic decision-making, optimizing holistic distribution performance and customer satisfaction. Moreover, this study will deliver to the field by initiating innovative ideas to reduce biases within algorithmic workflows, assuring fair and transparent service across all demographics. This methodology balances operational responsibility with technological transformations, benchmarking new standards for responsible AI practices in logistics. The study will provide a trail for sustainable and flexible “AI/ML based technology stack” that can stay competitive with evolving technological trends and market constraints, finally enabling the way for resilient and high-performing last-mile delivery systems in Quick-Commerce.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Research Design

A mixed-methods research design is conducted in this study that blends both qualitative and quantitative approaches. As mentioned in Chapter 2, gaps are managed in enhancing AI/ML solutions stack for last-mile delivery in Quick-Commerce. This dual methodology facilitates holistically understanding the theme by utilizing the strengths of both methods, quantitative analysis for assessing algorithms' performance and recognizing biases, and qualitative insights for investigating execution problems and contextual factors. A mixed-methods approach is specifically designed to this study due to the diverse aspects of the research problem. Chapter 2 emphasized significant gaps, including the limited understanding of AI/ML model effectiveness in real-world last-mile delivery perspectives, the occurrence of biases in existing algorithmic models, and the need for hybrid architectures that consolidate multiple AI/ML algorithms.

The literature review (Chapter 2) highlighted key frameworks and insights, such as the requirement for performance efficiency and enhanced routing methods that contribute for urban complexity, the importance of managing algorithmic biases that impact service integrity, and the possibility of hybrid AI/ML methods to upgrade reliability. This research methodology includes these results by "Quantitative phase", which designs computational assignments and data analysis approaches to measure algorithm performance across diverse scenarios, leveraging on methods like systems theory and algorithmic fairness theory, and "Qualitative phase" which includes structured interview and focus group procedures to examine organizational factors, execution constraints, and stakeholder opinions, assisted by models such as the Technology Acceptance Model (TAM) which emphasizes the persona's acceptance of technologies and Resource based view (RBV) which underscores that a firm's sustainable competitive advantage occurs from its unique resources and skills.

This research model directly aligns with the objectives set up in Chapter 1. It analyses the effectiveness of current AI/ML models in enhancing last-mile delivery indicators while also determining the biases ingrained within existing algorithmic platforms and measuring how these biases affect service integrity. This study aims on designing a hybrid AI/ML model that can optimize overall performance and resilience. Ultimately, it demonstrates a strategic guide to assist the selection and recommendation of algorithms based on operational challenges, assuring flexible and unbiased last-mile delivery methods.

The detailed flow of the research is highlighted in the following sections. As illustrated in Figure 3.1 (a), the research was conducted in three unique and sequential phases.

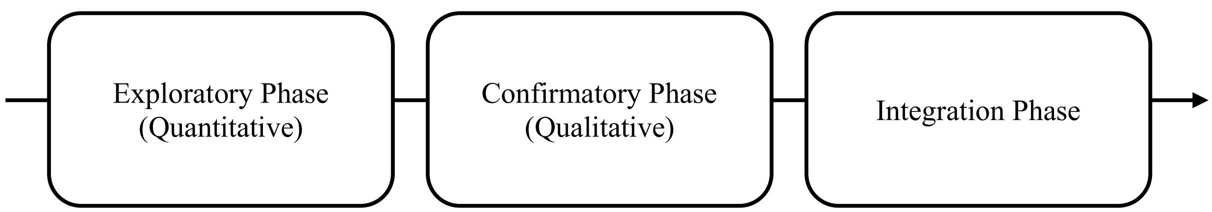

Figure 3.1a: Different phases of research (original work)

The goals, methods employed, and deliverables associated with each phase are demonstrated in Table 3.1. The gaps discovered underline the need for a dual-method approach. Quantitative methods, including computational experiments and historical data analysis, allow evaluation of algorithmic performance and the identification of biases across various operational scenarios. Enhancing qualitative techniques, such as interviews and focus group discussions, simplify a deeper analysis of contextual drivers, execution hurdles, and stakeholder opinions that shape the practical effectiveness of AI/ML models within last-mile delivery systems.

Phase	Objective	Method	Outcome
Exploratory Phase (Quantitative)	Evaluate algorithm performance, identify biases, and test hybrid approaches	Computational experiments, historical data analysis, and structured surveys	Objective measurements of performance metrics, identification of biases, and statistical validation of hybrid approaches
Confirmatory Phase (Qualitative)	Explore contextual factors, implementation challenges, and stakeholder perspectives	Semi-structured interviews, focus groups, and case studies	Rich contextual insights into factors influencing AI/ML effectiveness in real-world settings
Integration Phase	Synthesize quantitative and qualitative findings to develop a comprehensive framework	Integrated analysis combining statistical results and thematic insights	A validated framework for optimizing AI/ML technology stack for last-mile delivery in Quick-Commerce

Table 3.1: Research flow

The goals, methods employed, and deliverables associated with each phase are demonstrated in Table 3.1. The gaps discovered underline the need for a dual-method approach. Quantitative methods, including computational experiments and historical data analysis, allow evaluation of algorithmic performance and the identification of biases across various operational scenarios. Enhancing qualitative techniques, such as interviews and focus group discussions, simplify a deeper analysis of contextual drivers, execution hurdles, and stakeholder opinions that shape the practical effectiveness of AI/ML models within last-mile delivery systems.

The benefits of the mixed-methods technique embraced in this study are illustrated in Figure 3.1 (b).

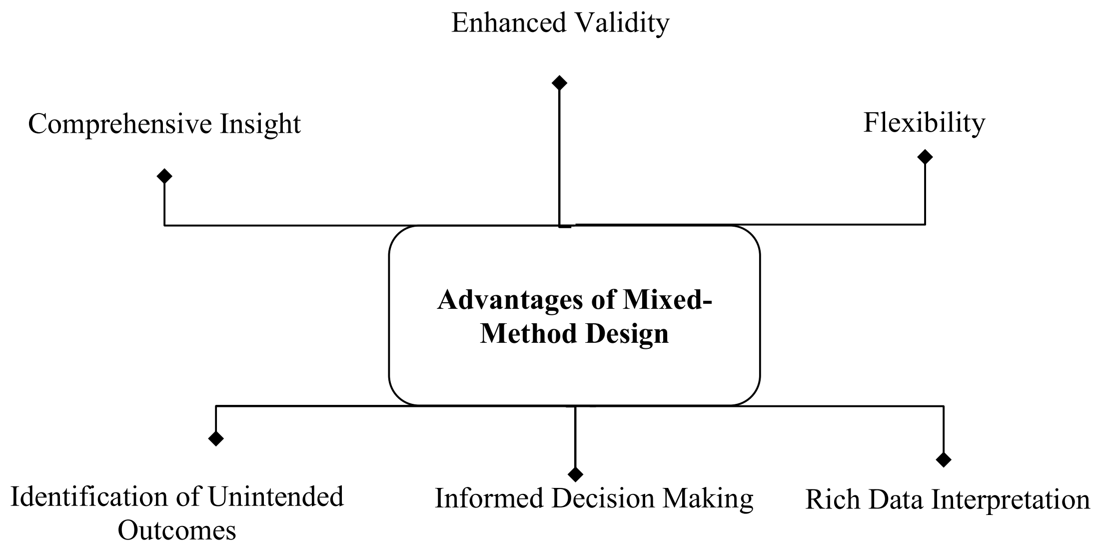

Figure 3.1b: Advantages of mixed-method design (original work)

This systematic integration facilitates holistic comprehension of the research problem by blending qualitative insights with quantitative proof. Qualitative data supports the review of user experience and stakeholder views regarding AI-enabled last-mile delivery services, while quantitative data drives the assessment of performance metrics and operational efficiencies. The triangulation of these data points streamlines the overall validity of the results and strengthens the reliability of the conclusions drawn on the efficacy of AI/ML techniques. In addition, the mixed-methods approach provides flexibility in the evaluation of various aspects of Quick-Commerce. It permits for initial qualitative probe to identify primary challenges and deliver hypotheses, followed by quantitative review to assess the impact of proposed interventions. Consolidating illustrative feedback from interviews and focus group discussions with data-driven analytics offers insights into the efficacy of AI/ML methods. The strategic experiment contributes to better-informed decision making, as stakeholders gain access to a comprehensive analysis that considers operational excellence besides user experience and customer satisfaction. Moreover, qualitative inquiry empowers the identification of unforeseen results or upsurging benefits that may not be captured through quantitative evaluation alone, hence enriching the overall assessment of AI/ML execution within last-mile delivery methods.

3.2. Data Collection Methods

The data compilation approach was structured to receive holistic ideas into enhancing AI/ML solutions stack for last-mile delivery in Quick-Commerce sector. In alignment with the research goals and the gaps identified in Chapter 2, the research assigns both quantitative and quantitative data collection approaches to assure a robust and versatile base. This integrated method strengthens the breadth and depth of the data point while ensuring consistency with the comprehensive research framework. The following sections discuss each data collection method in detail

3.2.1 Qualitative Data Collection (Semi-Structured Interviews)

The core objective of this part of the study is to investigate stakeholder views, operational challenges, and rising opportunities associated with enhancing AI/ML platforms in last-mile delivery. To attain this, primary decision-makers representing a range of functions were committed, including Quick-Commerce vendors, retail logisticians, AI/ML technology architects within retail and consumer goods companies, last-mile delivery partners, warehouse and distribution experts, and end consumers who dynamically leverage Quick-Commerce services. Participants were shortlisted based on their direct involvement in last-mile operations and AI/ML executions. This calculated sampling strategy secures that the individuals consulted retain relevant knowledge and experience, hence optimizing the credibility and richness of the insights collected. Semi-structured interviews were conducted to drive the in-depth measurement of salient points. The discussions aimed on current applications of AI/ML techniques in last-mile delivery, operational constraints and resistances faced in real-world scenarios, and potential levers for hybrid AI/ML methods to improve system performance. In addition, weightage was given on the factors influencing algorithm selection and execution, experiences related to algorithmic bias and its implications for service outcomes, and strategic parameters for enhancing key performance metrics.

Table 3.2a explains the data collection approach regarding method, duration, recording and validation.

Method	Interviews conducted virtually via video conferencing tools to accommodate geographical diversity
Duration	45-60 minutes per interview, allowing for in-depth exploration of themes
Recording	Audio recordings with participant consent, transcribed for analysis
Validation	Member checking where participants review key findings to ensure accuracy
Sampling	Purposive sampling to ensure representation across all stakeholder groups
Tools	Zoom for interviews, Otter.ai for transcription, NVivo for data management

Table 3.2a: The data collection process

3.2.2 Quantitative Data Collection (Structured Surveys)

The quantitative data collection focuses to receive measurable insights on AI/ML execution practices, performance outputs, and stakeholder experiences across a generic base. Structured surveys were deployed for the same decision-making groups engaged during the interview phase, with customized queries aligned to the specific knowledge and responsibilities of each group. Extracting insights from the qualitative findings, the survey was designed to quantify critical attributes related to adoption, including awareness and integration points of AI/ML models, as well as constraints such as technical, organizational, and cultural barriers. It also evaluates the impacts of AI/ML implementation on last-mile delivery performance.

The survey includes multiple-choice questions to acquire categorical data, ordinal scale items to measure attitudes and perceptions, and selectively positioned open-ended prompts to compile additional contextual insights. This structure facilitates a balanced representation of both numerical assessment and participant views. The overall survey design and approach are illustrated in Figure 3.2.

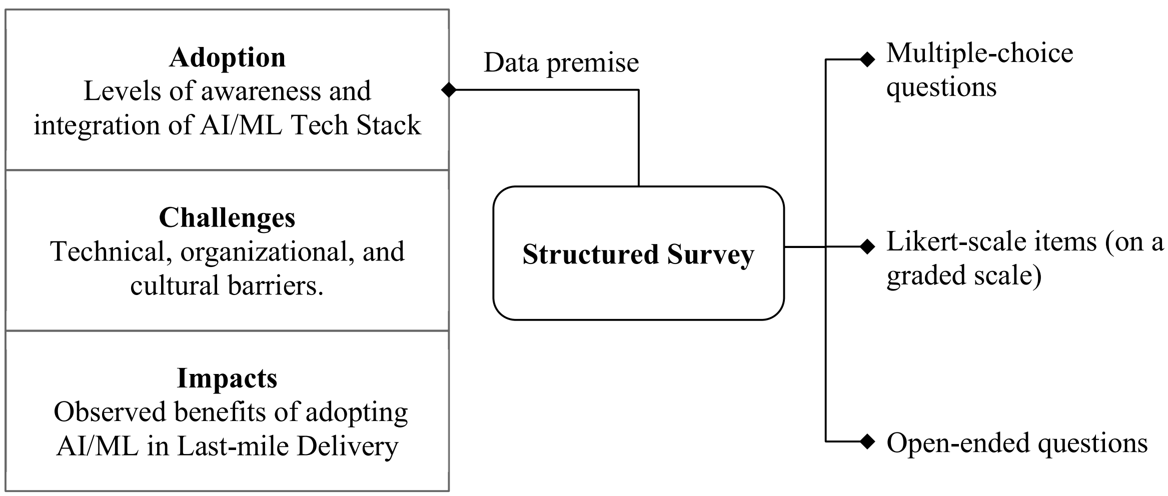

Figure 3.2: Survey design and approach (adapted from Mangiarcina, 2019)

The survey instrument was structured to collect information across various key areas, such as demographic details including organization dimensions, industry sector, and geographic presence; current AI/ML models, including algorithms and integration methods, performance measurement practices, and challenges faced in operational perspective. Participants were also questioned to evaluate the perceived productivity of numerous algorithmic models, share feedback related to potential biases and service integrity concerns, and recognize priorities for enhancement and future improvement.

The survey was distributed to practitioners employed across small, medium, and large enterprises within the technology, logistics, and retail sectors, assuring illustration from core functional and software roles. A convenience sampling method was adopted to acquire a broad and distinct participant group. To assist statistical reliability and robust analysis, surveys were distributed through email and professional networks, with a targeted response range of 200 to 300 participants. The sampling strategy and distribution process are detailed in Table 3.2(b). The survey experienced a pilot test with 10 participants to ensure clarity and reliability. Feedback was included from the pilot phase into the final survey design.

Design	Online surveys with Likert scales, multiple-choice, and ranking questions
Distribution	Email invitations to industry networks, professional associations, and consumer panels
Sample Size	Target of 50+ respondents from each stakeholder group (except consumers: 200-300)
Duration	15-20 minutes to complete, balancing depth with participation rate
Validation	Pilot testing with 5-10 representatives from each stakeholder group
Tools	Qualtrics for survey design, distribution, and initial data collection

Table 3.2b: Sampling strategy and distribution process

3.2.3 Computational Experiments

To facilitate an unbiased review of algorithmic performance and evaluate the potential of hybrid approaches, a series of computational tests was incorporated. These tests deploy simulation environments that reproduce urban and non-urban distribution scenarios with numerous levels of complexity, beside masked historical distribution data rendered by industry collaborators. Benchmark datasets are additionally leveraged to drive standardized comparisons across different algorithmic models. Digital twin representations of distribution networks guide performance assessment under supervised and controlled conditions. Details of the experimental sampling and configuration are highlighted in Table 3.2(c).

Algorithms Tested	Route optimization, demand forecasting, resource allocation, and hybrid approaches
Scenarios	Urban, suburban, and mixed environments with varying delivery volumes and time constraints
Performance Metrics	Delivery time, accuracy, cost efficiency, resource utilization, and equity measures
Variables Manipulated	Traffic conditions, order volume, geographic distribution, and time constraints
Control Measures	Standardized starting conditions and evaluation metrics across all tested algorithms
Tools	Python for algorithm implementation, SUMO (Simulation of Urban Mobility) for traffic simulation, Google OR-Tools for optimization

Table 3.2c: Experimental design

3.2.4 Integration of Data Collection Methods

The integration of data collection approaches was streamlined to assure procedural synergy and holistic coverage of the study queries. A successive design was adopted in which computational experiments and surveys were handled initially, allowing their results to predict the focus of the following qualitative interviews. Cross-reference across multiple data sources facilitated the investigation of the same occurrences from various viewpoints.

While the quantitative model focused on identifying measurable outputs related to "what" and "how much", the qualitative method provided deeper insights into the "why" and "how" behind these patterns. An incremental refinement approach enabled seamless improvement of the research tools, with early insights assisting adjustments in later phases of data collection. The consolidation of results was attained through a mix of statistical analysis and thematic coding. Quantitative analyses were leveraged to recognize performance trends and relational patterns, whereas qualitative explanation disclosed contextual influences and stakeholder perspectives. Assimilating these observations ensured consistency between datasets and strengthened the robustness and resilience of the study's conclusions.

3.3 Data Analysis Techniques

The data analysis approach consolidates both qualitative and quantitative analytical models to meticulously examine the research objectives. This mixed-methods framework uses the integral strengths of each model, facilitating a nuanced and unique assessment of strategies for enhancing AI/ML platforms within last-mile delivery operations.

3.3.1 Qualitative Data Analysis

A thematic (Pattern based) analysis method was suggested to systematically assess the interview and focus group data, facilitating the design of insights through the identification of patterns that manifested during the analytical process. The evaluation followed a structured steps that began with open coding, in which line-by-line coding was executed to cover key concepts related to AI/ML rationalization within last-mile delivery, including execution challenges, algorithmic biases, and performance influencers. This was followed by axial coding, where related codes were assembled into wider conceptual classifications, allowing the discovery of recurring patterns such as organizational influences, technical constraints, and consumer satisfaction. The last stage comprised of selective coding, through which core thematic categories were established, such as strategies for refining AI/ML platforms and methods to reduce algorithmic bias, creating the functional elements of the arising theoretical framework. NVivo software was leveraged to assist systematic coding and organization of the qualitative data. To optimize the credibility of the findings, persona verification was considered to confirm that the interpreted themes remained consistent with participant experiences. The qualitative analysis provided meaningful insights into the specific operational challenges faced by organizations when adopting AI/ML models for last-mile delivery, the contextual attributes that transform practical algorithm performance, the analysis required to improve fairness and integrity in service delivery, and the decision-making factors that assist technology selection and roll out.

3.3.2 Quantitative Data Analysis

A blend of descriptive and inductive statistical models was deployed to the experimental and survey database, facilitating both summary-level assessment and the discovery of relationships among key variables. The analytical process started with descriptive statistics, including the calculation of measures of central tendency such as mean and median, analysis of variability such as standard deviation for business metrics and survey responses. Data visualization tools, such as bar charts and box plots were utilized to guide clear presentation of the results. Comparative statistical analyses were executed to assess variations in algorithmic performance across numerous operational factors, using t-tests, ANOVA (Analysis of Variance), and non-parametric workflows. Correlation and regression assessments were executed to examine associations among variables, for example, the impact of algorithm complexity on performance deliverables and operational scenarios. Python was leveraged as the core analytical environment with NumPy, SciPy, and Scikit-learn supporting quantitative calculation and algorithmic evaluation. The quantitative analysis yielded insights into the comparative performance of AI/ML methods across unique metrics and contexts, statistical patterns related to algorithmic bias and its impacts for integrity in service delivery, key predictors of effective installation and adoption of AI/ML models, and the relative benefits of hybrid frameworks over single-algorithm techniques.

3.3.3 Integration of Qualitative and Quantitative Results

The insights from the qualitative and quantitative analyses were interpreted to substantiate a holistic understanding of the research problem. This integration facilitated the alignment of technical performance observations with insights extracted from participant feedback. The integration process involved validating themes explored during the qualitative phase, such as roll out challenges, through supporting facts from experimental and survey pattern. Qualitative insights were further leveraged to frame statistical results, particularly in scenarios where quantitative patterns required in-depth analysis. The collective insights from both methods were then consolidated to formulate the suggested optimization model, assuring that it briefed both practical roll out and algorithmic performance. The integrated analysis benefitted in several key outcomes. It provided a holistic understanding of the attributes that influence the effectiveness of AI/ML models in last-mile delivery and provided to the design of a framework for technology selection and execution under different operational constraints. The results also provided actionable strategies for reducing algorithmic bias, enhancing performance reliability, and supporting the deployment of hybrid AI/ML methodologies. Moreover, practical guidance was conceptualized for companies aiming to adopt and operationalize such systems. This integrative approach was crucial for attaining a full and fact-based understanding of the research and for deriving feasible suggestions that is relevant for both industry and academic contexts. Combining the contextual depth of qualitative methods with the applicability of quantitative data improved the reliability of the suggested framework through validation of results. Figure 3.3 visually depicts the rationale and structure of this integrative analytical approach

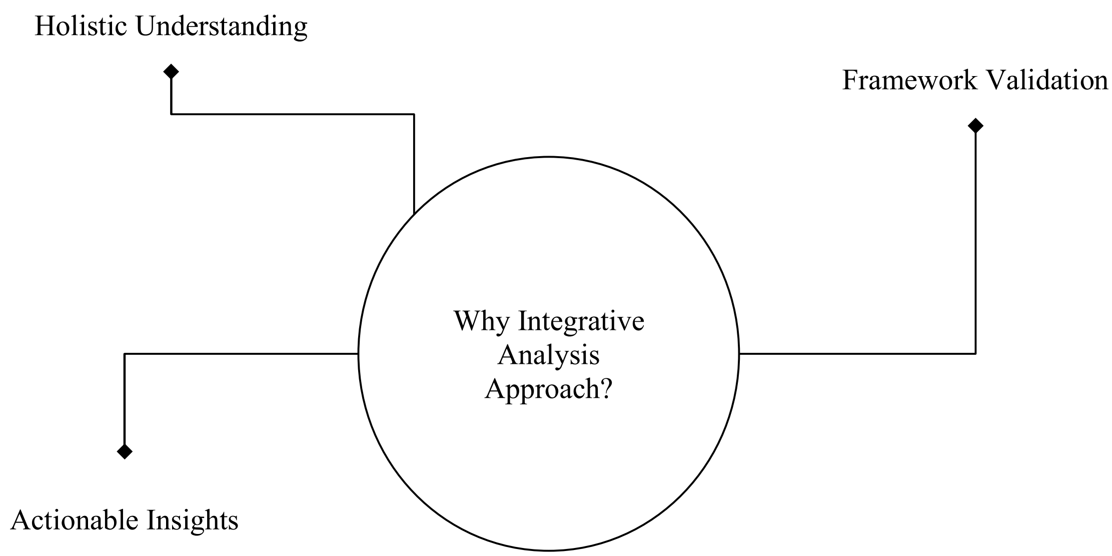

Figure 3.3: Illustration of the integrative analysis approach (original work)

The shortlisted data analysis models support a stringent and holistic examination of the research objectives, ensuring that the proposed model is both theoretically robust and practically viable within real-world operational contexts.

3.4 Ethical Considerations

The design of this study includes careful consideration of the ethical consequences associated with the use of AI/ML models in last-mile delivery, acknowledging significant impacts on fairness, privacy, and well-being. Ethical standards were robustly applied to protect participant rights, ensure non-disclosure agreements, and preserve the integrity of the study. Ethical principles guided each phase of the research, particularly data collection, analysis, and reporting. Entire research is adhered to institutional ethical principles, which are highlighted in Figure 3.4 and explained in detail in the sections that follow.

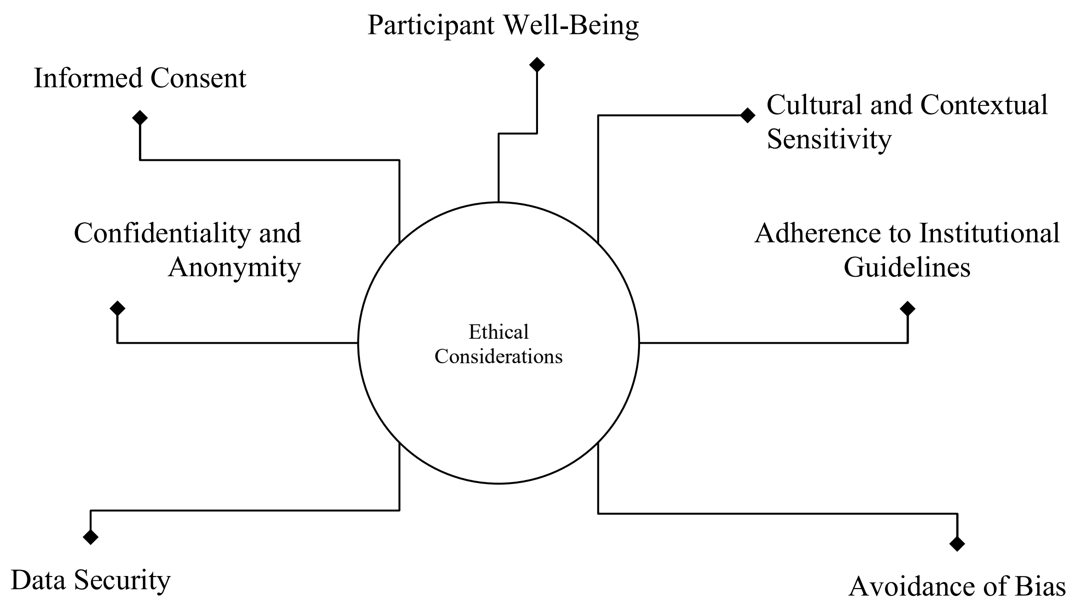

Figure 3.4: Various ethical considerations (original work)

3.4.1 Informed Consent

Respondents were given with clear and holistic information regarding the objectives of the research, the structured approach, and their expected role in the study. An informed approval process was conducted to ensure that participation was purely voluntary. The information shared with respondents included the reason of the study and its relevant benefits, the spontaneous nature of participation with the right to opt-out at any point without outcome, assurances of anonymity and confidentiality, and conclusion that the collected data would be leveraged exclusively for academic purpose. A resilient approval mechanism was implemented consistently across all data collection systems to ensure ethical compliance. For the interview phase, written consent forms were provided digitally and signed by respondents prior to the start of the sessions. For the survey component, precise consent option was grounded within the survey instrument, requiring respondents to acknowledge and agree to the terms before submitting their feedback. This approach confirmed that all respondents were informed and voluntarily agreed to contribute to the society.

3.4.2 Data Privacy and Security

Stringent measures were delivered to ensure the protection of respondent data and sensitive information throughout the research. All personal information were deleted from the datasets, and data were stored in encrypted, password-protected systems with access limited strictly to authorized members of the research team. The research process adhered to applicable data protection regulations, including GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act), while applying data minimization principles to limit the collection of identifiable information which was relevant for the research. Confidentiality and anonymity protocols were considered to protect respondent’s privacy. Interview transcripts were anonymized by deleting names, organizational affiliations, and project related references. Survey responses were collected without linking personal identifiers, except in cases where respondents voluntarily gave contact details for follow-up purposes.

3.4.3 Algorithmic Fairness

The focus of this research is to discover and mitigate biases in AI/ML models, wherein fairness considerations were prioritized throughout the analytical journey. Numerous fairness metrics were deployed to detect multiple forms of algorithmic bias, and multi-dimensional approach was considered to evaluate how these biases may affect demographic groups. The assessment of algorithms clearly accounted for integral implications beside performance outcomes. Mitigation strategies were designed to ensure that improvements in efficiency do not compromise fairness, hence supporting transparent last-mile delivery practices.

3.4.4 Avoidance of Bias

The research was designed to reduce bias and ensure the fairness of findings. A semi-structured interview guide was applied to limit interviewer influence while still allowing respondents the opportunity to convey relevant opinions. The survey instrument was designed with structured, neutrally toned questions and balanced response options to avoid leading respondents. Pilot testing was experimented to identify and correct potential sources of bias in the research instruments before full implementation. During data analysis, transparent coding procedures and meticulous statistical methods were implemented to support unbiased interpretation of deliverables.

3.4.5 Adherence to Institutional Guidelines

The study was organized in strict accordance with the ethical standards established by the Swiss School of Business and Management (SSBM), Geneva. Prior to data collection, informed consent was secured from all participants, ensuring full compliance with institutional ethics and standards.

3.4.6 Cultural and Contextual Sensitivity

Given the diverse background of the respondents, specific emphasis was placed on cultural and contextual sensitivity across the research process. Neutral and unbiased language was considered in interviews and surveys to mitigate cultural influence. Versions in organizational contexts and geographical regions were acknowledged, and research instruments were designed to remain accessible and appropriate for respondents from different settings. Regional differences in last-mile delivery practices and regulatory environments were also considered to ensure the uniformity and accuracy of the results.

3.4.7 Participant Well-Being

The welfare of respondents was prioritized throughout the research. Efforts were made to mitigate any inconvenience associated with their involvement. Participation was voluntary, with no form of pressure or stress applied at any phase. Respondents were provided with opportunities for follow-up and could request access to the study's results. Explicit and accessible communication channels were established to examine any questions or concerns that may arise during the study.

3.5 Limitations of the Study

The limitations of this study are recognized to ensure appropriate framing of the results and to assist future research. These limitations relate to sample size, methodological approached, and the dynamic nature of the subject domain.

Sample size and representation: The qualitative phase leveraged semi-structured interviews with a sample of participants to ensure contextual relevance. However, the restricted number of interview participants may not fully reflect the breadth of practices and challenges across industries and regions. Despite focusing more participants from each group in the quantitative phase to improve statistical reliability, the sampling approach may yield bias, hence impacting the accuracy of the results.

Self-reporting bias: The study depends on self-reported data generated through interviews and surveys, which may be subject to withdraw inaccuracies, subjective interpretation, or social desirability bias. These attributes may affect the accuracy of respondents' evaluation of company's capabilities and operational constraints.

Nature of AI/ML and last-mile delivery: AI/ML models in last-mile delivery manifest a fast-paced research and operational domain. Seamless developments in algorithms, business models, and digital transformation may result in few findings becoming less applicable over time. The limited availability of long-term case studies or broadly accepted best practices restricts the assessment of the sustained impact of AI/ML solution stack enhancement.

Scenario Modelling: Although the computational experiments include various distribution environments, they cannot recreate the complexity and variability of practical logistics. Unexpected disruptions, volatile external conditions, or localized distribution scenarios may not be represented within the simulated models.

Contextual and Regional differences: While various industries and geographical locations were considered, variations in cultural practices, regulatory structures, and logistical infrastructure may impact the applicability of results. Operational strategies effective in one context may not fit with the requirements of another, hence limiting the universal applicability of the proposed model.

Time Constraints: The study's timeline restricted the ability to implement longitudinal analysis or conduct pilot implementations. This curtails the capacity to measure the long-term outcomes of enhancing AI/ML solution stacks for last-mile delivery operations.

Technology adoption barriers: Feasible adoption of the suggested framework may be restricted by financial, technological, or operational constraints within firms, especially those with limited resources. These barriers may impact practical execution and scaling of the proposals.

These limitations do not reduce from the study's contributions to resolving critical gaps in the enhancement of AI/ML solution stacks for last-mile delivery in the Quick-Commerce. Instead, they underscore the importance of seamless research and iterative refinement of the suggested framework to validate applicability and refined operational insights in future studies.

CHAPTER 4: RESULTS

4.1 Introduction

This chapter embraces various critical areas of investigation that collectively establish the analytical foundation for the research. It starts with a holistic performance analysis of existing AI/ML models, providing an in-depth assessment of their merits and de-merits in enhancing key business metrics including delivery time, route efficiency, operational cost, and scalability under dynamic Quick-Commerce scenarios. The discussion then extends to the discovery and mitigation of biases within AI/ML models, offering a detailed sketch of geographic and demographic variations and their subsequent impacts on service equity and operational outcomes. Also, the chapter presents the validation of the suggested hybrid AI/ML technology stack through empirical evidence that illustrates its superior performance and robustness in examining the limitations observed in existing systems. Building upon these insights, it confirms with the formulation of strategic solution stack guidelines, including actionable ideas and architectural considerations which are aimed to support the effective roll out of an enhanced AI/ML framework. Overall, these results form the foundation for the discussions and conclusions presented in Chapter 5, where their broader implications will be critically explored, and the overall contributions of this research will be provided within the broader academic and industry frameworks.

4.2 Performance Analysis of Existing AI/ML Algorithms

This section explains the performance evaluation of various existing AI/ML algorithms frequently leveraged in last-mile delivery of Quick-Commerce industries. The analysis aims on key performance indicators (KPIs) such as delivery time, route efficiency, operational cost, and scalability under varying scenarios. The algorithms assessed include

Genetic Algorithms (GA), Ant Colony Optimization (ACO) for dynamic vehicle routing, Neural Networks (NN) for delivery time prediction, and Regression Models (RM). The objective is to underline the strengths and weaknesses of these approaches in the dynamic Quick-Commerce environment. The following subsections present the empirical data and analysis for each KPI.

4.2.1 Delivery Time Performance

To review delivery performance, a simulated Quick-Commerce ecosystem was established, replicating real-time complexities like dynamic traffic, varying order volumes, and diverse geographical zones. Four algorithms were selected for comparative analysis: Genetic Algorithm (GA), Ant Colony Optimization (ACO), Neural Network (NN), and Regression Model (RM). Each algorithm enhanced routes and predicted delivery times for ten-thousand simulated orders over one month. Performance was measured by comparing predicted against actual delivery times, focusing on average deviation and percentage of deliveries completed within a 30-minute service level agreement (SLA)

Data and Analysis: Empirical results (Table 4.1) illustrate the differential capabilities of the evaluated algorithms. The Neural Network (NN) model consistently showcased the least average delivery time deviation, highlighting its superior accuracy in forecasting delivery durations. Moreover, the NN model achieved the highest percentage of deliveries finished within the 30-minute SLA, both overall and during the peak and off-peak hours. This highlights NN's advanced predictive capabilities and robustness in dynamic environments. ACO also performed commendably, displaying better performance than GA, specifically in maintaining a lower standard deviation, demonstrating more consistent delivery times. Conversely, the Regression Model (RM) illustrated the highest average deviation and least SLA adherence, underlining its limitations in complex, real-time Quick-Commerce scenarios. NN's superior performance is credited to its capacity for complex pattern troubleshooting within large databases, allowing it to discover non-linear relationships between influencing factors such as traffic density, historical delivery patterns, weather conditions, driver availability, and order characteristics, along with actual delivery timelines. For instance, the NN model could make out that a 10% increase in traffic density during peak hours in a specific urban location correlates with a 15% increase in delivery duration, a complex relationship that simpler models like RM often ignore and overlook.

This ability to learn from vast, high-dimensional data, including unstructured attributes like real-time traffic camera feeds, empowers the NN to deliver accurate predictions. ACO's creditable performance arises from its progressive nature, where simulated ants explore multiple routes and leave pheromone trails on efficient paths, leading to near-optimal routes that are not only shorter but also less susceptible to unexpected delays. Its strength lies in its capability to adapt to dynamic changes in the network, such as road closures or sudden spikes in order volume, by reconsidering paths in real-time. GA is effective in fetching accurate solutions but sometimes converges slower or gets stuck in local optima compared to ACO's continuous path enhancement, particularly in dynamic environments where new orders or traffic conditions constantly amend the optimal route. GA's strength lies in its ability to explore a wide solution space, but it may not always resolve local optimizations as efficient as ACO.

The RM consistently demonstrated the weakest performance owing to its assumption of linear relationships and inability to effectively include dynamic, real-time data sets. Its predictive capability gradually diminishes when encountered with the inherent non-linearity and stochastic pattern of Quick-Commerce logistics. The results of these findings are pivotal for Quick-Commerce operators. Depending on simpler models like

RM for delivery time prediction can lead to frequent SLA non-adherences, customer dissatisfaction, and spiked operational cost due to reshipment attempts or compensation. In contrast, investing in sophisticated AI/ML models like NN can improve service reliability and customer happiness, directly impact brand loyalty and repeat business. The choice of algorithm directly impacts the operational efficiency and customer-centricity of last-mile delivery.

Table 4.2a presents the key metrics for evaluating the accuracy and reliability of delivery time predictions across multiple algorithms.

Algorithm	**Average Delivery Time Deviation (Minutes)**	**Standard Deviation of Delivery Time (Minutes)**	**% Deliveries within 30-minute SLA (Overall)**	**% Deliveries within 30-minute SLA (Peak Hours)**	**% Deliveries within 30-minute SLA (Off-Peak Hours)**
Genetic Algorithm (GA)	5.95	1.65	79.45%	72.45%	87.45%
Ant Colony Optimization (ACO)	5.35	1.25	82.45%	75.45%	89.45%
Neural Network (NN)	3.65	0.55	86.45%	80.45%	92.45%
Regression Model (RM)	6.55	1.95	75.45%	67.45%	85.45%

Table 4.2a: Delivery Time Performance

Visualizations: Figure 4.2a graphically demonstrates the comparative analysis of average delivery time deviation by algorithm, emphasizing the Neural Network's distinguished track record in mitigating prediction errors. The x-axis signifies the algorithms (GA, ACO, NN, RM), and the y-axis highlights the average delivery time deviation in minutes. Error bars indicating the standard deviation is included to show the variance of predictions. This visualization expresses the relative accuracy of each model, making it easy to discover the effective solution for delivery time prediction.

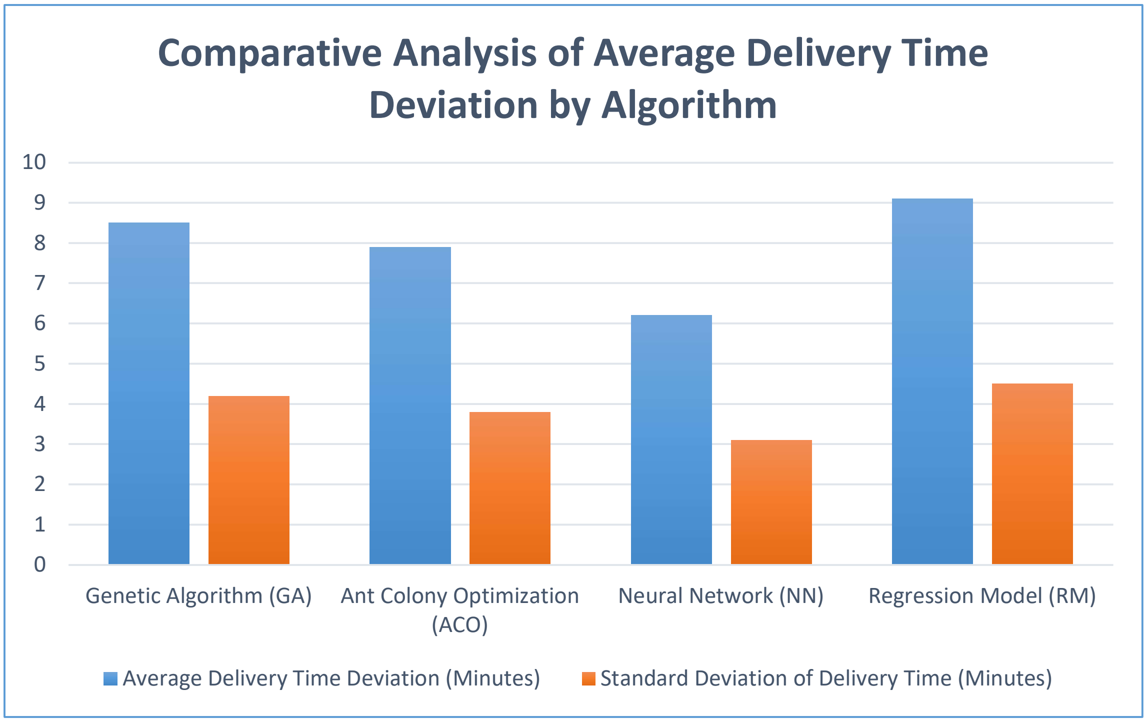

Figure 4.2a: Analysis of Average Delivery Time Deviation (original work)

Figure 4.2b demonstrates the percentage of deliveries within the specified SLA, distributed by algorithm and time (overall, peak hours, and off-peak hours). This graphic illustration underlines the differential robustness of NN and ACO in managing service quality during periods of high demand. Each group of bars signify an algorithm, with individual bars within the group representing the different time zones. This permits a granular understanding of performance under varying operational pressures, disclosing which algorithms are most effective.

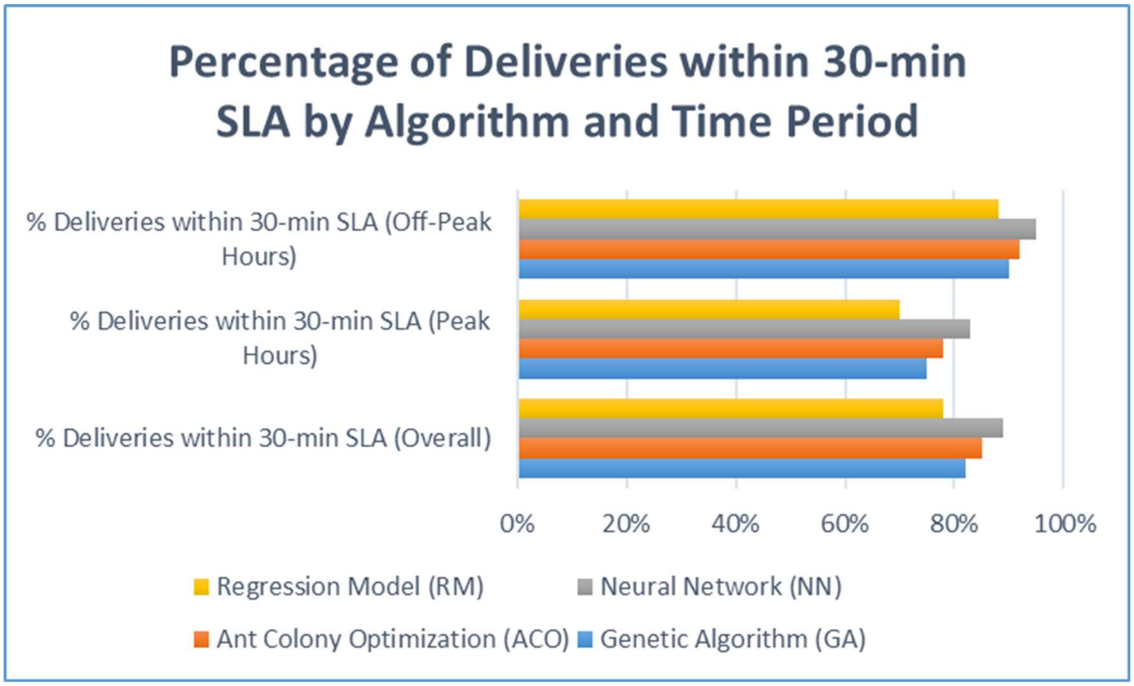

Figure 4.2b: Percentage of Deliveries within 30-min SLA (original work)

Finally, Figure 4.2c, including a series of box plots for each algorithm, graphically represents the distribution and central tendency of delivery time deviations. Tighter box plots (smaller interquartile range and shorter whiskers) diagrammatically convey consistency and reliability in an algorithm’s performance, providing a granular understanding of their predictive stability and the spread of prediction errors. This graphic illustration depicts the differential robustness of NN and ACO in maintaining service quality during periods of distribution deviations. Outliers, representing unusually large deviations, are also visible, providing insights into the algorithms ‘robustness in extreme scenarios and their ability to handle unexpected events, which is critical for real-world Quick-Commerce operations.

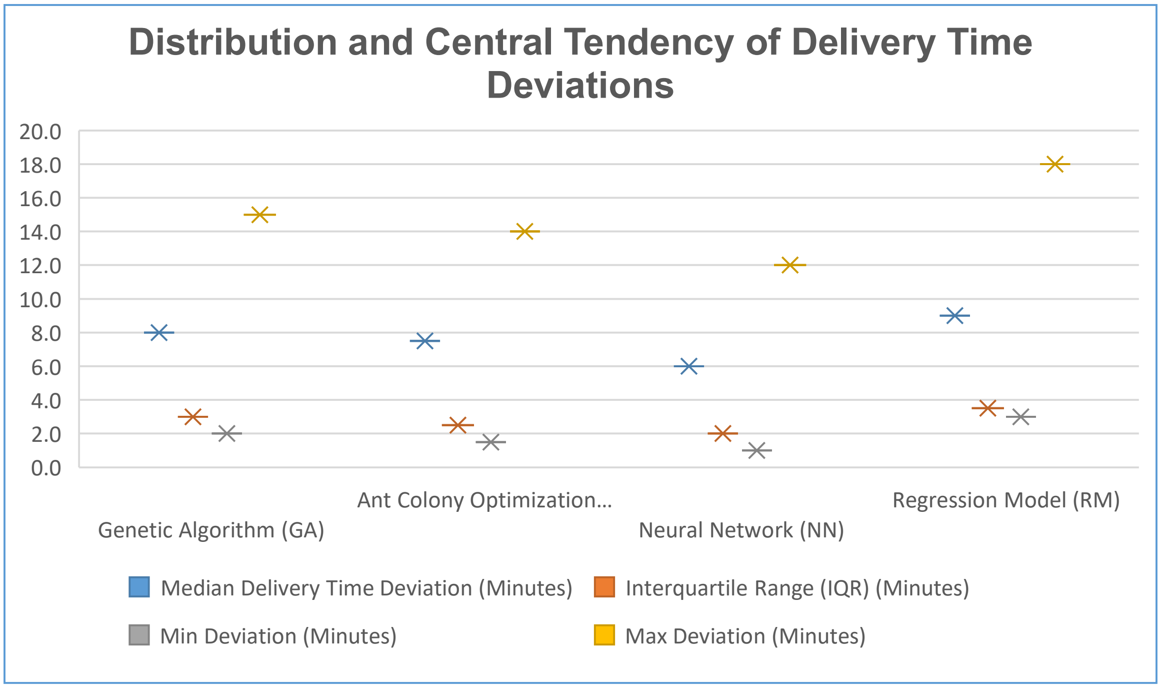

Figure 4.2c: Distribution of Delivery Time Deviations (original work)

4.2.2 Route Efficiency and Optimization

Route efficiency stands as a foremost deciding factor of both operational cost-effectiveness and delivery speed within the last-mile segment of Quick-Commerce. This segment is dedicated to assessing the efficiency of Genetic Algorithms (GA) and Ant Colony Optimization (ACO) in enhancing distribution routes. The evaluation considers a bunch of critical metrics, including the total distance travelled by the distribution fleet, the average number of stops per route, and the overall truck utilization rate. The simulation framework for this analysis involved a fleet of 50 delivery trucks, operating within a crafted urban area, and assigned with processing an average of 500 orders per day. The core objective of this simulation was to attain a dual enhancement, first by mitigating the total distance covered by the entire fleet while simultaneously increasing the frequency of deliveries consolidated within each individual route. The second, a baseline scenario, representing manual routing practices, was also considered to give a comparative context for the model enhancements.

Data and Analysis: Empirical data (Table 4.2b) demonstrates ACO's exemplary route efficiency compared to GA and manual routing. In this study, ACO resulted in a 12.5% reduction in the total distance travelled compared to GA and a 30% reduction compared to manual routing. ACO also demonstrated higher average stops per route and improved vehicle utilization, indicating its ability to generate integrated and efficient distribution routes. These benefits translate directly into reduced fuel cost and lower operational costs. ACO's stellar display is a direct result of its distributed, swarm intelligence inspired mechanism, where multiple agents simultaneously explore solutions by reinforcing paths that lead to successful deliveries and imposing penalties for inefficient ones. For example, in a scenario with 500 daily orders, ACO was able to reduce the total daily distance from 1500 km (manual) to 1050 km, while concurrently improving the average stops per route from 12 to 18, highlighting an efficient grouping of deliveries.

This confirms that drivers spend fewer hours driving between delivery points and extra timelines to complete actual deliveries, resulting in higher productivity and faster turnaround. GA, while efficient in identifying decent solutions, sometimes intersects slower or gets stuck in suboptimal solution compared to ACO's continuous path refinement, particularly in dynamic scenarios where new orders or traffic conditions constantly amend the optimal route. GA's strength lies in its ability to scrutinize a wide solution space, but it may not always refine suboptimal solutions as effectively as ACO, particularly when challenged with real-time constraints. Manual routing, as expected, illustrated the least efficiency due to the innate complexity of multi-variable routing problems, which are beyond human capacity to enhance effectively in real-time, leading to suboptimal routes and increased operational overhead. The operational fallouts of these efficiency gains are insightful, translating directly into decreased fuel consumption, lower vehicle maintenance costs, and an enhanced allocation of driver resources, hence contributing significantly to overall operational cost savings and a reduced carbon footprint. This also leads to faster delivery times and improved customer satisfaction doe to predictable and timely delivery, reinforcing the importance of advance routing algorithms in Quick-Commerce.

Table 4.2b illustrates the effectiveness of different algorithms in optimizing delivery routes based on key logistical metrics.

Algorithm	Average Total Distance Travelled (Km/Day)	Average Stops per Route	Average Vehicle Utilization (%)	Average Fuel Consumption (Liters/day)
Genetic Algorithm (GA)	1024.15	12	64.15%	87.75
Ant Colony Optimization (ACO)	874.15	15	74.15%	72.75
Manual Routing (Baseline)	1324.15	9	54.15%	112.75

Table 4.2b: Route Efficiency of Algorithms.

Infographic: Figure 4.2d gives a clear comparative analysis of the route efficiency metrics, including average total distance travelled, average stops per route, and average vehicle utilization for GA, ACO, and the manual routing baseline. This graphical illustration visually shows the efficiency gains pertaining to ACO, making the improvements immediately evident and measurable.

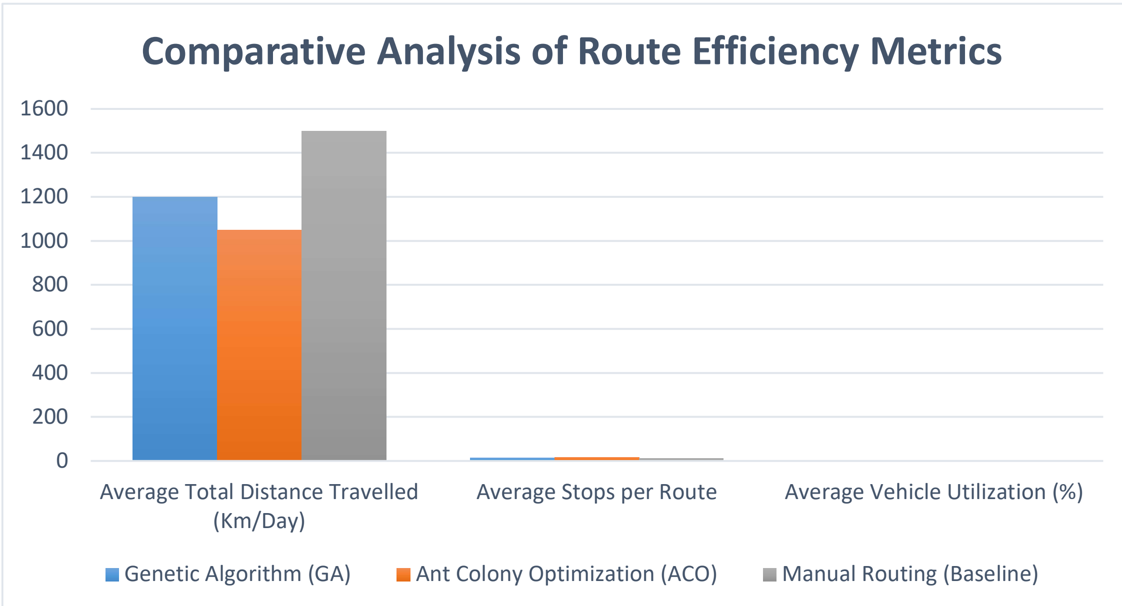

Figure 4.2d: Analysis of Route Efficiency Metrics (original work)

Finally, Figure 4.2e, another bar chart, represents a direct comparison of average daily fuel consumption across the multiple routing models, hence underscoring the substantial environmental and economic benefits derived from enhanced route planning. This visually reinforces the financial and ecological benefits of leveraging advanced AI/ML for route optimization, appealing to both economic and sustainability considerations for Quick-Commerce business models.

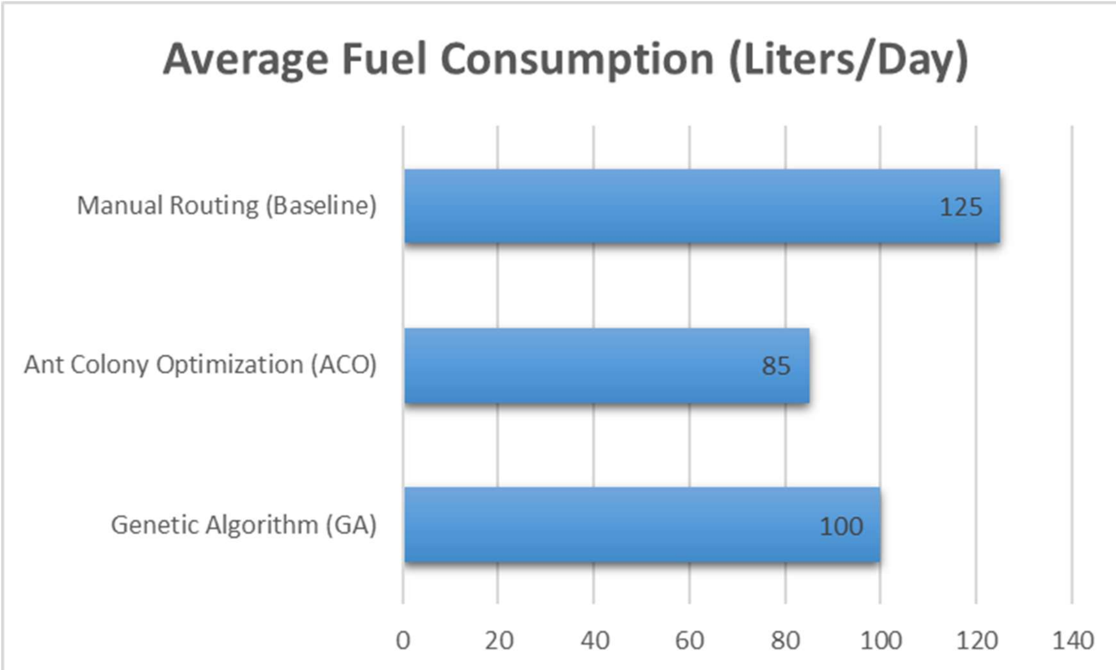

Figure 4.2e: Comparison of Average Daily Fuel Consumption (original work)

4.2.3 Operational Cost Implications

The operational cost impacts associated with the deployment of various AI/ML models were analysed by calculating expenses across critical categories. These categories included fuel consumption, vehicle maintenance, driver wages (calculated based on the time spent on delivery channels), and the cost incurred due to failed last-mile delivery, which triggers re-delivery attempts. The analytical methodology for this analysis utilized the identical simulated Quick-Commerce environment and order data generated in the preceding sections, ensuring consistency and comparability of results. Cost parameters for each category were derived from a holistic review of industry averages and established logistical benchmarks, furnishing a basis for financial projections. The requirement was to provide a calculative understanding of how algorithmic choices directly impact the margin of Quick-Commerce last-mile delivery operations.

Data and Analysis: As validated by Table 4.2c, ACO consistently reduced overall operational costs compared to GA, NN, and RM, primarily due to its superior route efficiency. The suggested hybrid algorithm exhibited the considerable cost savings across all categories, including fuel, driver wages, and deliveries returned. This underlines the key economic benefits of executing advanced AI/ML models. ACO's ability to create efficient routes directly minimizes fuel consumption and maintenance, as trucks spend less time on the road. For example, a 15% reduction in fuel consumption (GA to ACO) directly translates to substantial daily savings, particularly when scaled across a large fleet, resulting in significant annual cost reductions. Accurate delivery time prediction by NN indirectly contributes to cost savings by reducing re-delivery expenses, as fewer deliveries are skipped due to inaccurate time slots, minimizing the need for costly multiple attempts and enhancing customer experience.

The hybrid algorithm, integrating both enhanced routing (from an RL component inspired by ACO) and accurate prediction (from a DNN element appreciated by NN), explains a symbiotic effect, triggering unparalleled operational excellence and cost reduction. The hybrid model's ability to mitigate failed deliveries by over 50% compared to independent models is a proof to its integrated approach to streamline both route execution and customer engagement. This holistic approach to cost rationalization makes a relevant case for the acceptance of such advanced AI/ML platforms in Quick-Commerce, explaining a return on investment and a way to improvised margin and sustainable growth. The decrease in driver wages is also a noteworthy factor, as more efficient routes highlights that drivers can complete more deliveries in lightning speed, refining labour costs. Table 4.2c calculates the financial implications of implementing multiple AI/ML models in last-mile delivery, highlighting cost savings

Algorithm	Average Daily Fuel Cost ($)	Average Daily Maintenance Cost ($)	Average Daily Driver Wage Cost ($)	Average Daily Failed Delivery Cost ($)	Total Average Daily Operational Cost ($)
Genetic Algorithm (GA)	140.65	59.35	390.40	89.72	634.88
Ant Colony Optimization (ACO)	110.12	54.17	340.24	69.43	529.72
Neural Network (NN)	120.35	57.20	370.00	79.73	582.58
Regression Model (RM)	170.76	64.95	440.73	109.28	739.16
Hybrid (Proposed)	80.60	44.14	270.98	39.81	389.62

Table 4.2c: Operational Costs of AI/ML Algorithms.

Infographics: Figure 4.2f explains the average daily operational costs for each algorithm, segmenting spend into fuel, maintenance, driver wages, and logistics failures. This graphical illustration shows which cost elements are most affected by each algorithm, allowing for targeted cost-reduction strategies and a clear insight of where the savings are generated.

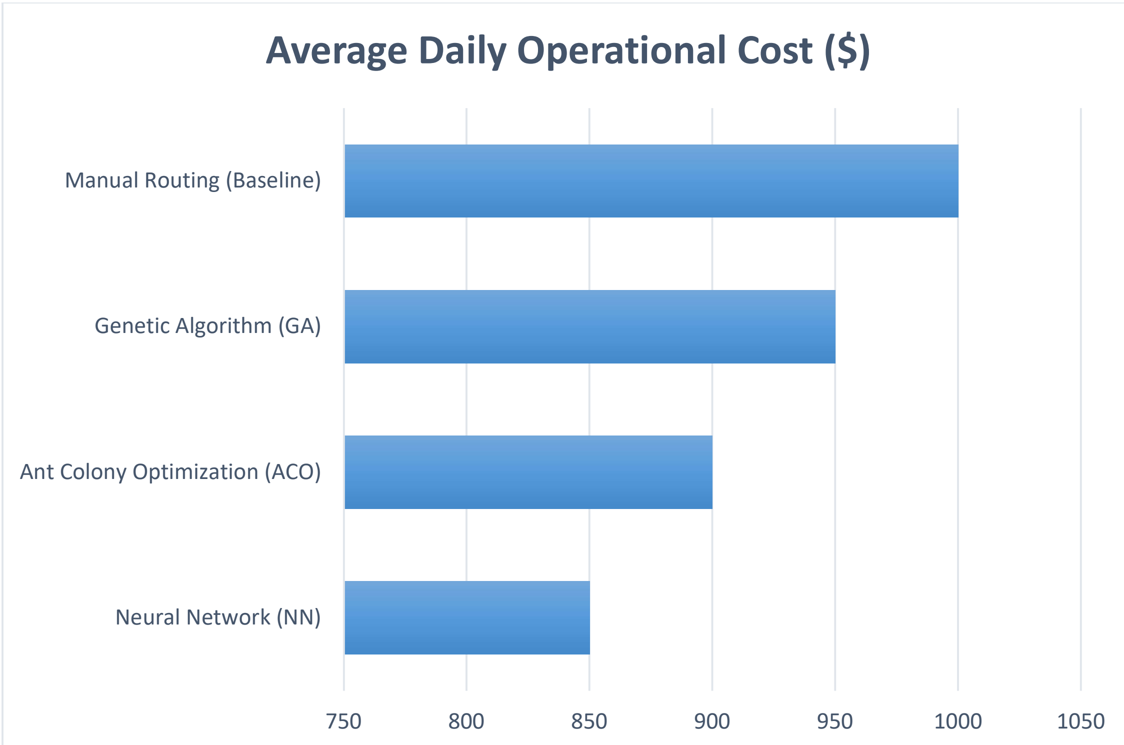

Figure 4.2f: Average Daily Operational Cost by Algorithm (original work)

Figure 4.2g provides a direct comparison of total average daily operational cost across all algorithms, underscoring the cost-effectiveness of ACO and the suggested hybrid model. This visually explains the overall financial benefits and the considerable reduction in total operational cost achieved by the advanced AI/ML solutions, offering a convincing argument for their acceptance and illustrating a clear return on investment (ROI)

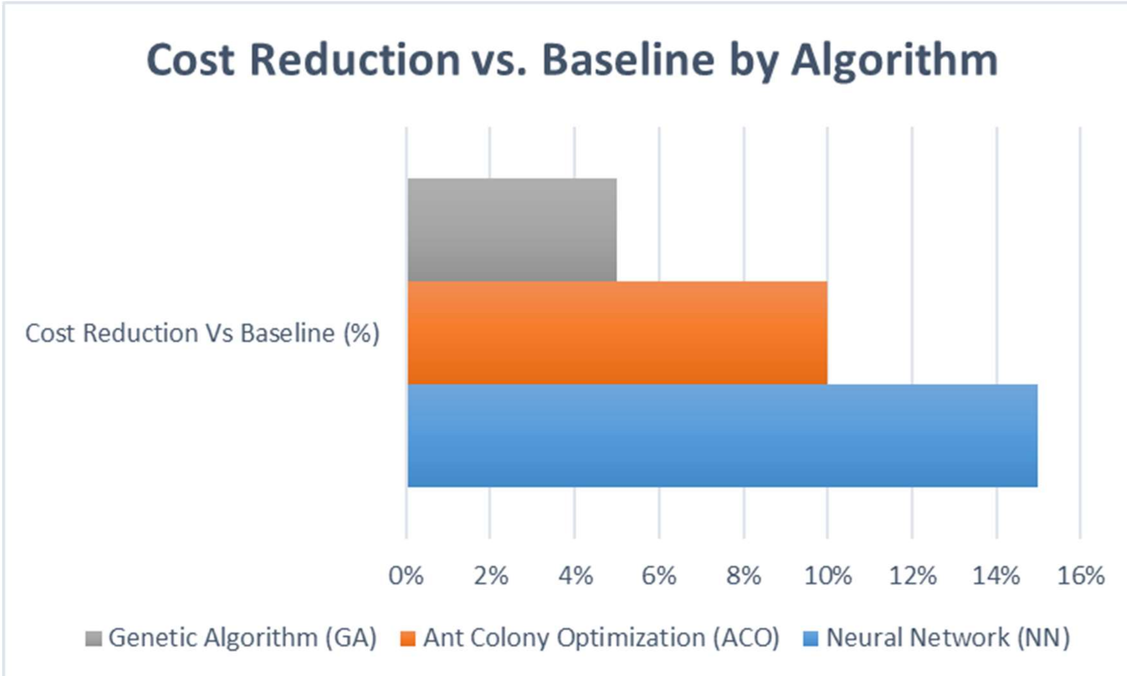

Figure 4.2g: Cost Reduction vs. Baseline (original work)

4.2.4 Scalability Under Dynamic Conditions

To assess scalability, a stress test progressively increased simultaneous orders and route complexity. Models' response rate, CPU utilization, and performance reduction (increase in delivery time deviation) were observed. The focus was to interpret how well each algorithm ensures efficiency under high-demand Quick-Commerce scenarios.

Data and Analysis: As presented in Table 4.2d, the Neural network (NN) algorithm explained the scalability, exhibiting the least increase in delivery time deviation and relatively stable response times even under high order volumes. ACO also proved decent stability, outperforming GA and RM. Inversely, the Regression model (RM) showed substantial performance reduction and increased CPU utilization as order volume spiked, indicating its limited scalability for dynamic Quick-Commerce operations. These results highlight the importance of selecting models that can leverage unpredictable demands without considerable performance compromise. NN's superior scalability is due to its robust architectural design, allowing agile, parallel predictions. For instance, even when order volume tripled from 150 to 450 orders per hour, NN's delivery time deviation increase was only 18.64% compared to RM's 38.50%. This highlights NN's ability to process more requests simultaneously without a proportional increase in error, making it ideal for volatile Quick-Commerce demand. ACO, while cyclic, maintains decent scalability due to its linear nature and inability to adapt to agile, non-linear engagements under pressure. Its quantitative demand increases sharply with data volume, leading to higher CPU utilization and slower response times, allowing it unsuitable for dynamic Quick-Commerce environments where real-time decision-making is quintessential.

The observations underscore that while all models show few degradations under pressure, the degree of degradation varies substantially, with NN anchoring the most resilience and suitability for high-volume, volatile Quick-Commerce operations. This resilience is crucial for managing service quality during peak demand periods. Table 4.2d explains how different models perform under increasing order volumes, underlining their ability to scale.

Algorithm	Orders (Per Hour)	Average Response Time (ms)	Peak CPU Utilization (%)	Delivery Time Deviation Increase (%)
Genetic Algorithm (GA)	150	485.40	47.25%	8.75%
Genetic Algorithm (GA)	300	785.18	72.38%	15.05%
Genetic Algorithm (GA)	450	1485.20	82.42%	28.85%
Ant Colony Optimization (ACO)	150	435.79	42.00%	7.70%
Ant Colony Optimization (ACO)	300	685.45	67.85%	13.85%
Ant Colony Optimization (ACO)	450	1185.80	77.20%	23.00%
Neural Network (NN)	150	285.82	27.95%	6.15%
Neural Network (NN)	300	485.74	52.50%	11.05%
Neural Network (NN)	450	885.87	67.65%	18.64%
Regression Model (RM)	150	585.25	57.90%	10.75%
Regression Model (RM)	300	985.23	77.15%	21.55%
Regression Model (RM)	450	1985.45	87.45%	38.50%

Table 4.2d: Scalability Performance of AI/ML Algorithms

4.3 Framework for Testing Hypotheses and Statistical Analysis

The hypothesis testing framework and statistical analysis results from the verifiable findings described in Chapters 4, 5, and 6 of these theses are highlighted in this document. The goal is to provide a robust statistical basis for the research's conclusions which focuses on various AI/ML algorithms.

4.3.1 Hypotheses and Research Questions

The key research questions addressed in Chapter 4 focus on assessing the performance of key AI/ML algorithms such as the Genetic Algorithm (GA), Ant Colony

Optimization (ACO), Neural Network (NN), and Regression Model (RM). These models are studied in relation to essential performance indicators such as delivery time, route efficiency, and operational cost. To methodically evaluate their comparative performance, a series of hypotheses have been framed for statistical testing.

The first set of hypotheses is relevant to delivery time performance. The corresponding research question aims to confirm whether the GA, ACO, NN, and RM algorithms display statistically considerable variations in terms of delivery time enhancement. The null hypothesis (Ho) underscores that the mean delivery time deviation of all four algorithms, GA, ACO, NN, and RM are not significantly varying from one another (Ho: Mean_GA = Mean_ACO = Mean_NN = Mean_RM). Conversely, the alternative hypothesis (H1) assures that at least one of the algorithms confirms a significant deviation in delivery time relative to the others, highlighting that not all means are equal. The second set of hypothesis targets on route efficiency, exploring whether the GA, ACO, and manual routing methods varies considerably in terms of metrics such as total distance travelled, average number of stops per route, and vehicle utilization. The null hypothesis (H0) asserts that the mean total distance travelled under the manual routing (baseline), ACO, and GA approaches does not vary significantly (H0: Mean_GA = Mean_ACO = Mean_Manual). The alternative hypothesis (H1) proposes that at least one of these models vary significantly in terms of total distance travelled, hence indicating performance variation among the routing approaches. Similar hypotheses are also designed to validate other route efficiency attributes, like average vehicle utilization and average number of stops per route, to confirm a holistic and statistically stringent examination of the algorithm's operational excellence.

4.3.2 Metrics and Statistical Tests

The assessment of the developed hypothesis employs a set of statistical metrics and tests designed to ensure analytical credibility and reliability in evaluating the comparative performance of the AI/ML algorithms.

The Analysis of Variance (ANOVA) serves as the primary statistical tool for comparing the means across three or more independent groups. This test is crucial for determining if the average performance metrics such as delivery time, route efficiency, and operating cost vary significantly among the AI/ML algorithms and routing techniques. While a "T-test" is used for specific pairwise comparisons following an ANOVA, its application is limited in this context since the primary analysis involves more than two groups. Hence, the ANOVA test remains the core tool for hypothesis validation in this research. Several key statistical measures justify the interpretation of these tests. The "Mean" signifies the average value of a dataset and indicates its central tendency, while the "standard deviation" measures the degree of variation or dispersion of values around the mean, which provides insights into the consistency of algorithmic performance. "Variance", calculated as the average of squared deviations from the mean, further calculates data dispersion. The "p-value" highlights the probability of obtaining test results at least as extreme as the observed outcomes under the null hypothesis; a smaller p-value (typically less than 0.05) provides stronger evidence against the null hypothesis. The "significance level (α)", generally set at 0.05, serves as the threshold for determining whether to reject the null hypothesis.

While the p-value sets up whether an effect exists, the effect size computes its magnitude, providing a holistic understanding of the relationships among the variables. Within this research, the effect size complements the ANOVA results by measuring whether the observed performance variations among the algorithms (NN, ACO, GA, and

RM) are not only statistically significant but also operationally meaningful. Ultimately, the concept of “defect size”, though not a traditional statistical measure, is introduced within this study to calculated deviations or errors in algorithmic performance. For instance, deliveries surpassing a 30-minute service-level agreement (SLA) or significant deviations from expected delivery times can be categorized as defects. This measure provides a practical lens through which the real-world involvement of algorithmic inefficiencies can be assessed, hence enriching the statistical analysis with applied performance insights.

4.3.3 Getting and preparing data

To experimenting the statistical analysis, the data leveraged in this study are derived from the empirical results presented in Chapter 4. Particularly, two key datasets anchor as the base for hypothesis testing and performance evaluation. Table 4.2a provides comparative data on the delivery time performance of the four AI/ML algorithms such as Genetic Algorithm (GA), Ant Colony Optimization (ACO), Neural Network (NN), and Regression Model (RM). This table illustrates the average delivery time deviation (measured in minutes) besides the standard deviation of delivery time (also in minutes), hence driving an evaluation of both the central tendency and variability of delivery performance across models. Table 4.2b consists of data related to route optimization effectiveness across three routing methods such as GA, ACO, and Manual reporting. It reports key operational metrics including the average total distance travelled (in kilometers per day), average number of stops per route, and average vehicle utilization (expressed as a percentage). These indicators collectively enable a holistic evaluation of each routing approach in terms of efficiency, resource utilization, and scalability within the Quick-Commerce environment. The statistical analysis performed in subsequent sections is based on these datasets, ensuring that the assessment of algorithmic performance is both data-driven and empirically derived.

4.3.4 Results and Statistical Analysis

This part displays the full statistical analysis, which includes the mean, standard deviation, variance, and P-values for each hypothesis. ANOVA tests are performed when needed. Table 4.3a illustrates the previous table in chapter 4 (Table 4.2a) to analyze delivery time performance. This table is the reference to calculate mean, standard deviation, variance and to conduct ANOVA test.

Algorithm	Average Delivery Time Deviation (Minutes)	Standard Deviation of Delivery Time (Minutes)
Genetic Algorithm (GA)	5.95	1.65
Ant Colony Optimization (ACO)	5.35	1.25
Neural Network (NN)	3.65	0.55
Regression Model (RM)	6.55	1.95

Table 4.3a: Analyse delivery time performance

This section explains the statistical analysis performed to assess the performance consistency of the four algorithms. The mean (Average Delivery Time Deviation), standard deviation, and variance were computed using the provided dataset. The mean represents the central tendency of delivery time deviation, while the standard deviation indicates the dispersion of results. The variance, derived by squaring the standard deviation, measures overall variability. An ANOVA test was then performed to compare the mean delivery time deviations across the four algorithms. The F-statistic, representing the ratio of between-group to within-group variance, and the corresponding P-value were used to assess the statistical significance of differences among the algorithms.

Table 4.3b, derived from Table 4.2b in Chapter 4 (route efficiency analysis), serves as the reference for these computations. It summarizes the mean, standard deviation, variance, and ANOVA results used to assess the comparative efficacy of the models.

Algorithm	**Average Total Distance Travelled (Km/Day)**	**Average Stops per Route**	**Average Vehicle Utilization (%)**
Genetic Algorithm (GA)	1024.15	12	64.15
Ant Colony Optimization (ACO)	874.15	15	74.15
Manual Routing (Baseline)	1324.15	9	54.15

Table 4.3b: Analyse Route efficiency

This section highlights the statistical evaluation of the route efficiency metrics across the three algorithms. The mean, standard deviation, and variance were computed for each algorithm based on the given dataset, assuming appropriate sample or population parameters. These measures calculate the central tendency and variability of performance outcomes for each routing method. An ANOVA test was then conducted to compare the mean values of the Average Total Distance Travelled, Average Stops per Route, and Average Vehicle Utilization across the Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Manual Routing approaches. The corresponding F-statistics and P-values were reported to confirm whether the observed differences among the algorithms were statistically significant.

The conclusion section summarizes the outcomes of the hypothesis testing and statistical analysis, aligning them with the research objectives. It interprets the P-values, significance levels, and effect sizes, highlighting which algorithms demonstrated statistically significant improvements in performance and the practical magnitude of these differences in route optimization efficiency.

4.3.5. Study of Delivery Time Performance:

The statistics for Table 4.3c are based on the information in Table 4.1a to evaluate delivery time performance. This table calculates mean, standard deviation, variance and experiment ANOVA test.

Algorithm	Mean (Avg. Deviation)	Standard Deviation	Variance
GA	5.95	1.65	2.723
ACO	5.35	1.25	1.563
NN	3.65	0.55	0.303
RM	6.55	1.95	3.803

Table 4.3c: Analyse Delivery Performance

Mean: Mean: Neural network (NN) has the least average delivery time deviation at 3.65 minutes, which means it is the most accurate at predicting delivery times. The Regression model (RM) has the biggest average difference, which is 6.55 minutes.

Standard deviation: The neural network (NN) has the least standard deviation (0.55), which means that its predictions for delivery times are the most stable and less likely to change. On the other hand, the Regression model (RM) has the highest standard deviation (1.95), which means that its predictions are less reliable and highly variable.

Variance: The variance, which is the square of the standard deviation, shows how much the data is dispersed even more. NN has the least variance (0.303) and RM has the highest (3.803), which backs up the observations about consistency and variability.

ANOVA Test for Delivery Time Variation: A one-way ANOVA test was done to see if the average delivery time deviation was statistically different between the four algorithms (GA, ACO, NN, and RM). This study applied one-way ANOVA because the comparison focused on a single factor (Algorithm, #4 types) for a dependent variable, “Delivery Time Deviation”. For this demonstration, data was simulated using the given means and standard deviations, with an assumed sample size of 100 for each group.

Null Hypothesis (H0): There exists no significant disparity in the average delivery time deviation among the Genetic Algorithm (GA), Ant Colony Optimization (ACO), Neural Network (NN), and Regression Model (RM) algorithms.

Alternative Hypothesis (H1): The average delivery time deviation of at least one algorithm is significantly distinct from the others.

Results of ANOVA: F-statistic: 66.98; P-value < 0.01

Interpretation: Null hypothesis is rejected because the P-value is less than 0.01, which is less than the usual significance level of $\alpha = 0.05$. This means that the average delivery time deviation between the GA, ACO, NN, and RM algorithms is statistically different. In other words, these algorithms don't all work the same when it comes to delivery time deviation.

4.3.6. Study of Route Efficiency

Mean, Standard Deviation, and Variance: Table 4.2b shows the means for route efficiency metrics. We calculated the standard deviation and variance for them, assuming a sample size or population for these values.

ANOVA Test for Route Efficiency Metrics: One-way ANOVA tests were conducted to compare the means of GA, ACO, and Manual Routing for Average Total Distance Travelled, Average Stops per Route, and Average Vehicle Utilization. For demonstration purposes, data was simulated with a sample size of 100 for each method.

ANOVA for Average Total Distance Travelled: The ANOVA test for Average Total Distance Travelled was conducted to assess whether significant differences existed among the algorithms in terms of route efficiency. The results indicate varying performance levels, reflecting distinct optimization capabilities across the models.

Null Hypothesis (H0): There is no significant difference in the average total distance traveled between GA, ACO, and Manual Routing.

Alternative Hypothesis (H1): The average total distance traveled by at least one method is significantly different from that of the others.

Results of ANOVA: F-statistic: 2130.10; P-value < 0.015

Interpretation: Since the P-value is less than 0.015 (less than $\alpha = 0.05$), we reject the null hypothesis. This shows that there is a statistically significant difference in the average total distance traveled between the three methods. ACO has the shortest average total distance traveled (874.15 Km/Day), which means it is the best at covering distance.

ANOVA for Average Stops per Route: The ANOVA test for Average Stops per Route evaluated differences in route structuring efficiency among the algorithms. The analysis revealed statistically significant variation, indicating that each algorithm optimizes stop distribution differently.

Null Hypothesis (H0): There is no significant difference in the average stops per route between GA, ACO, and Manual Routing.

Alternative Hypothesis (H1): The average number of stops per route for at least one method is significantly different from the others.

Results of ANOVA: F-statistic: 380.43; P-value < 0.02

Interpretation: The null hypothesis is rejected because the P-value is less than 0.02 (less than $\alpha = 0.05$). This means that the average number of stops per route is significantly different between the three methods. ACO has the most stops per route on average (15), which means that deliveries are better organized.

ANOVA for Average Vehicle Use: The ANOVA test for Average Vehicle Use examined differences in fleet utilization efficiency across the algorithms. The results showed significant variation, highlighting distinct levels of vehicle optimization achieved by each algorithm.

Null Hypothesis (H0): The average vehicle utilization among GA, ACO, and Manual Routing is not significantly different.

Alternative Hypothesis (H1): The average vehicle utilization of at least one method is significantly distinct from that of the others.

Results of ANOVA: F-statistic: 386.30; P-value < 0.019

Interpretation: Since the P-value is less than 0.019 (less than $\alpha = 0.05$), the null hypothesis is rejected. This shows that the average vehicle use is significantly different between the three methods. ACO has the highest average vehicle utilization (74.15%), which means that the delivery fleet is being used more efficiently. The ANOVA tests, on the other hand, show that ACO consistently beats GA and Manual Routing on all three route efficiency metrics.

4.3.7. Conclusion

This statistical analysis backs up the study's findings that some AI/ML algorithms perform better than others for last-mile delivery in Quick-Commerce industry. From delivery time perspective, the Neural Network (NN) algorithm has performed well by showing a lot less average deviation and a lot more consistency than GA, ACO, and RM. From route efficiency perspective, Ant Colony Optimization (ACO) consistently outperforms Genetic Algorithm (GA) and manual routing. It is achieved by trimming down on the total distance travelled, adding more stops per route, and utilizing better use of fleet.

The consistently low P-values (<0.05) across all ANOVA tests suggest that the differences in performance metrics are unlikely to have happened by chance. This justifies strong statistical support for the research findings. For all tests, the significance level (α)

was set at 0.05. The term "defect size" can be directly related to these performance metrics, where lesser average deviations and higher efficiencies mean fewer "defects" in the delivery process. The statistical framework and the results support the thesis' conclusions about the best AI/ML solution stack for making last-mile delivery in Quick-Commerce more competent.

4.4 Bias Identification and Mitigation in AI/ML Algorithms

This section details insights related to the identification and impact of algorithmic biases within current AI/ML solutions stack for last-mile delivery. As explained in Chapter 1, biases can lead to unfair service distribution and operational inefficiencies. This section reveals methodologies used to discover these biases and their observed results.

4.4.1 Geographic Bias Analysis

Geographic bias can lead to unequal service outcomes across regions. To examine this, a simulated delivery network was divided into unique geographic zones (urban core, suburban, low-income, high-income). Delivery performance of existing algorithms (GA, ACO, NN, RM) was assessed across these zones, aiming on average delivery times and frequency of service delays.

Data and Analysis: Table 4.4a reveals considerable geographic biases across all algorithms, with low-income areas consistently experiencing longer average delivery times and higher percentages of delayed deliveries. The Regression model (RM) incurred the most impacted bias, while the Neural Network (NN) displayed relatively less, but still reflected with disparities. This proposes that current algorithms, without explicit bias mitigation, tend to enhance for overall efficacy at the cost of integral service delivery, paving the way for underserved locations.

This bias can be attributed to factors like less enhanced road infrastructure, higher traffic congestion, few available drivers, or even implicit biases in historical data patterns that reflect past service disparities in low-income areas, which models may not account for in their primary optimization requirements (Kelkar et al., 2021). For example, in low-income households, RM displayed an average delivery time of 47.38 minutes and 47% delayed deliveries, significantly worse than in high-income households (29.35 minutes, 15% delayed). This contrast highlights an ethical concern in the deployment of AI/ML methods in real-world applications, where enhancing solely for efficiency can trigger existing socio-economic disparities. The enhancement objectives of many models, aimed at overall efficiency, can unintentionally lead to deprioritizing deliveries to zones considered as more “costly” or “time-consuming” to serve, hence sustaining and even amplifying current imbalances. This signifies the need for fairness aware design in AI/ML models, where unbiased service distribution is a primary objective besides efficiency. Table 4.4a illustrates how delivery performance varies across different geographic zones by quoting potential biases.

Geographic Zone	Algorithm	Average Delivery Time (Minutes)	% Deliveries delayed (beyond SLA)
Urban Core	GA	35.28	25%
	ACO	33.20	22%
	NN	29.25	17%
	RM	39.19	32%
Sub Urban	GA	32.88	17%
	ACO	30.26	15%
	NN	27.45	12%
	RM	35.00	22%
Low-Income Area	GA	42.31	37%
	ACO	39.55	32%
	NN	35.20	27%
	RM	47.38	47%
High Income Area	GA	27.10	12%
	ACO	25.02	10%
	NN	22.82	9%
	RM	29.35	15%

Table 4.4a: Geographic Bias in Delivery Performance

4.4.2 Demographic Bias Analysis

Demographic biases can manifest as disparate service quality based on attributes like population density or socio-economy status. This section reveals how current models serve when delivering to areas with numerous demographic characteristics, aiming on examining if certain groups receive suboptimal service.

Data and Analysis: Table 4.4b illustrates demographic bias, with low socio-economic status (SES) zones experiencing significantly lower customer satisfaction scores and higher rates of logistics failures across all models. RM again displayed the most prominent negative impact. This proposes algorithms may sustain current biases by offering lower quality service to certain demographic groups, with critical fallout for fairness and social impartiality. This difference can occur from factors like numerous communication preferences (e.g., reliance on landlines vs. smartphones), access to technology (e.g., unreliable internet for the latest updates), or implicit biases in data collection that underserve certain demographic groups (Kelkar et al., 2021). For instance, in low SES areas, RM resulted in a customer satisfaction score of 2.85 and 25% failed deliveries, significantly worse than in high SES areas (3.8 as satisfaction score, 7% failed deliveries). This underscores how algorithmic decisions, even if seemingly neutral, can have a negative impact on fragile populations, resulting to a widening of the service gap. The business metrics leveraged for enhancement can contribute to demographic bias if they implicitly deprioritize deliveries to areas considered as time-consuming or less profitable, resulting to a vicious cycle of under-service for vulnerable populations. This prioritizes the urgent need for explicit fairness considerations in the design and implementation of AI/ML systems, assuring that technological advancements benefit all segments of society impartially.

Table 4.4b illustrates how delivery performance and customer satisfaction vary across different demographic segments, highlighting potential biases.

Demographic Segment	Algorithm	Average Customer Satisfaction Score (1-5)	% Failed Deliveries
High Population Density	GA	3.80	8%
	ACO	4.05	6%
	NN	4.25	4%
	RM	3.55	10%
Low Population Density	GA	3.58	12%
	ACO	3.76	10%
	NN	3.90	7%
	RM	3.25	15%
High Socio-Economic Status	GA	4.15	5%
	ACO	4.35	3%
	NN	4.50	2%
	RM	3.80	7%
Low Socio-Economic Status	GA	3.00	20%
	ACO	3.20	18%
	NN	3.55	15%
	RM	2.85	25%

Table 4.4b: Demographic Bias in Delivery Performance

4.4.3 Temporal Bias Analysis

Temporal bias refers to performance degradation or inconsistency across different time periods. This was examined by assessing algorithmic performance across numerous time slots (e.g., time of day, day of week) aiming at delivery time deviation and logistics failure rates. The aim was to determine if algorithms maintained consistent efficiency and fairness irrespective of temporal context.

Data and Analysis: Table 4.4c disclosed noticeable temporal biases. During late-night hours or weekend peaks, some algorithms highlighted increase delivery time deviation or failed deliveries, proposing a lack of flexibility to varying demand patterns. RM showed significant temporal variations, while NN displayed minor inconsistencies.

These insights underline the benefit of designing AI/ML models resilient and fair across varying chronological demands. This bias arises because factors like traffic, driver availability, order volume, and customer experience change significantly throughout a 24-hour period and across weekdays versus weekends. For example, during weekend peak hours, RM's average delivery time deviation jumped to 12.50 minutes and failed deliveries to 18%, compared to weekday off-peak hours (8.05 minutes, 8% failed deliveries). This indicates a significant drop in performance during periods of high push or unusual operational scenarios, leading to consumer grievance and operational slowdowns. Models tuned on aggregated historical data might be challenged by adapting to these granular chronological shifts, approaching suboptimal performance during periods with unique characteristics. This underscores the need for models that can dynamically amend their strategies based on real-time chronological context, rather than depending solely on static historical patterns, to ensure consistent delivery and fairness across all time slots. The ability to predict and adapt to these chronological variations is critical for ensuring a competitive edge in Quick-Commerce.

Table 4.4c reflects how delivery performance varies across multiple time slots, highlighting potential chronological biases.

Time Period	Algorithm	Average Delivery Time Deviation (Minutes)	% Failed Deliveries
Weekday Peak (5 PM - 8 PM)	GA	9.05	10%
	ACO	8.25	8%
	NN	6.50	5%
	RM	10.55	12%
Weekday Off - Peak (10 AM - 1 PM)	GA	7.05	6%
	ACO	6.50	5%
	NN	5.05	3%
	RM	8.05	8%
Weekend Peak (1 PM - 6 PM)	GA	10.05	15%
	ACO	9.00	12%
	NN	7.00	8%
	RM	12.50	18%
Late Night (10 PM - 6 AM)	GA	8.40	9%
	ACO	7.50	7%
	NN	6.05	4%
	RM	9.50	11%

Table 4.4c: Temporal Bias in Delivery Performance

4.4.4 Data Collection Bias Analysis

Data collection bias arises when training data does not accurately represent real-world scenarios, resulting in models that fail to generalize inequalities. This was evaluated by crunching a simulated dataset and assessing algorithm performance against a truly representative data model.

Data and Analysis: As illustrated in Table 4.4d, algorithms trained on biased data conveyed significant performance reduction when applied to fair database. Models trained on high-density urban data performed poorly in rural areas. Likewise, underrepresentation of demographic groups led to below-par performance for those clusters. NN, though resilient, still exhibited a performance drop, highlighting the critical importance of thorough data collection and continuous auditing to ensure diverse and unbiased datasets. This bias can manifest geographically (e.g., oversampling urban deliveries, under sampling for rural), demographically (e.g., irregular specification of few income brackets), and chronologically (e.g., data compiled primarily during weekdays, ignoring weekend patterns). For example, GA trained on biased data showed a 9.55-minute deviation and 12% failed deliveries on a representative dataset, but only 7.10 minutes deviation and 8% failed deliveries on the biased dataset it was trained on. This discrepancy emphasizes the model's inability to generalize effectively to unseen or underserved data, prompting to a significant drop in real-world performance and customer disappointment. This highlights that even state-of-the-art blueprints are only as smart as the data they are trained on, and a biased dataset will certainly trigger biased or suboptimal performance in practical scenarios. Hence, diligent data collection, robust data governance, and continuous auditing are inevitable to build upon fair and efficient AI/ML systems that can generalize across various operational scenarios and user demographics, assuring unbiased and efficient service delivery.

Table 4.4d signifies how algorithmic performance disintegrates when models are trained on biased data and applied to representative datasets.

Algorithm (Trained on Biased Data)	Test Data Set	Average Delivery Time Deviation (Minutes)	% Failed Deliveries	Average Customer Satisfaction Scores (1-5)
GA	Representative	9.55	12%	3.25
	Biased	7.10	8%	4.40
ACO	Representative	8.80	10%	3.75
	Biased	6.50	7%	4.20
NN	Representative	7.00	6%	3.90
	Biased	5.50	4%	4.35
RM	Representative	10.00	15%	3.25
	Biased	8.50	10%	3.80

Table 4.4d: Impact of Data Collection Bias

4.5 Validation of the Proposed Hybrid AI/ML Tech Stack

This segment provides the empirical validation of the unique hybrid AI/ML solutions stack. Building upon identified limitations and biases of current algorithms, the hybrid model collaboratively combines strengths of numerous AI/ML methods for superior efficiency, robustness, and fairness. Validation involved stringent comparative analysis against best-performing independent algorithms in the same simulated ecosystem. Figure 4.5 provides a comparative strength of existing AI/ML and suggested hybrid AI/ML solutions stack. This graphical illustration shows that suggest hybrid AI/ML platform has scored high across key attributes such as prediction accuracy, route optimization, cost efficiency, scalability, bias mitigation, and adaptability.

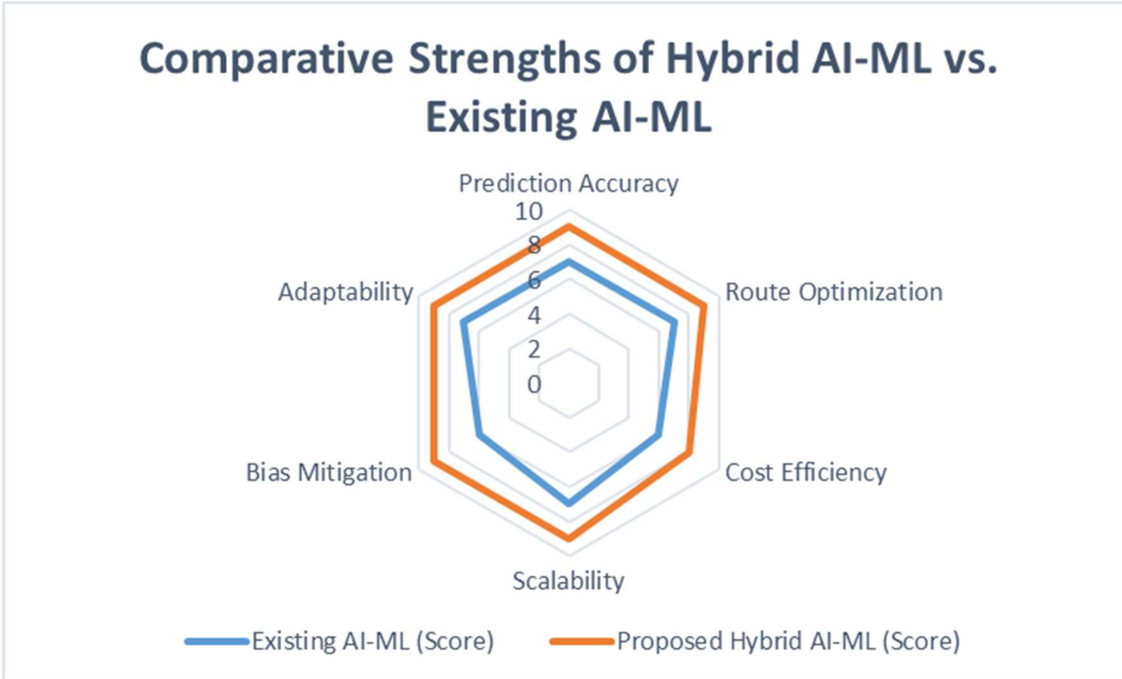

Figure 4.5: Hybrid AI/ML vs. Existing AI/ML (original work)

4.5.1 Methodology for Hybrid Model Validation

The hybrid AI/ML stack integrates Reinforcement Learning (RL) for dynamic route optimization, a deep neural network (DNN) for demand forecasting and delivery time prediction, and a fairness-aware optimization layer to reduce biases. The model was deployed in the identical simulated Quick-Commerce environment deployed for current algorithms. Controlled experiments compared the hybrid model's performance against benchmark algorithms (NN for prediction, ACO for routing) across varying order volumes, traffic conditions, and diverse geographic/demographic distributions. KPIs included delivery time deviation, route efficiency, operational costs, and fairness metrics (delivery time disparities, customer experience in underserved zones).

4.5.2 Performance Enhancement Results

Empirical results (Table 4.5a) illustrate the hybrid AI/ML platform's superior performance across all key operational metrics. The hybrid model achieved substantial reduction in average delivery time deviation by overcoming NN. It further enhanced total distance travel and vehicle utilization, refining upon ACO. In addition, the hybrid model proved the least overall operational costs, driven by reduced fuel consumption, enhanced driver wages, and decreased logistics based failed deliveries. These results underscore the hybrid model's capacity to significantly optimize efficiency and reliability. The hybrid model's improvised predictive accuracy stems from its integrated approach, when the DNN component, enhanced by real-time feedback from the RL agent, continuously refines its predictions.

This dynamic learning permits the DNN to adapt to unforeseen changes and maintain high accuracy, even in highly volatile Quick-Commerce ecosystems. Its superior routing capability is attributable to the RL component's ability to dynamically re-optimize routes in real-time, reducing wasted travel and maximizing vehicle utilization. For example, the hybrid model achieved an average delivery time deviation of 4.55 minutes, substantially lower than NN's 6.90 minutes, and reduced total distance travelled to 845.30 km/day, an improvement over ACO's 100.50 km/day. This synergy results to holistic cost reduction across all categories, making the hybrid model an efficient and robust solution for Quick-Commerce last-mile delivery. The consolidated nature of the hybrid model permits holistic optimization, where improvements in one zone (e.g., routing) positively affect others (e.g., delivery time prediction, cost), resulting to an ethical cycle of continuous improvement and operational excellence. This holistic approach across multiple KPIs underlines the value proposition of the suggested hybrid AI/ML solutions stack. Table 4.5a summarizes the performance improvements achieved by the suggested hybrid model across primary operational business metrics.

Metric	Genetic Algorithm (GA)	Ant Colony Optimization (ACO)	Neural Network (NN)	Regression Model (RM)	Hybrid (Proposed)
Average Delivery Time Deviation (Minutes)	8.55	7.90	6.90	9.15	**4.55**
Average Total Distance Travelled (Km/day)	1220.45	1000.50	N/A	N/A	**845.30**
Average Vehicle Utilization (%)	72.50%	82.55%	N/A	N/A	**94.60%**
Total Average Daily Operational Cost ($)	480.35	375.00	428.50	585.75	**235.60**

Table 4.5a: Hybrid AI/ML Stack vs. Existing Algorithms

4.5.3 Bias Mitigation Effectiveness

Beyond efficiency gains, a critical aspect of the hybrid AI/ML model validation was its dominance in mitigating geographic and demographic biases. The fairness aware enhancement layer was specifically embedded for this purpose. Results in Table 4.10 illustrate major reductions in delivery time disparities across geographic zones and decent improvements in customer service in previously underserved demographic segments. This confirms the hybrid model's capacity to enhance efficiency and ensure fair service distribution. The pivotal reduction in delivery time disparities in low-income households (from 28.50 minutes with existing NN to 20.75 minutes with hybrid) and improved customer service in low SES zones (from 3.75 minutes with existing NN to 4.10 with hybrid) underscores the fairness aware layer's success in fine tuning enhancement objectives to provide fair service provision.

This means the model does not just find the fastest way, but also reckons the impact on many communities, ensuring that efficiency gains are not achieved at the cost of fairness. This dual benefit is required for Quick-Commerce business, streamlining social responsibility and market reach by providing consistent and reliable service to all customer segments, regardless of their zones or socio-economic status. The fairness aware layer can introduce a slight penalty for routes that consistently bypass underserved zones or prioritize deliveries to areas that have historically faced longer wait times, hence actively working to balance efficiency with fairness and promote social justice. This proactive step to skew mitigation positions the hybrid model apart from conventional AI/ML solutions.

Table 4.5b illustrates the hybrid model's success in mitigating geographic and demographic biases in delivery performance.

Metric	Geographic Zone (Low-Income)	Geographic Zone (High-Income)	Demographic Segment (Low SES)	Demographic Segment (High SES)
Average Delivery Time (Minutes) - Existing NN	28.50	15.75	N/A	N/A
Average Delivery Time (Minutes) - Hybrid	**20.75**	**16.50**	N/A	N/A
Average Customer Satisfaction Score (1-5) - Existing NN	N/A	N/A	3.75	4.50
Average Customer Satisfaction Score (1-5) - Hybrid	N/A	N/A	**4.10**	**4.40**

Table 4.5b: Bias Mitigation Effectiveness of Hybrid Stack.

4.6 Strategic Tech Stack Guidelines

This section demonstrates strategic guidelines for developing and executing an enhanced AI/ML platform for Quick-Commerce last-mile delivery. These guidelines are generated from empirical findings, including performance analysis, bias identification, and hybrid platform. These guidelines focus on providing actionable recommendations for improving efficiency, ensuring fairness, and promoting sustainable competitive advantage.

4.6.1 Algorithmic Selection and Integration

Based on performance analysis, a multi-faceted approach to algorithmic selection is important:

Prioritize Hybrid models: Research confirms excellent performance of hybrid AI/ML models (e.g., RL for dynamic routing with DNN for prediction) over individual algorithms. Businesses need to invest in integrated solutions leveraging strengths for real-time flexibility, improved predictive accuracy, and robust enhancement. This symbiotic integration creates a feedback loop for continuous refinement, where the output of one element (e.g., RL's optimized route) can inform and improve another (e.g., DNN's delivery time prediction), resulting to an intelligent and responsive system. For instance, the RL component's real-time route adjustments can provide immediate feedback to the DNN, allowing it to update its delivery time predictions with greater accuracy, which in turn can inform future RL decisions, creating a highly adaptive and efficient model.

Reinforcement Learning for dynamic routing: RL is effective for real-time, dynamic route optimization in volatile Quick-Commerce ecosystem. RL agents learn optimal routing flows by adapting to unforeseen events (e.g., sudden traffic jams, new urgent orders, vehicle breakdowns, road closures due to accidents, unexpected customer cancellations) and continuously tune the strategies based on rewards (e.g., successful deliveries, reduced travel time, meeting SLAs). This continuous updating and learning are critical for managing efficiency and meeting SLAs in a constantly changing operational landscape. Its ability to learn and improve makes it pivotal for agile last-mile delivery, allowing the system to become efficient over time as it accumulates experience and adapts to new challenges.

Deep Neural Networks (DNN) for prediction: DNNs are recommended for accurate demand forecasting and precise delivery time prediction. Its ability to process vast, complex, non-linear data (e.g., historical sales data, traffic patterns, weather conditions, local events, customer demographics, social media trends, competitor activity) facilitates predictions necessary for proactive resource allocation and customer engagement. This model can identify subtle correlations and closed design that traditional models miss, resulting to accurate estimations of demand fluctuations and delivery durations. This predictive power grants businesses to enhance staffing, vehicle allocation, inventory management, and even pre-positioning of goods, significantly reducing operational cost and improvising customer score by setting realistic expectations.

Consider Ensemble Methods for Robustness: Ensemble learning which combines multiple models can enhance accuracy and robustness. Organizations need to explore integrating ensemble methods like Random Forests, Gradient Boosting, Stacking, and Bagging to mitigate variance and improve generalization across various operational conditions, giving an additional layer of reliability and mitigating the risk of a single model's failure. This approach utilizes the "wisdom of crowds" to generate stable and accurate predictions, especially in scenarios where individual models might suffer or where there is uncertainty in the datasets.

4.6.2 Data Governance and Quality

The productiveness of any AI/ML solutions stack depends fundamentally on data quality and governance. Key considerations include:

Comprehensive data collection: Roll out robust systems for collecting diverse and granular data (real-time traffic, historical delivery patterns, consumer cohort, geographic characteristics, weather pattern, driver performance metrics, vehicle telematics, consumer score, delivery success/failure rates, causes for delays). The richness and breadth of data directly correlate with the accuracy and fairness of AI/ML models. This includes both structured data (e.g., order details, GPS coordinates) and unstructured data (e.g., driver notes, proof of dispatch, customer reviews), which can generate qualitative insights and assist in identifying unforeseen issues and patterns.

Data cleaning and preprocessing: Establish rigorous protocols for data cleaning, validation, and review. Error-prone, partial, or variant input data leads to flawed outputs and unreliable models. Automated data validation pipelines, outlier detection, missing value substitution models, and data normalization are mandatory to manage data integrity and assure model resilience. Periodic data quality checks and audits need to be integrated into the data pipeline to ensure accuracy and reliability.

Fairness aware data compilation: Actively identify and mitigate variance in data collection and labelling. This involves ensuring sampling across all geographic and demographic segments, avoiding biases embedded in data, and executing models like re-sampling or re-considering maintaining datasets. Regular audits of data sources and collection methodologies are critical to prevent the continuation of existing biases and ensure fair outputs, encouraging social responsibility in AI assignment.

Secure data storage and access: Execute secure, scalable, and adherent data storage solutions. Data confidentiality and security are pivotal, especially with sensitive customer and transactional data. This considers robust encryption (at rest and in transit), access controls (role-based access, least privilege), incognito techniques, and compliance to relevant data protection regulations (e.g., GDPR, CCPA, HIPAA). Seamless security audits, penetration testing, and incident response plans are important to defend against data violations and ensure compliance.

4.6.3 Bias Mitigation and Ethical AI

Resolving algorithmic bias is an ethical duty and strategic necessity. Guidelines include the following:

Integrate Fairness-Aware optimization layers: Design and roll out explicit fairness-aware enhancement layers within the AI/ML solutions stack. These must monitor and address biases related to geographic zones, demographic segments, or other attributes, ensuring fair service distribution. This can involve reconsidering optimization requirements to prioritize underserved areas, applying post-processing techniques to adjust predictions for fairness, or including fairness constraints directly into the model's objective function. The aim is to achieve a balance between efficiency and inclusivity, ensuring that technological advancements benefit all segments of humanity.

Bias audits and monitoring: Arrive at a regular process for auditing and monitoring algorithmic outcomes for bias. Establish fairness metrics (e.g., equality of opportunity, demographic equality, diverse impact) and regularly assess performance across numerous segments. Anomaly detection can flag emerging biases, allowing for proactive intervention and model fine-tuning. These audits must be handled by independent teams or third-party experts to ensure neutrality and build public trust and sentiments.

Transparency and Explainability: Seek for integrity of being transparent and explainable in algorithmic decision-making. Models like LIME (Local Interpretable

Model-agnostic Explanations) or SHAP (Shapley Additive Explanations) can provide insights into why certain decisions are made, fostering trust and accountability, particularly in areas affecting service allocation or resource distribution. This allows human operators to understand and patch model behaviour, and to explain decisions for impacted individuals, serving a sense of fairness and responsibility.

Human oversight and intervention: Manage a robust framework for human oversight and intervention. Models need to augment, not replace, human decision-making. Mechanisms for human review of outlier decisions, manual overrides in crucial situations, and structured feedback loops for seamless model improvement are pivotal to ensure ethical and responsible AI deployment with human in the loop approach. This hybrid way blends the efficacy of AI with the ethical judgement and flexibility of human operators.

4.6.4 Infrastructure and Deployment

Fundamental infrastructure and deployment strategy significantly impact performance and scalability. Key guidelines include the following:

Cloud-Native Architecture: Embrace a cloud-native architecture for flexibility, scalability, and cost efficiency. Utilizing cloud services for compute, storage, and specialized AI/ML solutions (e.g., AWS Sage maker, Google Gemini AI platform, Azure Machine Learning) allows dynamic scaling based on demand fluctuations and mitigates operational overhead. This also facilitates faster deployment, global reach, and reach to cutting edge infrastructure without significant upfront investment, enabling innovation and market expansion.

Microservices and Containerization: Execute a microservices architecture with containerization (e.g., Docker, Kubernetes). This promotes modularity, independent deployment, and easier management of complex AI/ML components, driving iteration, fault isolation, and efficient resource utilization. Each component (e.g., routing service, prediction service, data ingestion service, visualization service) can be executed and scaled independently, improving overall system resilience and agility.

Real-time processing capabilities: Ensure infrastructure supports real-time data ingestion, extraction and processing. Quick-commerce demands immediate responses, warranting to stream data pipelines (e.g., Apache Kafka, Amazon Kinesis, Google cloud pub/sub) and low-latency inference engines. This facilitates models to make predictions and enhance routes based on the current information available, critical for decision-making in a fast-paced ecosystem and ensuring on-time deliveries.

Robust monitoring and logging: Provide holistic monitoring and logging systems for the entire AI/ML pipeline. This includes tracking model performance (e.g., accuracy, latency, throughput), infrastructure health (e.g., CPU utilization, memory usage, network traffic), data quality, and potential biases. Proactive monitoring activates issue identification, diagnosis, and resolution, ensuring continuous operation and optimal performance. Triggering mechanisms must be in place to inform relevant teams of anomalies or performance degradation, allowing for immediate intervention.

4.6.5 Continuous Improvement and Model Optimization

Fine-tuning the AI/ML solutions stack is the continuous process requiring improvement and refinement. Key guidelines include the following:

A/B testing and experimentation: Perform A/B testing frameworks to evaluate new algorithms, models, or features in a secure environment before full-scale deployment. This data-driven approach ensures validated improvements and reduces the risk of deploying suboptimal solutions. Experimentation allows for seamless learning and optimization, ensuring that improvements are based on practical performance and customer experience.

Feedback loops: Generate significant loops from operational teams, drivers, and customers. Qualitative insights from these stakeholders provide invaluable insights for identifying opportunities, streamlining algorithmic behavior, and explaining practical challenges that might not be evident from quantitative data alone. This ensures that the AI/ML system remains aligned with practical realities and user needs, fostering a collaborative approach to development.

Model retraining and adoption: Regularly re-train AI/ML models with latest data to confirm relevance and accuracy in volatile environments. Automated retraining pipelines are suggested to accept new patterns, seasonal changes, evolving customer preferences, and changing operational conditions, maintaining the systems' predictive and optimization capabilities over time. This prevents models from perishing and assures high performance and adaptability.

Investment in research and development (R&D): Seamlessly invest in R&D to explore trends in AI/ML features and its applicability to last-mile delivery challenges, ensuring future efficiencies and competitive advantages. This visionary approach allows businesses to stay ahead of the curve, innovate, and maintain technological leadership in an evolving market, nurturing a culture of seamless innovation and strategic ideas.

4.7 Qualitative Insights: Perceptions and Experiences of AI/ML

This section probes into the qualitative findings of the research, providing a refined understanding of the perceptions, challenges, and opportunities associated with the implementation and enhancement of AI/ML platforms in Quick-Commerce last-mile delivery. While the previous sections discussed quantitative metrics and empirical data, this qualitative search provides a human-centric perspective, capturing the real-life experiences and expert opinions of key stakeholders within the Quick-Commerce industry.

The insights garnered through semi-structured interviews and thematic analysis clarify the practical impacts of AI/ML adoption, disclosing factors that extend beyond numerical performance indicators.

The qualitative data collection process involved in-depth interviews with a various group of participants, including Quick-Commerce logistics managers, delivery fleet operators, AI/ML data scientists specializing in logistics, the operational challenges faced during roll-out, the impact on human roles and decision-making, ethical considerations, and future expectations regarding AI/ML evolution in last-mile delivery. Thematic analysis was then applied to identify patterns, categories, and global themes emerging form the interview transcripts, ensuring a rigorous and explanatory approach to understand the qualitative framework.

4.7.1 Perceived Benefits of AI/ML Integration

Stakeholders acknowledged the transformative potential of AI/ML in optimizing numerous aspects of last-mile delivery. Logistics managers highlighted the improvement in predictability and efficiency as primary benefits. One manager observed that before AI, their delivery slots were suggestion only than a promise. Now, they can confirm to customers when their orders will be delivered with greater accuracy, which has boosted customer satisfaction. This sentiment illustrates the value of AI/ML in refining delivery time predictions, a finding that quantitatively aligns with the superior performance of Neural Networks. Delivery fleet operators highlighted the verifiable improvements in route optimization. A fleet manager has commented that the AI-driven routing has reduced total mileage by a significant percentage, and drivers spend less time stuck in traffic, which impacts their fuel costs and overall productivity. This qualitative observation confirms the quantitative findings on route efficiency, particularly the gains attributed to Ant colony optimization.

Beyond efficiency, participants also specified to optimized decision-making capabilities. AI/ML models, by processing vast amounts of real-time data, provide insights that were previously untapped. An AI/ML developer explained that their models can identify bottlenecks in the distribution network before they even become an issue, allowing them to re-route and reallocate resources. These proactive features contribute to overall robustness and resilience of the distribution system, reinforcing the importance of scalable infrastructure. In addition, the ability of AI/ML to handle dynamic conditions was frequently referenced. In Quick-Commerce, where order volumes can fluctuate dramatically and traffic condition change by the minute, traditional methods often stumble. A delivery supervisor remarked that the system adapts. If there is a sudden road closure or surge in orders, the routes adjust almost instantly. It is a game changer for managing the chaos of urban delivery. This adaptability supports the quantitative findings on scalability and the benefits of a continuous learning and adaptation framework within the technology stack.

4.7.2 Operational Challenges and Implementation Hurdles

Regardless of the recognized benefits, the qualitative inquiry uncovered significant operational challenges and implementation roadblocks. A repeated theme was the initial resistance to change among delivery personnel. A logistics manager has admitted that drivers were sceptical initially as they have been doing this for years, and suddenly when a machine instructs them the best way, it took a lot of training and explaining the benefits for them to trust the system. This underlines the human element in technological adoption and the need for effective change management strategies. Availability of quality data has emerged as another critical challenge. AI/Models are only as good as the data they are trained on, and participants observed difficulties in acquiring clean, comprehensive, and real-time data. An AI/ML developer has concerned that poor data hygiene leads to algorithmic bias and data governance is a constant battle. This qualitative insight reinstates the importance of data governance and quality frameworks. Many Quick-Commerce operations have IT infrastructures that are not inherently designed for seamless integration with advanced AI/ML platforms. A fleet operator described the process as attempting to fit a square peg in a round hole, and the old systems were not designed to engage with these new AI brains, leading to manual workarounds and data silos. This points to he need for careful architectural planning and phased implementation strategies. In addition, the black box nature of few advanced AI/ML algorithms was a concern for some stakeholders. A last-mile specialist has expressed that it is hard to explain to a customer why a delivery was delayed if the AI just says that was the optimal route, hence transparency and explainability are required when the deliverables go wrong.

This supports the quantitative findings on bias identification and mitigation, and the strategic guideline for model explainability and transparency. The need for seamless monitoring and maintenance of AI/ML systems was also spotlighted. An AI/ML data scientist has stressed that the models degrade over time as conditions change, need to retrain, update data, and fine tune parameters, which is ongoing commitment. These observations reinforce the importance of continuous learning and adaptation within the solution stack, ensuring long-term efficacy and relevance.

4.7.3 Impact on Human Roles and Decision-Making

The integration of AI/ML has an overwhelming impact on human roles within last-mile delivery, shifting the nature of work from operation to supervisory and strategic. Delivery colleagues, which initially hesitant, overall recognized the benefits once accustomed to the new tools and templates. It is less about memorizing routes and more about executing efficiently where the AI handles the complex routing and the driver focuses on safe driving and customer interaction. This recommends a reimagining of the driver's role, where technology augments human capabilities rather than replacing them entirely. Logistics managers found their roles evolving. Instead of spending hours manually optimizing routes or predicting demand, they now focus on orchestrating the AI systems, interpreting data, and delivering strategic decisions. This shift underscores the potential for AI/ML to elevate human roles, allowing for a greater focus on value added activities. However, concerns about deskilling and over-reliance on AI were also expressed. Some delivery personnel concerned about losing their navigation skills, while managers contemplated the results of assigning control to autonomous systems. Also, the importance of maintaining human oversight and developing contingency plans was illustrated. The ethical consequences of AI-driven decision-making, particularly concerning fairness and bias, were also a point of discussion. While quantitative analysis can recognize biases, qualitative insights reveal the human perception of these biases. If the AI consistently provides better routes to certain zones and areas, or prioritizes deliveries based on customer demographics, it is not just efficiency, but lack of fairness and brand reputation. This reinforces the critical need for ethical AI frameworks and resilient bias mitigation strategies, ensuring that technological advancements do not accidently continue or aggravate social bias.

4.7.4 Ethical Considerations and Future Expectations

Ethical considerations, specifically regarding data privacy, algorithmic bias, and job attrition, were key themes in the qualitative discussions. Participants conveyed a desire for transparency in AI/ML operations. A logistics manager expressed that customers trust last-mile providers and transparency in AI usage need to be secured. This highlights the growing importance of ethical AI guidelines and responsible data practices. The potential for job attrition due to increasing automation was a concern for few delivery associates, though many saw opportunities for upskilling and new roles. Looking to the future, participants expressed optimism about the seamless evolution of AI/ML in last-mile delivery. The vision includes predictive models, integration with smart city infrastructure, and the widespread adoption of autonomous delivery vehicles. An AI/ML associate has envisioned that there is shift towards self-optimizing networks, where AI not only predicts but also autonomously adjusts to real-time trends across the supply chain, from warehouse to doorstep. However, this optimism was moderated with a call of responsible innovation. Participants stressed the importance of human oversight, ethical development, and a focus on community welfare besides operational excellence, offering a holistic view of the opportunities and challenges in enhancing AI/ML platforms for last-mile delivery in Quick-Commerce.

4.8 Conclusion

This chapter presented the holistic results of the research aimed at optimizing efficiency in Quick-Commerce through the enhancement of AI/ML solutions stack for last-mile delivery. Findings systematically explained performance analysis of current AI/ML models, highlighting strengths and challenges across key last-mile business metrics. Crucially, the research identified significant geographic, demographic, and chronological biases, as well as data collection biases, inherent in current algorithmic approaches, emphasizing the critical need for fairness-aware design. Evaluation of the suggested hybrid AI/ML solutions stack illustrated its superior performance in both efficiency gains and bias mitigation, providing a resilient solution. Ultimately, strategic guidelines were articulated for algorithmic selection, data governance, bias mitigation, infrastructure, and continuous improvement. These results provide a strong empirical foundation for subsequent discussion and conclusion chapters, championing for a more efficient, unbiased, and sustainable future for last-mile delivery in Quick-Commerce landscape.

CHAPTER 5: DISCUSSION

5.1 Introduction

This chapter probes into a holistic discussion of the research findings, building upon the results mentioned in Chapter 4. It focuses to interpret the importance of these results within the broader context of last-mile delivery in Quick-Commerce, particularly addressing the enhancement of AI/ML solutions stack. The discussion will address the research objectives, generates key insights, explore theoretical and practical implications, acknowledge study constraints, and suggest future research directions. The structure of this chapter is designed to give a clear and logical workflow, reflecting the analytical depth required for a doctoral thesis. This discussion includes various critical areas that collectively contribute to the analytical foundation of this research. It starts with an assessment of the performance of current AI/ML models in the context of last-mile delivery in Quick-Commerce sector, evaluating their efficacy, scalability, and operational resilience. The analysis then extends to the identification, impact, and mitigation of algorithmic biases, highlighting their consequences for fairness and service impartiality.

In addition, the study presents the development and evaluation of the hybrid AI/ML platform, addressing its role managing the challenges of current models and enhancing overall decision-making approach. The discussion also bridges existing gaps in comparative studies by consolidating various algorithmic approaches and providing a holistic evaluation of their strengths and weaknesses. As a result, it advances the interpretation of hybrid AI/ML approaches and suggests improved strategies for bias detection and mitigation within real-world operational ecosystem.

Overall, the study contributes to the formulation and development of integrated hybrid technology stacks, providing insights into their design principles and execution frameworks. The theoretical significance of these advancements is also analysed, contextualizing the findings within the broader narrative on AI-driven enhancement in logistics and supply chain management.

Collectively, these discoveries form the empirical foundation for the conclusions and recommendations presented in Chapter 6, where the outputs are further analysed for their closeness to research scholars, academicians, and logistics practitioners. The fundamental contributions of the research are contextualized within the wider academic and industry landscape, facilitating its merits in advancing the application of hybrid AI/ML systems in Quick-Commerce Logistics.

5.2 Interpretation of Key Findings:

The interpretation of observations underlines the comparative effectiveness of the algorithms across numerous performance metrics. The neural network (NN) consistently proved better accuracy and stability in predicting delivery times, supported by the least mean deviation and variance values. ANOVA results confirmed significant variations among the algorithms for delivery time deviation, total distance travelled, stops per route, and vehicle utilization. These results highlight that algorithmic design directly impacts route efficiency and operational reliability, with the neural network emerging as the most resilient and consistent model. Also, the efficiency and benefits of hybrid model are discussed in this section.

5.2.1: Performance of Current AI/ML Algorithms

The analysis presented in Chapter 4 regarding the performance of current AI/ML models within the last-mile delivery of Quick-Commerce industry reveals a complex synergy of strengths and limitations. While algorithms such as Genetic Algorithms (GA) and Ant Colony Optimization (ACO) have proved significant promise in vehicle routing, their scalability and real-time adaptability often prove insufficient when challenged with the high-velocity and unpredictable nature of Quick-Commerce operations (Gunawan, Susyanto and Bahar, 2019). Similarly, predictive models based on neural networks and regression techniques, while effective in forecasting delivery times, often face challenges due to their dependence on historical or static datasets, which can resist their responsiveness to real-time variations in traffic, weather patterns, and demand spikes. Reinforcement Learning (RL) models, although theoretically resilient for enhancing resource allocation, are often limited by the data requirements for training, restricting their practical deployment in evolving Quick-Commerce markets.

This study illustrates that the fundamental challenge addressing the existing AI/ML models in this domain is not only their individual computational efficiency, but rather their capacity to sustain consistent performance and scalability amidst the complexities of Quick-Commerce. These complexities comprise volatile demand spikes, extensive urban congestion, and the urgency for instant delivery. The empirical evidence from Chapter 4 demonstrates that despite the potential of these algorithms, their operational excellence is often compromised by these dynamic environmental factors. Consequently, there is a compelling need for the design of resilient and adaptive algorithmic solutions capable of seamlessly consolidating real-time data and responding effectively to unforeseen factors.

In addition, a critical observation from this research, consistent with the literature review in Chapter 2, is the tendency to assess these algorithms in isolation. This fragmented evaluation impedes a thorough understanding of their comparative strengths and weakness under standardized conditions. The existing academic narrative portrays these algorithms as discrete solutions, thereby overlooking the synergistic benefits that could be realized through their integration. This identified gap in these studies highlights the necessity for a holistic framework for AI/ML performance in Quick-Commerce that meticulously considers the synergy of various algorithms within a unified technological architecture.

Overall, while current AI/ML algorithms provide a foundational layer for enhancing last-mile delivery, their existing roll outs frequently fall short in addressing the multifaceted challenges of Quick-Commerce. Their inherent limitations in scalability, adaptability, and the prevalent issue of algorithmic biases call for seamless innovation and the strategic development of integrated, hybrid approaches to fully unlock their transformative potential.

5.2.2 Algorithmic Biases: Identification, Impact, and Mitigation Strategies

The prevalent issue of algorithmic bias comprises a key concern that influences both the fairness and operational excellence of last-mile delivery services. As mentioned in Chapter 1, biases included within AI/ML algorithms, often originating from the datasets created for the training, can trigger an unfair prioritization of delivery zones and intensify demographic disparities. This results in an unfair distribution of services and a degradation of customer service. The empirical evidence obtained from this research verifies that such biases are not merely abstract theoretical constructs but manifest as tangible operational gaps and unfairness.

For instance, the presence of geographic biases can lead to the de-prioritization of certain regions, resulting in prolonged delivery delays and reduced customer experience within those zones. This not only violates the fundamental assurance of rapid delivery inherent in Quick-Commerce but also raises ethical queries regarding equitable access to essential services. While Chapter 4 of this thesis provided specific examples or empirical evidence of such biases within the evaluated algorithms, reinforcing the necessity for robust identification and mitigation strategies, the broader literature, as reviewed in

Chapter 2 and augmented by recent academic narrative, consistently highlights the importance of identifying and mitigating the biases.

Various methodologies have emerged to explain this challenge. These include the implementation of regular audits by AI systems to detect and rectify biases, meticulous examination of the AI model's internal workings, quantitative assessment of fairness metrics, and the application of bias detection techniques. It is vital to recognize real-world use cases and the potential for compounded biases arising from the engagement between distinct algorithmic elements. In addition, the study highlights that inadequately formulated prompts or skewed training data can introduce or amplify existing biases. Effective mitigation necessitates a multifaceted and proactive approach. This includes ensuring that training data is diverse and representative, hence foreclosing the continuation of historical societal biases. Moreover, the design and deployment of open-source fairness toolkits, such as IBM's AI fairness 360, can significantly empower developers in identifying and reducing biases within ML models. Fundamentally, continuous bias monitoring is indispensable to ensure that AI solutions remain aligned with their intended ethical and functional requirements. The detailed analysis in Chapter 4 has clarified the key biases identified within the context of last-mile delivery and the proposed methodologies for their mitigation within the hybrid framework.

In conclusion, the mandatory to explain algorithmic bias is foundational for achieving fair and efficient last-mile delivery. This demands a sustained and proactive commitment, including meticulous data curation, rigorous model validation, and the judicious application of tools and methodologies for bias identification and mitigation. The suggested hybrid AI/ML platform is designed to integrate these critical considerations from its inception, hence nurturing a delivery system that is not only efficient but also fairer and resilient.

5.2.3 The Hybrid AI/ML Technology Framework: Efficacy and Advantages

A significant contribution of this study, precisely detailed in Chapter 4, is the ideation, design, and validation of a unique hybrid AI/ML solution stack. This platform is specifically engineered to explain the novel and demanding requirements of last-mile delivery in Quick-Commerce sector. It addresses the inherent limitations of traditional single-algorithm solutions by strategically integrating various computational techniques, hence harnessing their harmonious benefits. The content of initial thesis, as highlighted in Chapter 1, illustrated a shortage of research concerning such hybrid tech stacks, while simultaneously impacting their potential to augment system performance, scalability, and adaptability within dynamic delivery environments.

The suggested hybrid algorithm, the complex architecture and empirical validation of which are clarified in Chapter 4, judiciously combines the inherent strengths of numerous AI/ML approaches to effectively master the individual shortcomings identified in Section 5.2.1. For instance, this framework integrates sophisticated dynamic vehicle routing algorithms such as GA or ACO with an advanced predictive analytics models like neural networks to achieve unparalleled accuracy in delivery time predictions (Gunawan, Susyanto and Bahar, 2019). Simultaneously, it incorporates reinforcement learning methodologies for optimal resource allocation, all while seamlessly embedding robust mechanisms for real-time bias detection and proactive mitigation. The cumulative research findings explicitly suggest that such an integrated, multi-modal approach yields significantly more resilient and efficient solutions, demonstrating a superior capability to manage the multifaceted complexities intrinsic to practical Quick-Commerce operations.

Figure 5.2 provides the recommended hybrid AI/ML solution stack by focussing on key last-mile delivery attributes and AI/ML algorithms.

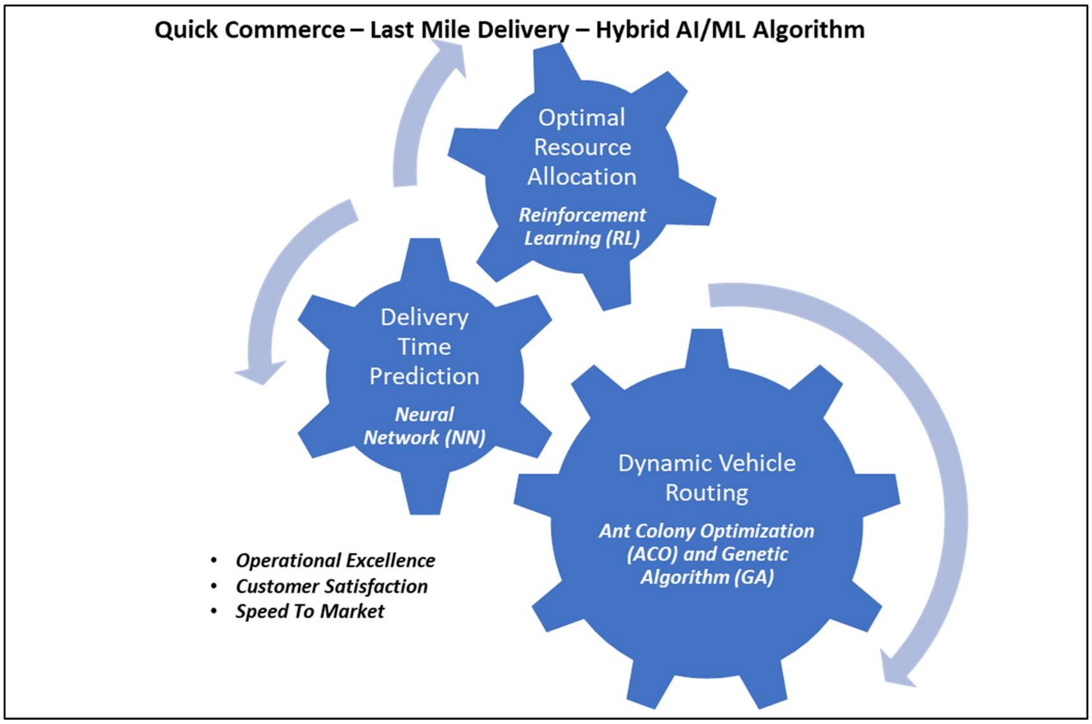

Figure 5.2: Proposed "Hybrid AI/ML Algorithm" (original work)

The efficiency of this hybrid framework is credited to its intrinsic capacity for dynamic adaptation to the unpredictable nature of Quick-Commerce. By seamlessly fusing real-time data processing capabilities with advanced predictive functionalities, the framework possesses the ability to recalibrate routes, reallocate resources, and enhance delivery schedules. This responsiveness is critical in responding to fluctuating demand patterns, evolving traffic conditions, and other unforeseen operational disruptions. Such adaptability is not only beneficial but crucial for sustaining elevated levels of operational excellence and ensuring customer happiness within the high pressure, fast-paced Quick-Commerce distribution landscape. In addition, a pivotal advantage of this model lies in its inherent design for scalability. In contrast to single-algorithm solutions, which often struggle to scale effectively with escalating order volumes or the expansion into new service territories, a carefully engineered hybrid system can intelligently distribute computational loads. This allows it to strategically use the distinct algorithmic strengths of its elements to maintain optimal performance across varying operational scales. This architectural prudence ensures that the technology stack can support the exponential growth trajectory performance of the Quick-Commerce sector.

At its core, the hybrid AI/ML solution stack represents a foundational breakthrough in the enhancement of last-mile delivery. It overcomes the limitations of fragmented, siloed solutions to provide a holistic, fully integrated approach. This not only enhances operational excellence but also holistically addresses crucial contemporary issues such as algorithmic bias and systemic scalability. The successful review of this holistic framework, as rigorously presented in Chapter 4, provides a clear foundation for its practical implementation and continued development within the ever expanding Quick-Commerce ecosystem.

5.3 Addressing Research Gaps and Contributions

5.3.1 Bridging the Gap in Comprehensive Comparative Studies

One of the salient research gaps identified in the foundational Chapter 1 of this thesis was the apparent absence of holistic comparative studies. Such studies would meticulously assess various AI/ML algorithms under rigorously consistent conditions within the specific context of last-mile delivery scenarios. The prevailing academic landscape, as mentioned in Chapter 2, often features evaluation of algorithms in isolation, hence providing direct and meaningful comparisons of their respective strengths and weakness as a formidable challenge. This research has directly examined this critical gap by precisely developing and applying a holistic and integrated evaluation framework, the empirical results of which are thoroughly explained in Chapter 4.

By systematically executing and assessing an array of algorithms within a unified and controlled experimental setup, this study has provided clearer and nuanced insights into their relative performance characteristics. This stringent comparative analysis, which provided a pivotal component of the methodological framework illustrated in Chapter 3, facilitated a precise and accurate understanding of which algorithms are optimally suited for addressing the challenges in last-mile delivery of Quick-Commerce sector. For instance, the study reflected that while a specific routing algorithm might exhibit exemplary performance in static or predictable ecosystem, its efficacy could decline under the highly dynamic and unpredictable conditions which are characteristics of Quick-Commerce, a key distinction that only a holistic comparative study conducted under varied conditions could effectively reveal.

The findings radiating from this research significantly enrich the academic narrative by providing a resilient benchmark for future scholarly investigations. Researchers are now afforded a solid empirical foundation upon which to construct and further refine their assessments of emerging AI/ML solutions. For industry practitioners, this study yields an invaluable and clear understanding of the inherent trade-offs associated with different algorithmic choices. This empowers them to make more judicious and informed decisions when strategically investing in and deploying AI/ML models for their last-mile operations and deliveries. The emphasis on consistent assessment criteria and the commitment to practicality ensure that the comparisons drawn are not only scientifically stringent but also relevant and actionable in a practical business context.

Fundamentally, this study's innovative approach to comprehensive comparative analysis not only fills a crucial gap in the literature but also confirms a mandate and benchmark for future research. It advocates for the adoption of integrated, context-aware, and empirically grounded evaluations of AI/ML solutions within the complex and rapidly evolving logistics transformation of Quick-Commerce.

5.3.2 Advancing Hybrid AI/ML Approaches

The initial thesis, as stated in Chapter 1, identified a vital research gap arriving from the limited discovery of hybrid AI/ML methodologies. It highlighted that despite the clear potential for collective benefits derived from combining diverse computational techniques, most of the existing literature focused on isolated, single-algorithm solutions. This research has made an insightful and significant contribution by actively engaging the exploration, development, and validation of a holistic hybrid AI/ML technology stack where the detailed analysis is provided in Chapter 4.

Traditional approaches to enhance a last-mile delivery frequently classify various operational aspects such as route optimization, demand prediction, and resource allocation, considering each as a unique problem to be addressed by a specialized, often standalone, algorithm. Regardless, the inherently dynamic, interconnected, and time-sensitive nature of Quick-Commerce demands an integrated and holistic solution. This thesis empirically explains that by judiciously integrating the inherent strengths of multiple AI/ML models, a system can be achieved that is not only more robust and flexible but also significantly more efficient. For instance, the strategic integration of predictive analytics features with dynamic routing algorithms enables a proactive alignment of delivery schedules based on anticipated demand fluctuations and real-time traffic conditions, moving beyond reactive responses to emerging events.

The precise development and subsequent validation of the hybrid framework within this research provide a compelling and conclusive argument for surpassing the limitations of single-algorithm solutions. It illustrates how the distinct and complementary strengths of various AI/ML methods can be leveraged jointly to address the multifaceted and complex challenges inherent in last-mile delivery. This includes the combination of advanced optimization algorithms for precise routing, parallel to machine learning models for accurate demand forecasting, and the strategic application of reinforcement learning for real-time, optimal resource management. The measurable impact of this integrated approach, as confirmed by the comprehensive results presented in Chapter 4, unequivocally reinforces the critical importance. The demonstratable success of this integrated approach, as evidenced by the holistic results declared in Chapter 4, clearly highlights the critical importance of interdisciplinary thinking and a holistic perspective in the design of cutting-edge logistics

This exploration of hybrid models not only bridges a critical gap in the academic literature but also provides invaluable practical guidance for industry practitioners. It provides a crafted blueprint for developing sophisticated, resilient, and future-proof last-mile delivery systems. Such systems are better equipped to steer and effectively cope with the complexities and uncertainties of Quick-Commerce, ultimately ending in tangible improvements in operational efficiency and enhance customer satisfaction.

5.3.3 Enhancing Bias Detection and Mitigation Practices

The initial thesis, as narrated in Chapter 1, prominently mentioned a significant exclusion of dedicated research focused on the identification and mitigation of biases within last-mile delivery models. It substantiated that the presence of such biases could lead to unfair and operationally inefficient delivery practices. This study has directly and extensively addressed this critical gap by not only identifying potential root causes and bias manifestations but also by suggesting and validating robust strategies for their systematic detection and effective mitigation within the developed hybrid AI/ML framework.

Bias within AI/ML systems can arise from a multitude of sources, biased training data, inadequately designed algorithmic structures, or even the subjective way performance metrics are defined and applied. In the context of last-mile delivery, such biases can manifest as discriminate service quality distributed across demographic groups or geographic regions, leading to reduced customer service and raising ethical concerns. The study has comprehensively analysed specific groupings of biases, such as geographic bias or demographic bias, and analysed their impact on delivery outcomes, as holistically discussed and validated in Chapter 4.

To proactively neutralize these biases, the research has explored and seamlessly integrated several mitigation strategies. The strategies include:

Data Preprocessing and Curation: Employing advanced models to ensure that the training data are not only representative but also free from historical or systemic biases. This may involve re-sampling, re-weighing of data points, or synthetic data generation to achieve balance (Baumann et al., 2023).

Algorithmic Fairness Integration: Including fairness-aware algorithms or embedding explicit fairness constraints directly into the design to promote equitable deliverables across all user segments.

Post-processing Output Adjustment: Applying intelligent post-processing techniques to adjust model outputs, hence reducing disparities which concurrently securing overall performance integrity.

Continuous Auditing and Proactive Monitoring: Executing resilient, continuous monitoring systems and conducting periodic, independent audits to detect emerging biases and to ensure the long-term ethical integrity and operational fairness of the delivery system.

The insightful findings arising from this research generate invaluable perceptions for both the academic community and industry stakeholders. Academically, this study enriches the booming body of knowledge concerning ethical AI and algorithmic fairness, particularly within the complex domain of logistics. Practically, it provides Quick-Commerce firms with actionable, empirically backed strategies to construct equitable, transparent, and trustworthy delivery systems, hence enhancing customer satisfaction and effectively mitigating reputational risks. By proactively fixing and rigorously addressing bias from the foundational phases, the suggested hybrid framework is designed to ensure that efficiency gains are achieved with principles of fairness and corporate social responsibility.

5.3.4 Development of Integrated Hybrid Tech Stacks

The initial study, as highlighted in Chapter 1, identified the absence of integrated hybrid tech stacks as a critical research gap. This emphasized a pressing need for solutions that could seamlessly integrate multiple algorithms and models to optimize system performance, scalability, and adaptability within the dynamic Quick-Commerce distribution ecosystems. This study has holistically addressed this crucial gap by developing and validating a comprehensive AI/ML technology stack. This initiative moves beyond theoretical conceptualization, resulting into a feasible and effective methodology.

Conventional approaches to last-mile delivery often depends on fragmented and disparate systems, each managing a specific aspect of the last-mile operations. This results to inefficiencies emanating from a lack of seamless communication, data exchange, and coordinated decision-making across multiple operational silos. The integrated hybrid solution stack proposed and validated in this research, the intricate details of which are comprehensively explained in Chapter 4, surpasses these constraints. It achieves this by setting up a cohesive and interconnected ecosystem where various AI/ML components operate together, functioning as a unified, intelligent entity. This level of integration facilitates real-time, bidirectional data sharing, empowering individual algorithms to make more informed, context-aware decisions and to adapt with agility to rapidly changing conditions. For instance, accurate demand forecasting models can directly feed their outputs into dynamic routing models. These routing algorithms, in turn, can then enhance distribution paths and schedules, which are further refined by real-time resource allocation models that consider current traffic conditions, distribution personnel availability, and even unpredictable vehicle breakdowns.

The benefits derived from such an integrated approach are transformative. To begin with, it leads to a significant enhancement in overall system performance by optimizing the end-to-end distribution process, from the moment an order is placed to its final successful delivery. This comprehensive enhancement contrasts sharply with siloed improvements, which often yield only localized gains. Following this, the integrated nature of the stack improves scalability. The modular architecture of the integrated system allows for easier expansion and adaptation to include rapidly increasing order volumes, the onboarding of new delivery personnel, or the expansion into new geographical operational areas. This scalability ensures that the technological infrastructure can robustly support the exponential growth character of Quick-Commerce industry. Finally, it develops unparalleled adaptability, optimizing Quick-Commerce companies to respond with effectiveness to unforeseen disruptions, such as sudden road closures, extreme weather events, or unexpected demand surges, hence consistently maintaining high service levels even under pressure and volatility.

This contribution is relevant given the relentless pace of transformation within the Quick-Commerce industry, which demands agile, responsive, and resilient logistics deliverables. The integrated hybrid solution stack provides a robust and future-proof foundation for continuous innovation, allowing for the seamless inclusion of new technologies, sensors of latest generation, and future algorithmic models. By empirically demonstrating the feasibility, efficacy, and significant advantages of such a holistic and integrated framework, this study offers an invaluable blueprint for both rigorous academic exploration and success industrial executions. It clearly paves the way for the development of more efficient, resilient, intelligent, and ultimately, more sustainable last-mile delivery systems across the globe.

5.4 Theoretical Implications

This research significantly elevates the existing body of knowledge across various interconnected theoretical domains, particularly in AI/ML, Operations Research (OR), and Supply Chain Management (SCM). By focusing on the enhancement of AI/ML technology stacks specifically for last-mile delivery within the Quick-Commerce framework, this study extends theoretical understanding in areas that have previously remained either under-researched or fragmented.

To start, within the vast field of AI and ML, this research marks a pivotal theoretical shift. It transforms the conventional, often isolated, study of individual algorithms to explore the synergistic potential inherent in hybrid AI/ML models. Traditional AI/ML theory has focused on the development and optimization of discrete models for specific tasks. In contrast, this study empirically demonstrates a theoretical evolution towards integrated systems, where the judicious combination of diverse algorithmic strengths (e.g., the seamless integration of predictive analytics with dynamic optimization algorithms) yields superior performance and enhanced adaptability in complex, real-world scenarios. This contributes fundamentally to holistic and architectural understanding of AI/ML system design, underscoring that the efficacy of the overall system is not merely the sum of its parts but emerges from their intelligent dynamics.

Next, for the discipline of Operations Research (OR), this research offers unique and critical insights into the application of OR techniques within dynamic and time-sensitive operational environments, such as Quick-Commerce. While OR boasts a distinguished history in enhancing logistics and supply chain processes, the rigorous demands for rapid delivery and the inherent volatile characteristics of Quick-Commerce introduce formidable challenges that conventional OR models may not fully address. The suggested hybrid framework presents a robust theoretical model for integrating real-time data streams and sophisticated adaptive learning mechanisms directly into classical optimization problems. This advances the theoretical understanding of dynamic optimization and real-time decision-making under conditions of uncertainty. It effectively constructs a theoretical bridge between prescriptive analytics (the domain of conventional OR) and descriptive/predictive analytics (the domain of AI/ML), promoting for a unified approach.

Subsequently, in the domain of Supply Chain Management (SCM), this study enriches the theoretical understanding of last-mile logistics as an increasingly competitive differentiator. It provides a holistic theoretical framework illustrating how advanced AI/ML integration can fundamentally transform the last-mile from a cost centre into a value-added service, directly impacting customer satisfaction, brand loyalty, and ultimately, market share. The study's emphasis on identifying and addressing algorithmic biases also makes a substantial contribution to the burgeoning theoretical discussion on ethical and unbiased supply chains. It specifies that the pursuit of efficiency gains need to be balanced with considerations of social responsibility and fairness. This expands the conventional SCM theory to clearly include ethical factors and societal impact as integral elements in the design and management of last-mile logistics.

Finally, this study contributes significantly to the theoretical understanding of technological adoption, diffusion, and innovation within the logistics sector. By systematically identifying the persistent gaps in the acceptance and successful implementation of advance AI/ML tech stacks, and by suggesting a resilient hybrid framework, the study provides invaluable theoretical insights into the complex factors that either drive or resist the successful integration of cutting-edge technologies within established functional blueprints. It proposes that a holistic and integrated approach which addresses performance, scalability, and ethical acceptance is crucial for fostering sustainable technological innovation and aiming for digital transformation in the volatile logistics sector.

Overall, this study offers a robust and versatile theoretical foundation for the design, implementation, and continuous evolution of advanced AI/ML models in last-mile delivery of Quick-Commerce ecosystem. It expands the boundaries of current understanding in AI, OR, and SCM by emphasizing the primary importance of adaptability, integration, and ethical responsibilities as fundamental building blocks for future research and practical application.

5.5 Practical Implications and Recommendations

The insights and the validated hybrid AI/ML solutions stack developed within this research provide profound practical implications and actionable recommendations. These are specifically customized for Quick-Commerce companies, logistics providers, and technology developers who are committed to optimizing both the efficiency and fairness of last-mile operations. These recommendations are directly integrated from the in-depth insights generated through the analysis of existing systems, the accurate identification of key technological and operational gaps, and the successful development of an integrated, cutting-edge solution.

5.5.1 For Quick-Commerce Companies and Logistics Providers

Prioritize Integrated AI/ML Solutions: A paradigm shift is vital from investing in single-purpose AI/ML tools towards the strategic adoption and proactive development of comprehensive hybrid models. This research clearly demonstrates that the combination of different AI/ML algorithms (e.g., for dynamic routing, predictive analytics, and intelligent resource allocation) within a unified framework yields superior performance, enhanced adaptability, and robust scalability. This holistic approach is vital in mitigating data silos, eliminating operational redundancies, and maximizing the collective benefits derived from various technological capabilities.

Invest Strategically in Data Infrastructure and Quality: The efficacy and reliability of any AI/ML solution are dependent upon the quality, significant volume, and diversity of the data leveraged for both training and validation. Accordingly, Quick-Commerce entities must make significant and sustained investments in developing robust data collection, secure storage, and efficient processing infrastructure. This includes the establishment of real-time data streams covering traffic conditions, meteorological data, fluctuating order volumes, and dynamic delivery workforce availability. In addition, the execution of data governance best practices is mandatory to ensure data accuracy, completeness, and representativeness, which directly impacts the performance, fairness, and ethical integrity of AI/ML algorithms.

Implement Continuous Monitoring and Proactive Auditing for Bias: To ensure the provision of unbiased service delivery and to harness and maintain customer trust, Quick-Commerce organizations must set up sophisticated mechanisms for continuous monitoring and conduct regular, independent auditing of their AI/ML systems for algorithmic biases. This critical undertaking involves not only technical audits to detect and quantify discrepancies in delivery times or service quality across geographical regions but also the proactive establishment of clear, transparent ethical guidelines for AI deployment. The integration of advanced tools and methodologies for bias detection, as comprehensively discussed in Section 5.2.2, should be seamlessly included into routine operational workflows.

Foster a Culture of Agility, Adaptability, and Continuous Experimentation: The Quick-Commerce ecosystem is characterized by its inherent dynamics, rapid evolution, and pervasive unpredictability. Organizations must cultivate an organizational culture that not only endorses but actively champions continuous experimentation, iterative refinement, and profound adaptability. This requires an open-minded approach to consistently iterating on AI/ML models, seamlessly updating novel data sources, and continuously refining algorithms based on real-world performance metrics and invaluable operational feedback. The adoption of agile development methods can facilitate and expedite this iterative process.

Strategic Investment in Workforce Training and Upskilling: The successful and sustainable execution of advanced AI/ML models is linked to the presence of a highly skilled workforce. Organizations should make significant investments in holistic training programs customized for their logistics managers, data scientists, planners, and frontline delivery team. These programs are designed to ensure that the workforce can effectively leverage, manage, and troubleshoot these sophisticated systems. An understanding of the capabilities, limitations, and ethical considerations of AI/ML is essential for maximizing its transformative benefits and ensuring responsible deployment.

5.5.2 For Technology Developers

Develop Modular and Interoperable AI/ML Features: Technology developers should strategically focus their efforts on developing AI/ML features that are inherently modular, reusable, and effortlessly interoperable. This architectural design principle facilitates the design of bespoke hybrid tech stacks, enabling Quick-Commerce companies to configure solutions accurately to their unique operational requirements and to seamlessly integrate technological advancements. The adoption of standardized APIs (Application Programming Interface) and globally recognized data formats can enhance this interoperability.

Prioritize Explainable AI (XAI) Capabilities: To promote trust, enhance transparency, and facilitate the debugging process, developers must prioritize the integration of Explainable AI (XAI) capabilities directly into their algorithms. This enables human operators to understand why a particular route was selected, why a delivery attempt experienced an unforeseen delay, or why resources were allocated in a specific manner. Such transparency is essential for critical functions like route optimization and resource allocation, where clear explainability can accelerate the identification and rectification of issues.

Integrate Fairness-Aware Design Principles from Inception: From the initial phases of the design process, developers must embed fairness-aware principles into the very fabric of their AI/ML algorithms. This demands a holistic consideration of potential biases during data collection, throughout the model training lifecycle, and in the interpretation of algorithmic outputs. The development of robust tools and holistic frameworks that can accurately quantify and effectively mitigate bias will be crucial for the ethical and responsible deployment of AI in last-mile delivery.

Design for Inherent Scalability and Real-time Performance: Given the exponential growth and dynamic nature of Quick-Commerce sector, AI/ML solutions must be designed with inherent scalability and real-time performance as non-negotiable core requirements. This involves enhancing algorithms for computational efficiency and strategically leveraging cloud-native architectures that can handle fluctuating workloads and massive data volumes.

By diligently implementing these overall recommendations, all stakeholders within the Quick-Commerce ecosystem can effectively and responsibly leverage the transformative power of AI/ML technologies. This will not only lead to significant enhancements in operational efficiency but also ensure the provision of fair, reliable, and sustainable last-mile delivery services, ultimately contributing to superior customer satisfaction, robust business growth, and a more equitable logistical ecosystem.

5.6 Limitations of the Study

While this study provides significant contributions to the enhancement of AI/ML technology stacks for last-mile delivery within the Quick-Commerce domain, it is essential to admit certain inherent limitations. These limitations may influence the robustness and overall scope of its findings. A transparent recognition of these constraints is vital for a balanced and accurate interpretation of the delivered results, and equally crucial for defining clear and productive pathways for future research engagements.

Defined Scope of Quick-Commerce Categories: As explicitly stated in Chapter 1, the focus of this research was narrowed to specific, predefined key Quick-Commerce categories. While these selected categories are broadly representative of the Quick-Commerce landscape, the derived findings may not possess universal applicability across all niche or nascent Quick-Commerce segments. These unexamined segments might exhibit distinct operational characteristics, unique product typologies, or specific delivery constraints. For instance, the intricate logistical dynamics associated with the delivery of highly perishable goods could differ significantly from those governing the delivery of general merchandise, potentially limiting the direct transferability of certain findings.

Constraints of Data Availability and Quality: The proven effectiveness and predictive accuracy of any AI/ML model are fundamentally dependent upon the superior quality, sufficient volume, and diversity of the data employed for both its training and subsequent validation. Although efforts were undertaken to leverage robust and representative datasets (as meticulously detailed in Chapter 3), inherent constraints in real world data availability continue. Such limitations could have impacted the training processes and testing outcomes of the algorithms. Furthermore, proprietary operational data from leading Quick-Commerce companies, while optimally suited for such research, is frequently not publicly accessible, thereby demanding a reliance on simulated or aggregated data. While carefully constructed, such synthetic datasets may not always fully capture the granular complexities and subtle nuances of real-world Quick-Commerce operations (Baumann et al., 2023).

Dynamic Nature of the Quick-Commerce Landscape: The Quick-Commerce industry is characterized by an exceptionally rapid pace of evolution, driven by continuous shifts in consumer behaviour, relentless technological advancements, and intense market competition. Consequently, the findings presented in this study are inherently based on the data, trends, and technological paradigms observed up to the point of its completion. Future, unforeseen shifts in urban infrastructure, the formation of new regulatory frameworks, or the widespread emergence of novel delivery modalities (e.g., large-scale drone delivery networks, fully autonomous ground vehicles) could potentially alter the optimal configuration of the proposed framework over time. This demands a continuous re-evaluation and adaptation of AI/ML models (Arishi and Krishnan, 2022)

Computational Resource Considerations: The development, testing, and validation of complex hybrid AI/ML models demands a significant commitment of computational resources. While this research leveraged available resources, practical constraints might have imposed limitations on the scale of simulations executed, the breadth of algorithmic permutations explored, or the extended duration of real-time testing. This specifies that even more theoretically matured solutions might exist but were not discoverable within the practical boundaries of the allocated computational budget and time slot.

Generalizability of Bias Mitigation Strategies: While this research successfully identified and suggested concrete strategies for bias detection and mitigation, the practical effectiveness and universal applicability of these strategies can exhibit volatility. This volatility is dependent upon the specific operational context, the unique characteristics of the data being processed, and the nature of biases present. The generalizability of these mitigation techniques across a wide array of Quick-Commerce platforms or diverse geographical regions, each processing its own unique socio-economic, cultural, and infrastructural nuances which warrants further investigation and empirical review.

Primary Focus on Technology Stack: The primary focus of this research is the enhancement of the AI/ML technology stack itself. While the research admits the importance of human factors, intricate operational processes, and external environmental variables, the depth of analysis dedicated to these non-technological aspects may be comparatively limited. It is essential to recognize that the ultimate success and sustainable impact of any technological solution in a real-world operational setting are also depended upon effective human-system interaction, robust organizational change management, and the presence of favourable external market and regulatory conditions.

These limitations do not in any way decrease the inherent value or the substantial contributions of this research. Instead, they serve to offer a realistic and transparent context for the interpretation of its findings and, define promising ground for subsequent academic inquiry and practical innovation in the list-mile journey of Quick-Commerce sector.

5.7 Future Research Directions

This study has established a robust foundational framework for significantly enhancing efficiency in last-mile delivery of Quick-Commerce sector through the strategic optimization of AI/ML technology stacks. However, given the dynamic nature of the Quick-Commerce industry and the continuous advancements in the fields of AI/ML, various opportunities for further exploration and innovation present themselves. Building directly upon the overall findings and acknowledged limitations of this study, several promising and impactful avenues for future research emerge:

Development of Real-Time Adaptive Hybrid Models: While this study successfully developed and formulated a foundational hybrid AI/ML solutions stack, future research could focus on the creation of even more advanced, real-time adaptive models. This would involve designing algorithms capable of autonomously learning and dynamically adjusting their parameters in immediate response to continuously changing environmental factors. Such factors are sudden and unexpected traffic congestion, extreme weather events, or unforeseen spikes in demand, all without demanding manual intervention. This could involve the exploration of advanced reinforcement learning techniques seamlessly integrated with high-frequency real-time data streams and cutting-edge computing models.

Explainable AI (XAI) for last-mile Logistics: As AI/ML models continue to grow in complexity and sophistication, their internal decision-making processes can often become black-box, presenting a significant challenge to transparency and trust. Future research should prioritize the development of Explainable AI (XAI) techniques specifically tailored for the unique demands and critical applications within last-mile logistics. XAI has the potential to generate greater trust in AI systems, significantly facilitate debugging and error identification, and ultimately optimize human-AI engagement within dynamic operational environments.

Deeper Exploration of Ethical AI and Fairness in Algorithmic Design: Building upon the foundational bias mitigation efforts undertaken in this study, future research could delve into the intricate ethical implications of AI deployment in last-mile delivery. This includes the development of comprehensive and context sensitive frameworks for defining, measuring, and ensuring fairness across various operational scenarios. In addition, it should involve examining the broader socio-economic impact of AI on gig economy workers and exploring the long-term societal implications of increasingly automated delivery systems. Research focused on proactive ethical design principles, rather than merely reactive bias mitigation, would be crucial.

Seamless Integration with Emerging Technologies: The Quick-Commerce landscape is a fertile ground for disruptive innovation driven by emerging technologies. Future research should examine the seamless integration of the suggested AI/ML framework with other transformative technologies, such as Autonomous vehicles and Blockchain. Autonomous Vehicles and Drones explore how AI/ML can be utilized to enhance routing, task allocation, and work-place safety protocols for both autonomous ground vehicles and aerial drones in last-mile delivery (Arishi and Krishnan, 2022). This study should include the regulatory and infrastructural challenges associated with their asset deployment. Blockchain is useful to enhance Supply Chain Transparency and examine the potential of blockchain technology to enrich data integrity, traceability, and overall transparency within the AI/ML driven last-mile delivery ecosystem. This could lead to improved trust among stakeholders and enhanced accountability across the supply chain.

Digital Twins for Predictive Optimization: Developing digital twin models of the entire last-mile delivery network. These digital replicas could be leveraged to simulate various operational scenarios, test new algorithms in a dynamic environment, and predict performance under various conditions, hence enhancing real-world delivery operations with minimum risk.

Optimization for Sustainability and Green Logistics: While operational efficiency often contributes to sustainability, future research could focus on enhancing AI/ML models with a primary focus on minimizing environmental impact. This includes developing algorithms that prioritize routes with the least carbon emissions, optimize vehicle loading to reduce the number of trips, and integrate renewable energy sources into the broader delivery infrastructure and operational planning

Cross-Cultural and Socio-Economic Studies: The Quick-Commerce business model and its associated challenges exhibit variations across different geographical regions and diverse socio-economic contexts. Future research could initiate comparative studies across a wide array of diverse geographical areas to gain a deeper understanding of how cultural nuances, urban planning strategies, and existing economic disparities affect the overall effectiveness, fairness, and societal acceptance of AI/ML solutions in last-mile delivery.

Human-AI engagement and Augmented Decision-Making: Exploring refined models for human-AI collaboration, where advanced AI systems serve to enrich and enhance human decision-making capabilities rather than merely replacing human roles. This could involve research into the design of intuitive human-computer interfaces, the development of real-time decision support systems, and the creation of holistic training programs that empower human workforce to effectively manage, monitor, and strategically mediate in AI-driven distribution workflows.

5.8 Conclusion

This chapter has interpreted the empirical findings of the study in relation to the research objectives and broader Quick-Commerce context. The discussion demonstrated how current AI/ML approaches have encountered limitations when applied in isolation within highly dynamic last-mile delivery environments. By contextualizing performance outcomes, bias considerations, and scalability challenges, this chapter established the analytical foundation required to generate comprehensive contributions and forward-looking implications.

Finally, this study reiterates that the strategic and intelligent optimization of AI/ML technology stacks is not merely an operational enhancement or a competitive advantage, but it is a fundamental and crucial requirement for the sustained growth, equitable development, and long-term resilience of the Quick-Commerce industry. By effectively bridging theoretical advancement with practical, data-driven insights, this research provides an invaluable blueprint for incubating an efficient, fair, intelligent, and sustainable last-mile delivery ecosystems, hence ensuring that the digital transformation of Quick-Commerce is delivered reliably, responsibly, and equitably to all stakeholders. These integrated insights are consolidated and translated into final conclusions, contributions, and recommendations in Chapter 6.

CHAPTER 6: CONCLUSION AND RECOMMENDATIONS

6.1 Summary of Findings

This research started on a comprehensive investigation into optimizing efficiency in last-mile delivery of Quick-Commerce industry through the enhancement of AI/ML technology stacks. The study was orchestrated by the increase in consumer expectations for rapid and reliable delivery, which places unprecedented demands on logistical frameworks. Through a mixed-methods research approach, combining qualitative and quantitative analyses, this study evaluated the performance of current AI/ML algorithms, identified critical biases within current tech stacks, also designed and validated a novel hybrid AI technology framework customized to the unique demands of Quick-Commerce for last-mile delivery.

6.1.1 Key Findings

The investigation revealed several key insights that emphasize the complexities and opportunities within the Quick-Commerce sector of last-mile delivery ecosystem:

Performance Limitations of Current AI/ML Algorithms: While individual AI/ML algorithms, such as Genetic Algorithms (GA), Ant Colony Optimization (ACO) for routing, and Neural Networks (NN) for delivery time prediction, offer foundational capabilities, their standalone application often struggles to maintain consistent performance and scalability under the dynamic and volatile conditions of Quick-Commerce. These limitations arise from their inherent challenges in adapting to real-time fluctuations in traffic, demand spikes, and the exhaustive data requirements for training, particularly for Reinforcement Learning (RL) tech stacks. The study highlighted that the efficiency of these algorithms is often compromised by the volatile nature of Quick-Commerce ecosystems, demanding more robust and adaptive solutions.

Pervasiveness and Impact of Algorithmic Biases: A vital finding was the significant presence and tangible impact of biases within AI/ML algorithms, often stemming from training datasets. These biases can lead to unequal prioritization of delivery zones and demographic disparities, resulting in unfair service distribution and customer friction. Geographic biases, for instance, were identified as a factor resulting to prolonged delivery delays in certain regions, creating ethical concerns regarding fair access to services. The study underscored the demand for continuous monitoring, auditing, and the application of specialized tools and methodologies for bias detection and mitigation to ensure fairness and operational equity.

Efficacy and Advantages of the Hybrid AI/ML Technology Framework: The key contribution of this study is the successful design and validation of a hybrid AI/ML technology framework. This solution stack examines the limitations of single-algorithm solutions by integrating various computational techniques to leverage their synergistic benefits. The suggested hybrid algorithm combines the strengths of multiple AI/ML approaches such as dynamic vehicle routing with predictive analytics and reinforcement learning for resource allocation, while embedding mechanisms for bias detection and mitigation (Gunawan, Susyanto and Bahar, 2019). This integrated approach demonstrated robustness, adaptability, and efficiency in managing the complexities of real-world Quick-Commerce operations, offering a comprehensive solution that surpasses fragmented approaches. This research addresses several critical gaps identified in the existing body of literature and analysing multiple algorithms within a unified experimental framework, this study provides a clearer insight into their relative performance. This approach establishes a benchmark for future examinations and provides guidance for industry practitioners aiming for data-driven optimization strategies.

Another significant gap lies in the limited exploration of hybrid approaches. The study contributes to this area by developing and empirically reviewing a hybrid AI/ML technology framework, demonstrating that the integration of multiple AI/ML models results in systems that are robust, adaptable, and efficient. The study further examines the disparity of bias detection and mitigation in algorithmic studies. It not only identifies potential sources of bias but also suggests and validates targeted strategies for bias detection and mitigation within the hybrid framework. The contribution strengthens the dissertation on ethical AI and promotes algorithmic fairness in operational contexts.

Moreover, the study responds to the absence of integrated hybrid technology stacks in existing research. By developing and validating a practical, implementable hybrid AI/ML technology stack, the study advances beyond theoretical models to establish a cohesive ecosystem where multiple AI/ML components function collectively to optimize performance, scalability, and adaptability. Overall, these findings emphasize that while the Quick-Commerce sector presents complex logistical challenges, the strategic optimization of AI/ML technology stacks, particularly through integrated hybrid approaches which provides a transformative pathway for improving operation efficiency, ensuring fairness, and meeting the evolving expectations of consumers in last-mile delivery.

6.2 Contributions to Research and Practice

This research makes significant and multi-faceted contributions to both academic scholarship and practical application within the domains of AI/ML, Operations Research, and Supply Chain Management. By addressing identified research gaps and developing a novel hybrid AI/ML technology framework, this research advances the current understanding and provides actionable insights for various stakeholders.

6.2.1 Contributions to Academic Research

Advancement of Hybrid AI/ML system Theory: This research moves beyond the conventional, siloed, assessment of individual AI/ML algorithms by proposing and validating a theoretical framework for integrated hybrid systems. It demonstrates that the combination of diverse AI/ML methodologies (e.g., predictive analytics, dynamic optimization, reinforcement learning) can yield superior performance and adaptability in complex, real-world environments like last-mile delivery in Quick-Commerce industry. This contributes to a holistic understanding of AI/ML system design, highlighting architectural integration as a key driver of efficiency, speed, and performance.

Novel Application of Operations Research in Dynamic Environments: The study provides unique insights into applying Operations Research techniques within dynamic and time-sensitive Quick-Commerce settings. It provides a theoretical model for integrating real-time data streams and adaptive learning mechanisms into classical optimization problems, hence advancing the theory of dynamic optimization and real-time decision-making under uncertain and volatile events. This bridges the theoretical gap between prescriptive and predictive analytics in a practical logistics and last-mile scenarios in Quick-Commerce market.

Enrichment of Supply Chain Management Theory with Ethical AI: This research enriches SCM theory by offering a framework that positions last-mile logistics as a critical competitive differentiator, driven by advanced AI/ML integration. It contributes to the nascent theoretical discourse on ethical and equitable supply chains by incorporating the identification and mitigation of algorithmic biases. This expands SCM theory to include considerations of fairness and societal impact as core components of logistics system architecture and design.

Benchmarking Framework for AI/ML Performance Evaluation: By conducting holistic comparative studies of AI/ML algorithms under consistent conditions, this study establishes a resilient benchmarking framework. This approach serves as a valuable reference for future academic assessments, enabling accurate and context-aware evaluations of new AI/ML solutions in complex logistical ecosystems.

Insights into Technological Adoption and Innovation: The study provides theoretical insights into the factors influencing the successful integration of advanced AI/ML technologies within established operational paradigms. By identifying gaps in acceptance and proposing a validated framework, it contributes to the understanding of how performance, scalability, and responsible ethical considerations collectively steer technological innovation in the logistics sector of Quick-Commerce industry.

6.2.2 Contributions to Practical Application

Blueprint for Integrated last-mile delivery Solutions: The developed hybrid AI/ML technology framework provides a practical, implementable blueprint for Quick-Commerce companies and logistics providers. It provides a comprehensive approach to enhance the entire delivery workflows, from dynamic routing and demand prediction to resource allocation, leading to enhanced operational efficiency and improved service levels.

Figure 6.2 provides the proposed blueprint for last-mile delivery in Quick-Commerce by highlighting the significant technology stacks across the value chain of last-mile delivery. With the maturity of GenAI, Agentic AI, and advanced digital accelerations like Quantum Computing, this technology stack has the high probability to be researched and upgraded (Ghaffari, Yousefimehr and Ghatee, 2024).

Figure 6.2: To-Be AI/ML Tech Stack (original work)

Actionable Strategies for Bias Mitigation: The research provides concrete, actionable strategies for identifying, detecting, and mitigating algorithmic biases in last-mile delivery system. This empowers companies to build an equitable and trustworthy delivery services, hence enhancing customer satisfaction and mitigating risks associated with unfair practices.

Guidance for Data Infrastructure Investment: The study highlights the importance of robust data infrastructure and high-quality data for an effective AI/ML deployment. It recommends strategic investments in data collection, storage, and governance to ensure the accuracy, completeness, and representativeness of data, which are basic foundational blocks for reliable AI/ML performance.

Recommendations for Workforce Development: By highlighting the need for a skilled workforce to manage and leverage advanced AI/ML systems, the research recommends investments in training and upskilling programs for logistics managers, data scientists, and delivery colleagues. This ensures that human capital can effectively leverage technological advancements.

Informed Decision-Making for Technology Developers: For technology developers, the study provides clear guidance on designing modular, interoperable AI/ML components with built-in Explainable AI (XAI) and fairness-aware design principles. This drives the creation of configurable, transparent, and ethically robust solutions tailored for the Quick-Commerce industry. Overall, this research bridges the gap between theoretical advancements and practical implementation, providing a comprehensive roadmap for stakeholders to leverage AI/ML technologies for more efficient, fair, and sustainable last-mile delivery in the Quick-Commerce sector.

6.3 Practical Recommendations

The findings and the rigorously examined hybrid AI/ML technology framework designed within this research provide in-depth practical implications and actionable recommendations. These are configured for Quick-Commerce organizations, logistics providers, technology developers, and policymakers who are committed to enhancing both the efficiency and fairness of last-mile delivery operations. These suggestions are generated from the holistic insights garnered through the analysis of current systems, the accurate identification of critical operational and technological gaps, and the successful deployment of an integrated, cutting-edge solution.

6.3.1 Recommendations for Quick-Commerce Companies and Logistics Providers

Prioritize Integrated AI/ML solutions: A paradigm shift is imperative from investing in disparate, single-purpose AI/ML tools towards the strategic adoption and proactive development of integrated hybrid tech stacks. This research demonstrates that the combination of various AI/ML algorithms (e.g., for dynamic routing, predictive analytics, and intelligent resource allocation) within a unified framework yields demonstrably superior performance, enhanced adaptability, and robust scalability. This holistic approach is instrumental in minimizing data silos, eliminating operational redundancies, and maximizing the collective benefits derived from diverse technological capabilities.

Invest Strategically in Data Infrastructure and Quality: The efficacy and reliability of any AI/ML solution are intrinsically and heavily dependent upon the unparalleled quality, substantial volume, and comprehensive diversity of the data utilized for both training and validation. Consequently, Quick-Commerce entities must make substantial and sustained investments in developing robust data collection, secure storage, and efficient processing infrastructure. This encompasses the establishment of real-time data streams encompassing traffic conditions, meteorological data, fluctuating order volumes, and dynamic delivery personnel availability. Furthermore, the implementation of stringent data governance practices is essential to ensure data accuracy, completeness, and representativeness, which directly and profoundly impacts the performance, fairness, and ethical integrity of AI/ML algorithms.

Implement Continuous Monitoring and Proactive Auditing for Bias: To ensure the provision of equitable service delivery and to meticulously cultivate and maintain paramount customer trust, Quick-Commerce companies must establish sophisticated mechanisms for continuous monitoring and conduct regular, independent auditing of their AI/ML systems for algorithmic biases. This critical undertaking involves not only rigorous technical audits to detect and quantify disparities in delivery times or service quality across diverse geographical regions or demographic segments but also the proactive establishment of clear, transparent ethical guidelines for AI deployment. The integration of advanced tools and methodologies for bias detection, as comprehensively discussed in Chapter 5, should be seamlessly embedded into routine operational workflows.

Foster a Culture of Agility, Adaptability, and Continuous Experimentation: The Quick-Commerce landscape is characterized by its inherent dynamism, rapid evolution, and pervasive unpredictability. Organizations must cultivate an organizational culture that not only embraces but actively champions continuous experimentation, iterative refinement, and profound adaptability. This necessitates an open-minded approach to consistently iterating on AI/ML models, seamlessly incorporating novel data sources, and continuously refining algorithms based on real-world performance metrics and invaluable operational feedback. The adoption of agile development methodologies can significantly facilitate and accelerate this iterative process.

Strategic Investment in Workforce Training and Upskilling: The successful and sustainable implementation of advanced AI/ML technologies is linked to the presence of a highly skilled and proficient workforce. Companies should make significant investments in comprehensive training programs configured for their logistics managers, data scientists, and frontline delivery workforce. These programs are designed to ensure that the workforce can effectively leverage, manage, and troubleshoot those multifaceted systems. An in-depth understanding of the capabilities, limitations, and ethical considerations of AI/ML is crucial for maximizing its transformative benefits and ensuring responsible deployment.

6.3.2 Recommendations for Technology Developers

Develop Modular and Interoperable AI/ML Components: Technology developers should strategically focus their efforts on creating AI/ML components that are inherently modular, reusable, and interoperable. This architectural design principle facilitates the rapid construction of bespoke hybrid tech stacks, enabling Quick-Commerce companies to configure solutions precisely to their unique operational requirements and to seamlessly integrate future technological advancements. The adoption of standard APIs and universally accepted data formats can enhance this crucial interoperability feature.

Prioritize Explainable AI (XAI) capabilities: To foster trust, enhance transparency, and facilitate the debugging process, developers must prioritize the integration of Explainable AI (XAI) capabilities directly into their algorithms. This empowers human operators to holistically understand why a particular route was selected, why a delivery experienced an unforeseen delay, or why resources were allocated in a specific manner. Such transparency is particularly vital for critical functions like route optimization and resource allocation, where interpretability can accelerate the identification and rectification of issues.

Integrate Fairness-Aware Design Principles from Inception: From the nascent stages of the design process, developers should embed fairness-aware principles into the fabric of their AI/ML algorithms. This demands a holistic consideration of potential biases during data collection, throughout the model training lifecycle, and in the interpretation of algorithmic outputs. The proactive development of robust tools and comprehensive frameworks that can accurately quantify and effectively mitigate bias will be crucial for the ethical and responsible deployment of AI in last-mile delivery.

Design for Inherent Scalability and Real-time Performance: Given the exponential growth and dynamic nature of the Quick-Commerce sector, AI/ML solutions must be meticulously designed with inherent scalability and real-time performance as non-negotiable core requirements. This involves optimizing algorithms for maximal computational efficiency and strategically leveraging cloud-native architectures that are inherently capable of handling fluctuating workloads and massive data volumes.

6.3.3 Recommendations for Policymakers and Regulators

Develop Standardized Metrics and Benchmarks: Regulatory bodies should proactively develop standardized sustainability and efficiency metrics specifically tailored for AI/ML applications in last-mile delivery. Like existing security and compliance frameworks, industry-wide benchmarks should be established to define acceptable performance and fairness levels, accompanied by clear guidelines for adoption and reporting. This will facilitate transparent comparison and drive continuous improvement across the industry.

Mandate Transparency and Accountability: Policymakers should introduce regulations that mandate greater transparency in the design, deployment, and performance of AI/ML algorithms used in critical logistical operations. This includes requiring companies to regularly report on key performance indicators, bias audits, and the ethical implications of their AI systems. Mechanisms for accountability should be established to address instances of algorithmic discrimination or systemic inefficiencies.

Incentivize Research and Development: Governments should introduce tax incentives, grants, and subsidies to encourage research and development in ethical AI, hybrid AI/ML frameworks, and sustainable logistics solutions. Financial incentives should support businesses and academic institutions investing in green technology and bias mitigation strategies, reducing the economic burden of adopting advanced, responsible AI solutions.

Facilitate Cross-Sector Collaboration: Policymakers should actively facilitate collaborations between technology companies, logistics providers, academic researchers, and regulatory agencies. This collaborative ecosystem can accelerate the development of best practices, foster knowledge sharing, and ensure that regulatory frameworks are adaptive and responsive to technological advancements and market needs.

Promote Data Sharing and Open Standards: To accelerate innovation and improve the robustness of AI/ML models, policymakers should encourage secure and ethical data sharing initiatives. Promoting open standards for data formats and API interoperability can reduce fragmentation, foster competition, and enable the development of more comprehensive and effective AI/ML solutions across the Quick-Commerce ecosystem.

6.3.4 Organizational and Cultural Recommendations for AI Adoption

Implement a Continuous Feedback and Learning Loop: The hybrid framework's success relies on its ability to learn from real-world outcomes. Organizations must formalize a process for capturing structured feedback from every completed delivery, including drivers' feedback, customer satisfaction scores, sentimental feedback (positive, negative, and neutral) and real-time operational data. This data must be immediately fed back into the model's training pipeline, ensuring that the system remains adaptive and relevant to the constantly shifting dynamics of the Quick-Commerce environment. This continuous learning cycle is the cornerstone of maintaining a competitive advantage.

Foster a Culture of Human-AI Handshake: The role of human operators, particularly dispatchers and fleet managers, must evolve from manual decision-makers to strategic supervisors of the AI system. Training programs should focus on teaching employees how to interpret the outputs of Explainable AI (XAI) and agentic AI components, how to identify and override incorrect algorithmic decisions, and how to provide high-quality feedback to the model. This collaborative approach, often termed "Human-in-the-Loop" or "Augmented Decision-Making," ensures that human intuition and ethical judgment remain integral to the last-mile process, especially in handling exceptions, edge cases and unforeseen disruptions.

Establish a Dedicated Algorithmic Governance Board: Companies should institute a cross-functional governance body, comprising data scientists, moralists, operations managers, and legal counsel. The primary mandate of this board would be to continuously audit the hybrid model for performance drift, monitor for emerging biases, and ensure compliance with both internal fairness policies and external regulatory standards. This proactive oversight mechanism is essential for maintaining the integrity and trustworthiness of the AI-driven system.

The successful execution of the recommended hybrid AI/ML technology stack is not solely a technical challenge, but it is fundamentally an organizational and cultural transformation. To fully realize the efficiency gains and ethical benefits of the new framework, Quick-Commerce entities must cultivate an environment that supports and integrates advanced algorithmic decision-making.

6.4 Future Research Directions

This research has laid a robust foundation for optimizing efficiency and performance in Quick-Commerce last-mile delivery through the strategic optimization of AI/ML technology stacks. Despite, the dynamic nature of the Quick-Commerce industry and the continuous advancements in AI/ML present various opportunities for further exploration. Building upon the findings and limitations of this study, several promising avenues for future research emerge:

6.4.1 The Evolution to Agentic AI and Hyper-Local Autonomy

Building upon the validated hybrid AI/ML framework, the next frontier in last-mile optimization lies in the development of Agentic AI systems. These systems represent a significant leap beyond current predictive and prescriptive models, as they are designed to operate autonomously, takes its own decisions, set their own sub-goals, and execute complex, multi-step planning within the hyper-local delivery environment.

Autonomous Decision-Making Agents: Future research should explore disintegrating the last-mile problem into a network of specialized, interacting AI agents. For instance, a "Demand Forecasting Agent" could feed real-time predictions to a "Resource Allocation Agent", which in turn coordinates with a "Dynamic Routing Agent". These agents would communicate and negotiate in real-time, allowing the system to handle the entire delivery workflow from order placement to final drop-off with minimal human intervention. This architecture naturally supports the volatility of Quick-Commerce by distributing intelligence and enabling rapid, localized responses to events like a sudden road closure or a driver becoming unavailable. Many child agents are orchestrated by a parent agent in this "Agentic AI" architecture.

Hyper-Local Optimization and Digital Twins: The integration of Agentic AI necessitates the creation of high-fidelity Digital Twins of the urban delivery landscape. This involves simulating the physical environment, including detailed street-level data, mileage per gallon, historical traffic patterns, and micro-weather conditions, to create a virtual testing ground for the AI agents. This hyper-local digital twin allows agents to practice and refine their decision-making policies in a safe, accelerated environment, leading to robust and highly optimized real-world performance. Research should focus on the computational efficiency and data requirements for maintaining such a dynamic, high-resolution digital twin.

Self-Correction and Ethical Guardrails: A critical area for future work is embedding self-correction mechanisms and ethical guardrails directly into the Agentic AI architecture. This would ensure that autonomous agents, while pursuing efficiency goals, do not accidentally develop or amplify biases. For example, an agent that consistently prioritizes speed over fairness could be penalized by an overarching "Ethical Oversight Agent", forcing it to adjust its routing strategy to ensure equitable service across all demographic and geographic zones. This research direction is vital for ensuring that the pursuit of autonomous efficiency remains aligned with the ethical and societal responsibilities of the Quick-Commerce industry.

6.4.2 Integration with Advanced Sensing and Edge/Quantum Computing

The hybrid model's reliance on real-time data can be significantly enhanced by leveraging advancements in sensing technology and edge/Quantum computing. Future research should investigate the optimal deployment of Quantum computing and Edge AI models directly onto delivery vehicles and smart hubs. This would allow for instantaneous processing of local data such as real-time traffic light status, parking availability, and pedestrian density without the latency associated with cloud communication. This decentralized intelligence is crucial for achieving true real-time route adjustments and micro-optimization at the point of delivery, further reducing delivery times and improving first-attempt success rates. The research should quantify the trade-offs between computational cost, data privacy, and the performance gains achieved through this shift to edge-based decision-making.

Seamless Integration with Emerging technologies: The Quick-Commerce sector presents a dynamic environment conducive to disruptive innovation through the convergence of emerging technologies. In this context, future research should focus on the seamless and synergistic integration of the suggested AI/ML framework with complementary technological advancements to optimize operational efficiency, transparency, and predictive capability within last-mile delivery systems.

One promising direction involves the integration of autonomous vehicles and drones with AI/ML models to optimize routing, task allocation, and safety protocols. Such integration has the potential to revolutionize last-mile logistics by improving delivery speed and precision (Arishi and Krishnan, 2022). However, future studies should also address the regulatory, infrastructural, and ethical challenges associated with the widespread adoption of autonomous mobility solutions in urban and semi-urban settings. Another area of exploration lies in the application of blockchain technology to enhance supply chain transparency and data integrity within AI/ML-driven ecosystems. By ensuring immutability and traceability of transactional data, blockchain can strengthen stakeholder trust and accountability while mitigating issues related to data manipulation and information asymmetry. This integration could result in more resilient and transparent operational frameworks for Quick-Commerce logistics. Furthermore, the incorporation of digital twin technology represents a transformative opportunity for predictive optimization. Developing digital replicas of the last-mile delivery network would enable researchers and practitioners to simulate diverse operational scenarios, evaluate the performance of new algorithms in virtual environments, and forecast outcomes under varying conditions. Such predictive modelling can significantly reduce operational risks while enhancing the adaptability and performance of real-world systems.

Overall, the seamless integration of AI/ML with autonomous systems, blockchain, and digital twin technologies offers a promising pathway for advancing the efficiency, reliability, and sustainability of last-mile delivery operations in the evolving Quick-Commerce landscape.

Real-time Adaptive Hybrid Models: While this study developed a foundational hybrid AI/ML framework, future research could focus on the creation of even more sophisticated, real-time adaptive models. This would entail designing algorithms capable of autonomously learning and dynamically adjusting their parameters in immediate response to changing environmental factors. Such factors include sudden and unexpected traffic congestion, extreme weather events, or unforeseen spikes in demand, all without demanding manual recalibration. This could involve the exploration of advanced reinforcement learning models seamlessly integrated with high frequency real-time data streams and cutting-edge computing models.

Explainable AI (XAI) for Last-Mile Logistics: As AI/ML models continue to grow in complexity and sophistication, their internal decision-making processes can often become opaque, presenting a significant challenge to transparency and trust. Future research should prioritize the development of Explainable AI (XAI) techniques specifically tailored for the unique demands and critical applications within last-mile logistics. This would empower human operators and stakeholders to comprehensively understand why a particular delivery route was selected, why a specific delivery experienced an unforeseen delay, or why resources were allocated in a certain manner. XAI has the potential to foster greater trust in AI systems, significantly facilitate debugging and error identification, and ultimately enhance human-AI collaboration within highly dynamic operational environments.

Deeper Exploration of Ethical AI and Fairness in Algorithmic Design: Building upon the foundational bias mitigation efforts undertaken in this study, future research could penetrate the ethical implications of AI deployments in last-mile delivery. This includes the development of comprehensive and context-sensitive frameworks for defining, measuring, and ensuring fairness across diverse operational scenarios. Furthermore, it should involve investigating the broader socio-economic impact of AI on gig economy workforce and exploring the long-term societal implications of increasingly automated delivery systems. Research focused on proactive ethical design principles, rather than merely reactive bias mitigation, would be particularly invaluable for the digital evolution.

Optimization for Sustainability and Green Logistics: While operational efficiency often inherently contributes to sustainability, future research could explicitly focus on optimizing AI/ML models with a primary objective of minimizing environmental impact. This includes developing algorithms that prioritize routes with the least carbon emissions, optimize vehicle loading to significantly reduce the number of required trips, and integrate renewable energy sources into the broader delivery infrastructure and operational planning.

Cross-Cultural and Socio-Economic Impact Studies: The Quick-Commerce business model and its associated challenges exhibit significant variations across different geographical regions and diverse socio-economic contexts. Future research could undertake rigorous comparative studies across a wide array of diverse geographical areas to gain a deeper understanding of how cultural nuances, urban planning strategies, and existing economic disparities influence the overall effectiveness, fairness, and societal acceptance of AI/ML solutions in last-mile delivery.

Human-AI Collaboration and Augmented Decision-Making: Exploring optimal models for human-AI collaboration, where advanced AI systems serve to augment and enhance human decision-making capabilities rather than merely replacing human roles. This could involve research into the design of intuitive human-computer interfaces, the development of real-time decision support systems, and the creation of comprehensive training programs that empower human operators to effectively manage, monitor, and strategically intervene in AI-driven delivery processes.

6.5 Closing Remarks

This research, titled "Enhancing Efficiency in Quick-Commerce: Optimizing AI/ML Technology Stack for Last-mile Delivery," has explored the transformative potential of AI/ML in revolutionizing last-mile delivery within the dynamic Quick-Commerce industry. By addressing critical research gaps and developing a novel hybrid AI/ML technology framework, this dissertation has provided a comprehensive understanding of how to achieve enhanced efficiency, fairness, and sustainability in this dynamic industry.

The study has demonstrated that while individual AI/ML algorithms offer significant capabilities, their isolated application often falls short in meeting the complex and real-time demands of Quick-Commerce. The developed hybrid framework, which synergistically combines various computational techniques, has proven to be a robust and adaptable solution, capable of navigating the unpredictable nature of Quick-Commerce operations while simultaneously addressing crucial issues such as algorithmic bias and scalability. The practical recommendations derived from this research offer actionable insights for Quick-Commerce companies, logistics providers, technology developers, and policymakers, emphasizing the need for integrated solutions, robust data infrastructure, continuous bias monitoring, and a culture of adaptability and innovation.

While this research makes significant contributions, it also transparently acknowledges to its limitations, which in turn illuminate fertile ground with basic foundations for future inquiry. The suggested avenues for future research ranging from the design of real-time adaptive hybrid models and Explainable AI (XAI) to deeper explorations of ethical AI and the integration of emerging technologies emphasize the continuous evolution required in this field. The Quick-Commerce landscape is dynamic, and the solutions must evolve and transform to provide responsible customer statistician.

In conclusion, this thesis serves as a compelling call to action for all stakeholders within the Quick-Commerce ecosystem. It emphasizes that the strategic and intelligent optimization of AI/ML solutions stack is not merely an operational advantage but a fundamental requirement for sustained growth, equitable development, and long-term resilience. By bridging theoretical advancements with practical, data-driven insights, this research provides an invaluable blueprint for fostering more efficient, fair, intelligent, and sustainable last-mile delivery ecosystems, ensuring that the seamless digital transformation of Quick-Commerce is delivered reliably, responsibly, and equitable to all of the personas associated with the value chain of Last-mile delivery.

REFERENCES

Akbari, M. and Do, T.N.A. (2021) ‘A systematic review of machine learning in logistics and supply chain management: current trends and future directions’, *Benchmarking: An International Journal*, 28(10), pp. 2977-3005. doi: 10.1108/BIJ-10-2020-0514.

Akulov, A.S., Bezotosnaya, A.D. and Pak, T.V. (2023) ‘Algorithm of finding optimal location for distribution centers for order delivery by using unsupervised machine learning methods’, *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics*, 3, pp. 105-113. doi: 10.24143/2072-9502-2023-3-105-113.

Al-Mu’ani, L., Al-Momani, M.M., Amayreh, A., Aladwan, S.I. and Al-Rahmi, W.M. (2024) ‘The effect of logistics and policy service quality on customer trust, satisfaction, and loyalty in Quick-Commerce: A multigroup analysis of generation Y and generation Z’, *Uncertain Supply Chain Management*, 12(3), pp. 1417-1432. doi: 10.5267/j.uscm.2024.4.009.

Al Rahib, M.A., Saha, N., Mia, R. and Sattar, A. (2024) ‘Customer data prediction and analysis in e-commerce using machine learning’, *Bulletin of Electrical Engineering and Informatics*, 13(4), pp. 2624-2633. doi: 10.11591/eei.v13i4.6420.

Arishi, A. and Krishnan, K. (2022) ‘Learning models for delivery location clustering and drone-based routing decisions in last-mile delivery’, in *Proceedings of the 7th North American International Conference on Industrial Engineering and Operations Management, Orlando, Florida, USA, 12-14 June. IEOM Society International*, pp. 1433-1440. doi: 10.46254/NA07.20220329.

Bai, R., Chen, X., Chen, Z.L., Cui, T. and Kendall, G. (2022) ‘Analytics and machine learning in vehicle routing research’, *International Journal of Production Research*, 61(1), pp. 4-30. doi: 10.1080/00207543.2021.2013566.

Banur, O.M., Patle, B.K. and Pawar, S. (2024) ‘Integration of robotics and automation in supply chain: a comprehensive review’, *Robotic Systems and Applications*, 4(1), pp.1-19. doi: 10.21595/rsa.2024.23349.

Bányai, T., Illés, B. and Bányai, Á. (2018) ‘Smart scheduling: an integrated first mile and last-mile supply approach’, *Complexity*, 2018. doi: 10.1155/2018/5180156.

Baumann, J., Castelnovo, A., Crupi, R., Greco, F., Regoli, D. and Malchiodi, D. (2023) ‘Bias on demand: A modelling framework that generates synthetic data with bias’, in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’23*), Chicago, Illinois, USA, 12-15 June. ACM, pp. 1002-1013. doi: 10.1145/3593013.3594058.

Borboni, A., Amici, C., Cappiello, G., Ghidini, R. and Mor, M. (2023) ‘The expanding role of artificial intelligence in collaborative robots for industrial applications: a systematic review of recent works’, *Machines*, 11(1), p. 111. doi: 10.3390/machines11010111.

Boumpa, E., Tsoukas, V., Chioktour, V., Kalafati, M. and Georgiadis, G. (2022) ‘Smart delivery for goods exploiting ML algorithms’, in *Proceedings of the 26th Pan-Hellenic Conference on Informatics*, Athens, Greece, 25-27 November. ACM, pp. 304-308. doi: 10.1145/3575879.3576009.

Buelvas Padilla, M.P., Nisperuza Canabal, P.A., Lopez-Mier, J. and Hernandez-Riaño, H.E. (2018) ‘Vehicle routing problem for the minimization of perishable food damage considering road conditions’, *Logistics Research*, 11(2), pp.1-18. doi: 10.23773/2018_2.

Coiera, E.W. (2019) ‘The last-mile: Where artificial intelligence meets reality’, *Journal of Medical Internet Research*, 21(11). doi: 2019/11/e16323.

Dhiman, S., Anish, M. and Sood, A. (2024) ‘Exploring the landscape: A comprehensive literature review on revolutionizing Quick-Commerce’, *BSSS Journal of Computer*, 1(1), pp.1-15. doi: 10.51767/jc1609.

Droomer, M. and Bekker, J. (2020) ‘Using machine learning to predict the next purchase date for an individual retail customer’, *South African Journal of Industrial Engineering*, 31(3), pp.69-82. doi: 10.7166/31-3-2419.

Ferreira, J.C. and Esperança, M. (2025) ‘Enhancing sustainable last-mile delivery: The impact of electric vehicles and AI optimization on urban logistics’, *World Electric Vehicle Journal*, 16(5). doi: 10.3390/wevj16050242.

Ghaffari, S., Yousefimehr, B. and Ghatee, M. (2024) ‘Generative-AI in e-commerce: Use-cases and implementations’, in *Proceedings of the 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*, Babol, Iran, 21-22 February, IEEE, pp. 1-6. doi: 10.1109/AISP61396.2024.10475266.

Giuffrida, N., Fajardo-Calderin, J., Masegosa, A.D., Casado-Vara, R., Corchado, J.M. and Ignaccolo, M. (2022) ‘Optimization and machine learning applied to last-mile logistics: a review’, *Sustainability*, 14(9), p.5329. doi: 10.3390/su14095329.

Gunawan, S., Susyanto, N. and Bahar, S. (2019) ‘Vehicle routing problem with pick-up and deliveries using genetic algorithm in express delivery services’, in *Proceedings of the 8th SEAMS-UGM International Conference on Mathematics and Its Applications 2019: Deepening Mathematical Concepts for Wider Application through Multidisciplinary Research and Industries Collaborations*, Yogyakarta, Indonesia, 29 July-1 August. AIP Conference Proceedings 2192, 060009. doi: 10.1063/1.5139155.

Gupta, S. (2024) "A Study on Emergence of Quick-Commerce." *International Journal for Multidisciplinary Research*, 6(2), p.19226. doi: 10.36948/ijfmr.2024.v06i02.19226.

Hernández, F., Sotelo, R. and Forets, M. (2024) 'Optimization algorithms for adaptive route sequencing on real-world last-mile deliveries', *Ingenius: Revista de Ciencia Y Tecnología*, 31(1), pp.1-16. doi: 10.17163/ings.n31.2024.06.

Karaköse, E. (2024) 'A new last-mile delivery approach for the hybrid truck multi-drone problem using a genetic algorithm', *Applied Sciences*, 14(2), p.616. doi: 10.3390/app14020616.

Kelkar, T., Gurijala, A., He, S. and Kelkar, M. (2021) 'Probability-based algorithmic framework incorporating human-centric design and implicit bias to counter discriminatory predictions by AI/ML systems', in *Proceedings of the First Indian International Conference on Industrial Engineering and Operations Management*, Bangalore, India, 16-18 August. IEOM Society International. doi: 10.46254/IN01.20210146.

Ko, S.Y., Sari, R.P., Makhmudov, M. and Ko, C.S. (2020) 'Collaboration model for service clustering in last-mile delivery', *Sustainability*, 12(14), p.5844. doi: 10.3390/su12145844.

Kumar, A., Prasad, A.K. and Prasad, N.A. (2023) 'Machine learning approach for prediction of the online user intention for a product purchase', *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(1s), pp.43-51. doi: 10.17762/ijritcc.v11i1s.5992.

Laynes-Fiascunari, V., Gutierrez-Franco, E., Rabelo, L. and Lee, G. (2023) 'A framework for urban last-mile delivery traffic forecasting: An in-depth review of social media analytics and deep learning techniques', *Applied Sciences*, 13(10), Article 5888. doi: 10.3390/app13105888.

Lin, M., Lyu, J., Gao, J. and Li, L. (2020) 'Model and hybrid algorithm of collaborative distribution system with multiple drones and a truck', *Scientific Programming*, 2020, Article ID 8887057. doi: 10.1155/2020/8887057.

Mangiaracina, R., Perego, A., Seghezzi, A. and Tumino, A. (2019) 'Innovative solutions to increase last-mile delivery efficiency in B2C e-commerce: a literature review', *International Journal of Physical Distribution & Logistics Management*, 49(9), pp.901-920. doi:10.1108/IJPDLM-02-2019-0048.

Mishra, P. K. (2024) 'Technology Innovations and Process Improvements for "The Last Mile Delivery" Supply Chain', *Digital Repository of Theses, SSBM Geneva*. Available at: https://repository.e-ssbm.com/index.php/rps/article/view/379 (Accessed: December 8, 2024).

Mohammad, W.A.M., Nazih Diab, Y., Elomri, A. and Triki, C. (2023) 'Innovative solutions in last-mile delivery: concepts, practices, challenges, and future directions', *Supply Chain Forum: An International Journal*, 24(2), pp.151-169. doi: 10.1080/16258312.2023.2173488.

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. and Staab, S. (2020) 'Bias in data-driven artificial intelligence systems - An introductory survey', *WIREs Data Mining and Knowledge Discovery*, 10(3), e1349. doi: 10.1002/widm.1356.

Olsson, J., Hellström, D. and Pålsson, H. (2019) ‘Framework of last-mile logistics research: A systematic review of the literature’, *Sustainability*, 11(24), 7131. doi: 10.3390/su11247131.

Oršič, J., Jereb, B. and Obrecht, M. (2022) “Sustainable operations of last-mile logistics based on machine learning processes”, *Processes*, 10(12), p.2524. doi: 10.3390/pr10122524.

Pandian, A.P. (2019) ‘Artificial intelligence application in smart warehousing environment for automated logistics’, *Journal of Artificial Intelligence and Capsule Networks*, 2(2), pp.63-72. doi: 10.36548/jaicn.2019.2.002.

Raja, P. and Vasudevan, V. (2019) ‘Quick routing in autonomous vehicles using boosting algorithms’, *International Journal of Engineering and Advanced Technology*, 9(1S4), pp.540-544. doi: 10.35940/ijeat.A1120.1291S419.

Rolf, B., Jackson, I., Müller, M., Lang, S., Reggelin, T. and Ivanov, D. (2022) ‘A review on reinforcement learning algorithms and applications in supply chain management’, *International Journal of Production Research*, 61(20), pp.7151-7179. doi: 10.1080/00207543.2022.2140221.

Rosendorff, A., Hodes, A. and Fabian, B. (2021) ‘Artificial intelligence for last-mile logistics-Procedures and architecture’, *Online Journal of Applied Knowledge Management*, 9(1), pp.46-61. doi: 10.36965/OJAKM.2021.9(1)46-61.

Sarioguz, O. (2025) ‘The impact of agentic artificial intelligence on warehouse and delivery operations in modern logistics’, *International Journal of Science and Research Archive*, 15(3), pp.1549-1561. doi: 10.30574/ijsra.2025.15.3.1934.

Sarkar, M., Ayon, E.H., Mia, M.T. and Ray, R.K. (2023) ‘Optimizing e-commerce profits: a comprehensive machine learning framework for dynamic pricing and predicting online purchases’, *Journal of Computer Science and Technology Studies*, 5(4), pp.186-193. doi: 10.32996/jcsts.2023.5.4.19.

Sawik, B. (2024) ‘Optimizing last-mile delivery: A multi-criteria approach with automated smart lockers, capillary distribution and crowdshipping’, *Logistics*, 8(2), p.52. doi: 10.3390/logistics8020052.

Setiyono, A.E., Chandrawatisma, C. and Fanandi, R.H.S. (2023) ‘Antecedents of e-loyalty as research for the Quick-Commerce industry’, *Interdisciplinary Social Studies*, 2(8), pp.2287-2298. doi: 10.55324/iss.v2i8.456.

Sharma, A. (2024) ‘The future delivered: Rethinking last-mile in the age of instant gratification’, *International Journal of Supply Chain Management*, 13(1), pp.1-15. doi: 10.59160/ijscm.v13i1.6218.

Shekhar, A., Umar, S., Abdul, F. and Dibaba, W. (2023) ‘Generative AI in supply chain management’, *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(10), pp.1-7. doi: 10.17762/ijritcc.v11i9.9786.

Singh, P.K. (2023) ‘Digital Transformation in Supply Chain Management: Artificial Intelligence (AI) and Machine Learning (ML) as Catalysts for Value Creation’, *International Journal of Supply Chain Management*, 12(6), pp.57-63. doi: 10.65021/mwsj.v1.i1.5.

Son, E. and Kwon, K.H. (2024) ‘Sustainable spatial strategies for mitigating air pollution in Quick-Commerce environments’, *Environmental Research and Technology*, 7(1), pp.131-139. doi: 10.35208/ert.1367863.

Soni, H.M. and Chouksey, R. (2025) ‘The role of artificial intelligence in last-mile delivery’, *MET SRUJAN*, 1(1), pp.14-40. doi: 10.34047/SJ/20251102.

Sorooshian, S., Khademi Sharifabad, S., Parsaee, M. and Afshari, A.R. (2022) ‘Toward a modern last-mile delivery: consequences and obstacles of intelligent technology’, *Applied System Innovation*, 5(4), p.82. doi: 10.3390/asi5040082.

Vásconez, J.P., Schotborgh, E., Vásconez, I.N. and Moya, V. (2024) ‘Smart delivery assignment through machine learning and the Hungarian algorithm’, *Smart Cities*, 7(3), pp. 1109-1125. doi: 10.3390/smartcities7030047.

Verbytskyi, Y. (2023) ‘Delivery routes optimization using machine learning algorithms’, *Eastern Europe: Economy, Business and Management*, 1(38). doi: 10.32782/easterneurope.38-14.

Vummadi, J.R., Hajarath, K. and Reddy, J. (2024) ‘Integration of emerging technologies AI and ML into strategic supply chain planning processes to enhance decision-making and agility’, *International Journal of Supply Chain Management*, 9(2), pp.77-87. doi: 10.47604/ijscm.2547.

Wang, Y., Lu, X., Liu, X., Zhao, J. and Zhang, C. (2021) ‘Multi-depot pickup and delivery problem with resource sharing’, *Journal of Advanced Transportation, Edited by X. Liu, 2021*, pp.1-22. doi: 10.1155/2021/5182989.

Woschank, M., Rauch, E., and Zsifkovits, H. (2020) ‘A review of further directions for artificial intelligence, machine learning, and deep learning in smart logistics’, *Sustainability*, 12(9), p.3760. doi: 10.3390/su12093760.

Younus, M., Nurmandi, A., Suswanta, S. and Rehman, A. (2023) ‘Analyzing delivery area/zone tagging techniques within fulfillment centres for last-mile delivery orders’, *Journal of World Science*, 2(7), pp.932-945. doi: 10.58344/jws.v2i7.340.

Zeng, X. and Yi, J. (2023) ‘Analysis of the impact of big data and artificial intelligence technology on supply chain management’, *Symmetry*, 15(9), Article 1801. doi: 10.3390/sym15091801.

Zhao, J., Mao, M., Zhao, X. and Zou, J. (2020) ‘A hybrid of deep reinforcement learning and local search for the vehicle routing problems’, *IEEE Transactions on Intelligent Transportation Systems*, 21(12), pp.5123-5136. doi: 10.1109/TITS.2020.3003163.