CHAPTER 1

INTRODUCTION

1.1 Introduction

In modern supply chain management, many calculation and prediction methods have long appeared in the market to accurately predict the quantity demand of regional supply and the expected arrival time of shipment. However, there are still various delayed deliveries and inventory shortages in the market, which lead to supply chain paralysis. Most enterprises rely on traditional methods to do forecasts according to the assumptions and historical data, but the efficiency of obtaining prediction results is reduced and the cost is increased (Sani, S et al., 2023). Regarding the aspect of the capital chain, there will be daunting impacts since small and medium-sized enterprises cannot compete with large enterprises. Hence, this project will focus on the delayed prediction of inventory shortage in the sports equipment supply chain issue, and use the sports equipment datasets to minimize the possibility of out-of-stock, overstock, and delayed shipment (Abouloifa & Bahaj, 2022).

1.2 Background of Study

Global cargo supply has experienced tremendous challenges, among which shipment delays and insufficient inventory still cannot resolved. There is a great correlation between accurately predicting the required inventory to ensure the profitability of enterprises and the satisfaction of consumers who expect to receive the stuff. Usually, the expected delivery time of stuff was delayed due to suppliers, transportation companies, data management, and other factors. On the supplier side, the most common factors are machine failures and raw material shortages leading to production delays and insufficient inventory (Gabellini et al., 2022). In terms of

transportation factors, there might be issues of mismatch between the number of loaded goods and the number of transport vehicles, improper route planning, and severe weather conditions (Sani et al., 2023). Human factors are generally the cause of data management, such as slow manual processing of orders and inability to ship due to information asymmetry. Beyond a doubt, these problems significantly impacted both suppliers and consumers, and people would gradually lose confidence in this field if they continued (Abouloifa & Bahaj, 2022).

1.3 Statement of Problem

From the perspective of the current technologically advanced world, it is impossible to further achieve high-precision accuracy by continuing to rely on traditional methods when real-time and seasonal changes need to be taken into account. The accuracy of an enterprise's inventory forecast is equivalent to the enterprise's profit (Abouloifa & Bahaj, 2022). Likewise, delivery punctually is equivalent to a promise from the consumers' perspective. Since current forecasting methods may not be sufficient to face all the challenges in this highly competitive market landscape, the accuracy of forecasts is crucial and related to the company's reputation and development prospects. It can ensure that the enterprise's inventory is sufficient and not excessive while ensuring that the goods are delivered to customers on time, which improves the enterprise's profits and customer satisfaction, creating a win-win situation (Sani et al., 2023; Gabellini et al., 2022).

1.4 Aim of the project

This project aims to provide an accessible, affordable, and effective forecasting analysis method to predict market trends and demand by studying various calculation methods to correspond to different supply and demand, help enterprises optimize inventory management, minimize out-of-stock and overstock, accurately calculate the arrival time of goods, and promote data-driven decision-making.

1.5 Research Goal

The research goal will describe the research objectives and questions of the Delay Prediction on Inventory Shortages in the Sports Equipment Supply Chain project.

1.5.1 Research Objectives

The objectives of the research are :

(a) To identify historical sales and inventory data of enterprises and pre-process them to handle missing values, outliers, and seasonality

(b) To derive complex forecast results into actionable insights using Tableau and Power BI

(c) To evaluate the performance of various forecasting models using ARIMA and XGBoost, SARIMA, and LSTM based on forecast accuracy, reliability, and applicability

1.5.2 Research Questions

The questions of the research are :

(a) What are the primary factors contributing to delays in the sports equipment supply chain?

(b) How does the performance of hybrid models like ARIMA + XGBoost compare to traditional time-series models like SARIMA in predicting inventory shortages?

(c) Can deep learning models, such as LSTM, effectively capture long-term dependencies in supply chain data for better delay predictions?

1.6 Hypothesis

This project combines machine learning models (ARIMA + XGBoost) to improve the accuracy of delay prediction for inventory shortages compared to traditional models (SARIMA) in the sports equipment supply chain.

1.7 Assumption

Based on the theoretical perspective of this project, assume the following:

(a) Inventory shortages or overstock can be predicted in advance to determine the length of delays, which can confirm consumers' confidence in the platform

(b) Overloaded transportation, supplier reliability, and unstable weather factors bring uncertain variables, which require various forecasts to prevent the worst from happening

(c) The demand for sports equipment has a specific seasonal peak, which can be confirmed by calculating data through forecasting models

1.8 Significance of the study

The following are some of the contributions that this study will make to the research population, stakeholders, and specific fields:

(a) To contribute to optimizing existing supply chain management by comparing the hybrid model with traditional delay prediction methods, which ARIMA + XGBoost and SARIMA methods

(b) To predict the delay duration by studying how to calculate and turn uncontrollable factors into controllable factors

(c) To enable sports equipment enterprises to reduce inventory shortages, improve customer satisfaction, and minimize the costs associated with delays

1.9 Scope of the study

This project focuses on predicting the inventory instability of international logistics companies from 2015 to 2018. The research uses inventory data of sports equipment supply chains held by selected suppliers and uses forecast data accurately calculated from multiple algorithm models.

1.10 Theoretical and Conceptual Framework

Two frameworks provide a foundation for this project, which the conceptual framework is building up the concept of work through the literature, while the theoretical framework is the theory that supports each component of the mind map

1.10.1 Theoretical Framework

(a) Theory of Constraints (TOC): To analyze bottlenecks in the supply chain that cause delays

(b) Systems Theory: To understand how disruptions in one part of the supply chain impact overall inventory levels

(c) Predictive Analytics Theory: To justify the use of advanced forecasting models for delay prediction

1.10.2 Conceptual Framework

(a) Independent Variable (Cause): Transportation time, supplier reliability, demand variability, weather conditions

(b) Dependent Variable (Effect): Inventory shortages

(c) Moderating Variable: Seasonal demand, predictive accuracy

1.11 Definition of terms

Term	Definition
Delay Prediction	The process of accurately calculating and predicting the severity of supply chain delivery delays and the possibility of various factors (Keung et al., 2021)
Inventory Shortage	A phenomenon in a supply chain scenario where demand is greater than the existing inventory and leads to the orders not being fulfilled or supply shortage (Keung et al., 2021)
Autoregressive Integrated Moving Average (ARIMA)	A statistical model for time series forecasting (Gabellini et al., 2024) that combines: ● Autoregressive component (AR) - measures the relationship between the current data point and the past values of the data set ● Integrated component (I) - calculates differences to achieve stationarity ● Moving average component (MA) - calculates the dependency between the current data point and past forecast errors
Extreme Gradient Boosting (XGBoost)	A machine learning algorithm for predictive modeling that achieves scalable and accurate forecasting (Keung et al., 2021)
Seasonal ARIMA (SARIMA)	A statistical model for time series forecasting that can handle periodic changes in data, making up for the inability of the ARIMA to predict seasonal changes (Keung et al., 2021)
Long Short-Term Memory (LSTM)	A recurrent neural network (RNN) that handles long-term reliance on sequence data, and can handle problems with recurring delays and fluctuating demand (Keung et al., 2021)
Theory of Constraints (TOC)	A process that identifies and addresses key constraints to improve overall performance, such as inefficient supplier production, delayed shipments, and other bottlenecks (Gabellini et al., 2024)
Systems Theory	A theory that considers the supply chain as an interconnected system and assumes that changes in any system part will affect the entire system operation (Keung et al., 2021)

Predictive Analytics Theory	A theory that predicts future outcomes by combining statistical models, historical data, and machine learning algorithms (Keung et al., 2021)
Independent Variable	A variable that is manipulated or measured by the dependent variable effect (Hudnurkar et al., 2024)
Dependent Variable	A variable that changes in the independent variable will affect the response or outcome (Keung et al., 2021)
Moderating Variable	A variable that affects the direction or strength of the relationship between the independent and dependent variables (Hudnurkar et al., 2024)

1.12 Markmap

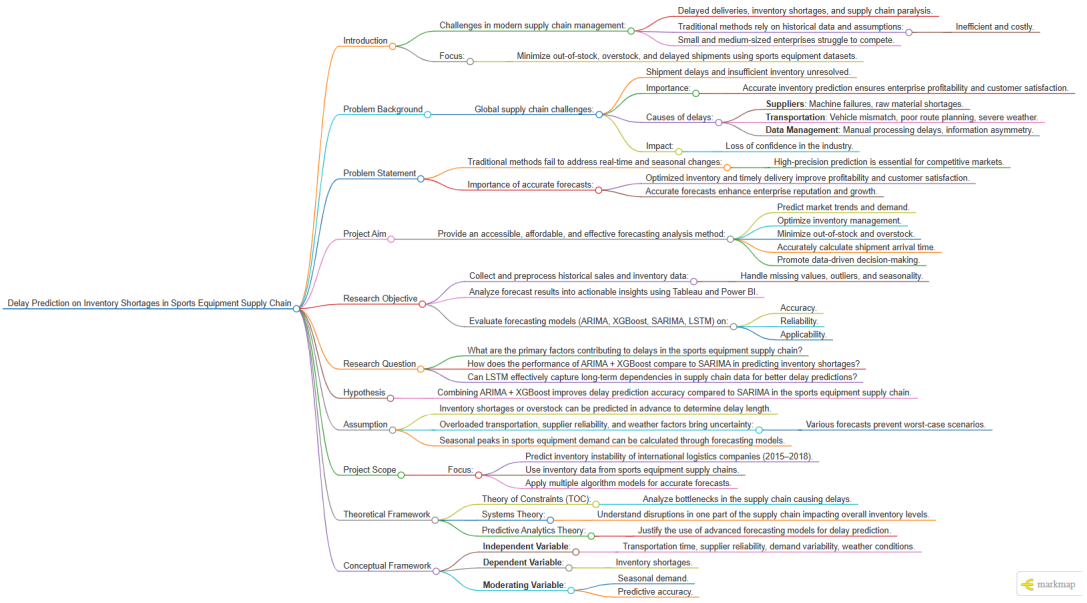

Figure 1.1 Introduction Markmap