Increasing Supply Chain Effectiveness: Forecasting Models for Order Quantity Prediction

MSc Research Project
Master of Science in AI for Business

Cem Akilli
Student ID:X23147016

School of Computing
National College of Ireland

Supervisor: Faithful Onwuegbuche

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Cem Akilli
Student ID:	x23147016
Programme:	AI for Business
Year:	2023-2024
Module:	MSc Research Project
Supervisor:	Faithful Onwuegbuche
Submission Due Date:	16/09/2024
Project Title:	Increasing Supply Chain Effectiveness: Forecasting Models for Order Quantity Prediction
Word Count:	6648
Page Count:	18

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Cem Akilli
Date:	14th September 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Increasing Supply Chain Effectiveness: Forecasting Models for Order Quantity Prediction

Cem AKilli
x23147016

Abstract

For the retail and whole sale industries, satisfying customer demands while preserving a competitive edge requires effective supply chain management. This management is greatly dependent on the precise demand forecasting as it foresees future requirement of goods. This study explores different modern and classical time series forecasting models like the Long Short-Term Memory networks networks, Autoregressive Integrated Moving Average in order to better a prediction accuracy. This study was motivated by the necessity in inventory management, production scheduling, and market entry choices of demand forecasting. Mistakes in demand forecasting can cause substantial operational inefficiencies and financial losses. In order to begin answering these questions, this study evaluates how well the cyclical models forecast US regional sales data utilizing historical information. It is an evaluation of all the models sens, spec and acc to help with which forecasting technique we can go for. These results suggest that while for simple, linear models such as ARIMA traditional models may still be effective LSTM could provide greater accuracy in complex and non-linear ones. This summarises the importance of selecting a forecasting model based on various data characteristics and forecasting needs. The study's findings offer supply chain professionals insightful information that will help them make better decisions, maximize inventory levels, and boost overall operational effectiveness. This research advances our understanding of forecasting model efficiency, which benefits supply chain management both theoretically and practically.

1 Introduction

Accurate demand forecasting plays a critical role in supply chain management in today's highly competitive retail and wholesale industries. In order to optimize inventory levels, plan production, and reduce operating costs, businesses must implement effective forecasting models. In this study, we focus on two specific forecasting models: ARIMA and LSTM networks. ARIMA is a traditional time series model widely used to predict linear trends, while LSTM, a recurrent neural network model, is more suitable for capturing complex, nonlinear patterns in data. This research will evaluate these models and explain how accurate demand forecasting is the key to accurate purchasing for accurate inventory, which is an important link in the supply chain.
The need to better comprehend and anticipate intricate, nonlinear patterns in sales data has fueled the development of forecasting tools Kilimci et al. (2019). Of course, this process could be impacted by a variety of external and demographic factors. One significant

one was the COVID-19 pandemic that struck us recently. We have demonstrated the importance of supply networks on a global scale, but we still need to ensure that these chains are strongly connected. Businesses used to base their predictions on expert opinion and basic biases. These methods, however, were frequently arbitrary and prone to serious mistakes. Time series analysis is one of the more popular statistical techniques as data gathering and storage technology have advanced Khan et al. (2020). By using historical data to spot trends and seasonal patterns, these techniques—which include ARIMA (Autoregressive Integrated Moving Average) and SARIMA (Seasonal ARIMA) —offered a more methodical approach to forecasting El Filali et al. (2022). Traditional statistical models are still widely used, but they have disadvantages, particularly when it comes to managing complicated and nonlinear interactions in data. Researchers and practitioners have resorted to machine learning (ML) approaches in order to address these obstacles. In recent years, a lot of research has focused on machine learning (ML) models, like XGBoost, Decision Trees, and Long Short-Term Memory (LSTM) networks, since they provide more accuracy and flexibility by automatically learning from big data sets and identifying intricate patterns that traditional models could overlook El Filali et al. (2022). These cutting-edge methods have proven to be highly promising in terms of increasing forecasting accuracy, and they are now valued additions to the toolset utilized by modern supply chain professionals. It is equally important to evaluate the effectiveness of these predictive models, which requires reliable performance measurements. The accuracy of a predictive model is often evaluated using metrics such as Mean Absolute Percent Error (MAPE) and Symmetric Mean Absolute Percent Error (SMAPE). SMAPE varies the mean absolute percentage error between expected and actual values to account for potential scale and data distribution issues; thus providing a more balanced perspective of model performance than MAPE, which evaluates prediction accuracy by calculating this difference. This makes final images easier to understand for subtle projections into the future Wiyanti et al. (2021). In the context of retail supply chain management, this study attempts to assess the effectiveness of both traditional and modern forecasting models, including ARIMA, SARIMA, XGBoost, Decision Trees, and LSTM networks. The objective is to identify the most effective methods for different scenarios by utilizing these models on historical sales data and assessing their performance with respect to MAPE, SMAPE, and comparable indicators. This study also looks at the circumstances in which traditional models can perform better than modern machine learning techniques and offers guidance on choosing the best prediction models depending on the properties of the data and the needs of the prediction Wiyanti et al. (2021). The findings from this study will have important implications for both theory and practice. From the point of view of theory, this study will add to the body of literature by comparing and contrasting modern and classic forecasting techniques within a particular setting. In reality, the findings will help supply chain specialists select the most effective forecasting methods, enhancing productivity and decision-making procedures. The study will also pinpoint important variables influencing various models' performance and offer a thorough comprehension of the circumstances and reasons behind why some models outperform others Khan et al. (2020). As a result, by offering useful knowledge and theoretical advancements that will be helpful in both academic research and commercial applications, this study will attempt to clarify the issues that must be taken into account while forecasting demand in the supply chain procedure. This study intends to enhance supply chain management and business outcomes by improving the accuracy and efficiency of demand forecasting by a thorough examination of different forecasting models.

1.1 Research Question and Objectives

Research Question: How can the accuracy of demand forecasting in supply chain management be enhanced by evualating the effectiveness of ARIMA and LSTM models in optimizing inventory levels and improving operational efficiency?

Objectives:

- **Evaluate Model Performance**: Compare the accuracy of ARIMA and LSTM models on historical sales data, specifically evaluating their performance using metrics such as Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE).
- **Optimize Inventory Management**: Provide supply chain professionals with insights on choosing the most effective forecasting model based on product demand characteristics, with a focus on high- and low-sales products.
- **Optimize Operational Efficiency**: Identify the model that minimizes forecasting errors and contributes to reducing operational inefficiencies and financial losses in supply chain management.

These goals aim to strengthen the work between supply chain forecasting theory and practice by utilizing state-of-the-art analytical techniques and improve business outcomes.

2 Related Work

Due to its major impact on customer retention, waste and overstocking avoidance techniques, and overall profitability, demand forecasting has drawn a lot of attention. Many methodologies have been proposed, put into practice, and discussed over time; these range from cutting-edge machine learning approaches to conventional statistical procedures.

2.1 Sustainability in Supply Chain Management

In order to enhance decision-making and operational efficiency, numerous research use machine learning approaches to assess and manage sustainability performance in supply chains and corporate operations. Studies underscore the significance of sustainability in supply chain management and its influence on the entire performance of businesses Ageeli et al. (2023). Creating sustainable supply chain practices that align with the Sustainable Development Goals (SDGs) of the UN, which are designed to enhance company performance and sustainability, is also crucial. Numerous scientific data have been gathered for this application, which is part of the "Business Process Framework for Sustainability" project and is being conducted at the University of Salerno. This study assessed the effects of distribution and purchasing operations on market and organizational performance. The study emphasizes the importance of supply chains in accomplishing the Sustainable Development Goals and unequivocally shows a positive association between sustainable practices and supply chain performance, it also emphasizes how important it is for supply chain participants to work together and how sustainable practices may give businesses a competitive edge Cammarano et al. (2022). By thoroughly addressing sustainable practices that are directly related to the Sustainable Development Goals, our research demonstrates the significance of our effort.

2.2 Comparative Analysis of Forecasting Models

Recently, many studies have tested and compared traditional and modern time series forecasting methods. According to these studies, how the datasets used are tested greatly affects the results because datasets have unique trends. While some data sets exhibit linear movements and are easier to predict, others have more complex and difficult-to-understand trends. For this reason, more advanced and complex machine learning algorithms or hybrid models have been frequently tried and recommended to improve the prediction accuracy of such datasets. The challenge of projecting demand for recently introduced products—for which prior sales data are frequently unavailable—is the subject of another study. To forecast demand patterns and measure uncertainty surrounding new items, the authors suggest a brand-new technique called DemandForest, which combines K-means clustering, Random Forest, and Quantile Regression Forest. To provide pre-launch forecasts, this hybrid machine learning approach leverages product attributes of both new and existing items, as well as historical sales data of previously released products Van Steenbergen and Mes (2020). The research uses real-world datasets from numerous businesses in a range of industries, such as retail, e-commerce, and wholesale, to assess the DemandForest approach. Based on delivery timings and service levels, DemandForest's estimates are more accurate than those of other benchmarking techniques, which might lead to inventory savings of about 15 percent Van Steenbergen and Mes (2020). By fitting theoretical distributions to quantiles, the suggested method expands upon the Quantile Regression Forest and improves prediction reliability. DemandForest not only increases prediction accuracy but also gives supply chain planners useful information by ranking similar goods and weighing the value of features Van Steenbergen and Mes (2020). This strategy facilitates better informed decision-making in inventory control and, in the end, aids in the successful launch of new products.

2.3 Deep Learning and AI in Demand Forecasting

A different investigation looks into Supply Chain Management 4.0 (SCM 4.0) and how deep learning techniques can be used for demand forecasting. The present study discusses the function of artificial intelligence (AI) in converting conventional supply chains into intelligent, cooperative, and interactive networks. In order to forecast future client requests, the study investigates the use of deep learning models—more particularly, Autoregressive Integrated Moving Average (ARIMA) and Long Short Term Memory (LSTM) networks—using transaction records from the past. The study draws attention to the shortcomings of traditional statistical techniques, which frequently fall short in handling complexity and demand variations Terrada et al. (2022). These restrictions are highly variable, and every day that a newborn is born brings with it a new, unanticipated restriction. In contrast, artificial intelligence and deep learning techniques widen the perspective of human knowledge and improve the quality of performance by exploiting enormous data sets and capturing non-linear relationships in data. The study compares the performance of ARIMA and LSTM models using a Kaggle dataset that contains monthly demand data for different items. Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) are examples of evaluation metrics Terrada et al. (2022). Later on in the review, we'll take a closer look at these evaluation metrics. The findings show that the LSTM model performs noticeably better than the ARIMA model, resulting in reduced error rates and increased demand forecasting accuracy Terrada et al. (2022). This research another prove that how deep learning techniques have a lot of promise for enhancing the effectiveness of demand forecasting systems in SCM 4.0.

2.4 Hybrid and Modern Approaches in Demand Forecasting

The goal of the project is to use artificial intelligence-based forecasting models to enhance supply chain performance and lessen the bullwhip impact, which will enhance decision-making and operational effectiveness. Here we see that the more modern machine learning model is clearly more successful than traditional methods. However, since this situation still does not satisfy us, let's continue to examine other studies. While we're talking about long short-term memories, a study in retail supply chain management examined the predicted accuracy of LSTM and seasonal auto-regressive integrated moving average (SARIMA) models. In order to determine which model performed better under various demand scenarios, the study examined over 37 months of actual retail sales data from an Austrian shop. Although SARIMA produced better results for products with seasonal demand patterns—which is not surprising given that this is precisely the context for which SARIMA is designed—the study discovered that the LSTM model generally performed better for products with steady demand Falatouri et al. (2022). Furthermore, adding outside variables like promotions using SARIMAX greatly increases prediction accuracy, particularly for products that are marketed Falatouri et al. (2022). The outcomes indicate that the prediction quality can be further enhanced by hybrid techniques that combine LSTM and SARIMA(X) models trained on pre-clustered store groups. This demonstrates how crucial it is to select hybrid algorithms or the best machine learning strategy based on the available data set rather of just one method. We investigate the impact of machine learning (ML) demand forecasting techniques on supply chain performance, concentrating on a hybrid strategy that combines ARIMAX and Artificial Neural Networks. In order to tackle the previously mentioned issue of operational inefficiency caused by variance amplification in multi-stage supply chains, also referred to as the bullwhip effect. The study compares ML-based and traditional forecasting techniques using data from a steel manufacturer, assessing the effects of each on forecast accuracy, inventory turns, and cash conversion cycles Feizabadi (2022). The results demonstrate how much supply chain performance is enhanced by ML-based techniques since they are more adept at managing intricate interdependencies and nonlinear interactions in the data. This improvement shows that the forecast error has decreased statistically significantly Feizabadi (2022).

The challenges posed by the COVID-19 pandemic, which has resulted in changing customer demand and severe unpredictability in the business climate, served as an inspiration for another similarly driven investigation. In order to increase forecast accuracy, this study suggests an improved LSTM model that adjusts hyperparameters via grid search. The performance of the suggested LSTM model is evaluated against other machine learning models, such as Recurrent Neural Networks (RNN), and conventional statistical techniques, such as Exponential Smoothing (ETS) and ARIMA, using past sales data from a Moroccan pharmaceutical manufacturing company El Filali et al. (2022). It's interesting to note that the optimized LSTM model performs better than the other approaches, obtaining lower RMSE and SMAPE values, proving a greater capacity to capture non-linear characteristics in time series data. The demand in the manufacturing industry is influenced by a wide range of intricate and diversified elements, including government subsidies, international collaboration, economic development, service levels, and technical advancements. In a study, LSTM networks were used to estimate future demand by building an index system with these factors. According to the study, LSTM networks performed better in terms of accuracy when measured by MAE,RMSE and MAPE than conventional techniques like Support Vector Machines (SVM), Backpropagation (BP) neural networks, Random Forest (RF), and autoregressive (AR) models Dou et al. (2021). It was determined that the LSTM model is especially useful for this application because of its capacity to manage nonlinear interactions and long-term dependencies Dou et al. (2021). It has been demonstrated that, in comparison to using historical demand data alone, forecast accuracy is greatly increased when numerous influencing factors are integrated into the forecast model. Though it seems idealistic, it is acknowledged that perfect prediction accuracy can be attained if we know every effect. However, as this is not achievable, we must incorporate all of the factors that we know exist and whose effects would produce particular outcomes into the algorithm. Another proposed strategy for efficient demand forecasting in enterprises is the integration of machine learning (ML) and business intelligence (BI). When we consider how business intelligence (BI) helps turn unstructured data into insightful knowledge, we find that it expands the scope of strategic decision-making and inevitably boosts operational effectiveness. The Deep AR algorithm in Amazon SageMaker, a supervised learning technique for time series forecasting that performs very well when handling a huge number of time series, is used in this model Khan et al. (2020). The model provides high demand forecasting accuracy by utilizing BI and ML, which is essential for lowering supply chain and inventory management expenses. The suggested approach can reach up to 92.38 percent accuracy, according to simulation findings, which is a major improvement over manual and Excel-based forecasting techniques Khan et al. (2020). In this literature review, while understanding the importance of demand forecasting, it is not only the sales, purchases and supply chains that are involved. For example, another research focuses on deep learning techniques to forecast electricity consumption. The research highlights the importance of smart grids that enable two-way communication between consumers and suppliers, thus facilitating efficient energy use and demand response. It makes use of a variety of deep learning models, most notably LSTM networks, because of how well they estimate energy consumption Aguiar-Pérez and Pérez-Juárez (2023). The study divides the period or forecast horizon into short-term (usually from one hour to one week), medium-term (usually from one week to one year) and long-term (usually more than one year) forecasts and processes its work on LSTM Aguiar-Pérez and Pérez-Juárez (2023). Furthermore, it reaffirms the significance of big data in enhancing smart grid operations and highlights the necessity of precise load forecasting to balance supply and demand, particularly during peak hours. As can be seen, many studies have made comparative analyses of traditional methods and modern methods. However, the high accuracy demand forecasting results of many studies on LSTM or hybrid methods combined with LSTM are quite intriguing. For this reason, let's take a closer look at these studies in the rest of the review.

2.5 LSTM Networks in Demand Forecasting

Let's first take a quick look at how LSTM functions in demand forecasting. With the use of input and output gates, forgetting, and keeping just the most crucial information, LSTM networks analyze historical demand data to forecast future demand. Through this process, long-term connections and correlations can be learned by LSTMs, which they can then use to their advantage to produce predictions that are more accurate. In a highly competitive corporate environment, the multilayer LSTM network is applied for demand forecasting. Grid search is used in a forecasting method to automatically determine which set of LSTM hyperparameters is optimal for a given time series, allowing nonlinear patterns in non-stationary data to be captured Abbasimehr et al. (2020). The proposed LSTM model with demand data is therefore compared with a number of well-known forecasting techniques, such as ARIMA, Exponential Smoothing (ETS), Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), Recurrent Neural Networks (RNN), Support Vector Machines (SVM), and single-layer LSTM Abbasimehr et al. (2020). Data from a furniture company is used in this process. The RMSE and SMAPE evaluation metrics are applied. The outcomes demonstrate how well the multilayer LSTM model works in comparison to alternative techniques and how well it can forecast demand data that varies greatly Abbasimehr et al. (2020). This shows a lot of promise. Our investigation also reveals that several studies assert that combining CNN and LSTM yields more accurate models. This hybrid model's benefit is that by taking into account both long-term trends and short-term patterns, it can produce predictions that are more accurate. The LSTM layer comprehends how these properties change over time, whereas the CNN layer rapidly extracts important aspects from the data. The model increases demand prediction accuracy in this way. With the goal of offering a more reliable and accurate forecasting framework, the hybrid CNN-LSTM model is made to collect both temporal dependencies from LSTM and spatial variables from CNN. As a result, a study makes use of historical sales data from a pharmacy covering six years (2014–2019) and 57 pharmaceutical products across eight categories MK et al. (2023). An hourly time series with Anatomical Therapeutic Chemical (ATC) classification features is created by pre-processing and resampling the data MK et al. (2023). Metrics like Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) are used to compare the model's performance to a number of conventional and deep learning models, such as Vanilla LSTM, Stacked LSTM, and Bidirectional LSTM MK et al. (2023). The findings demonstrate that the CNN-LSTM hybrid model performs better in terms of accuracy than the other models, particularly for certain medication categories. The study demonstrates how the model can reduce noise and variability in the input data, increasing its flexibility and producing demand estimates that are nearly perfect.

2.6 CNN and BiLSTM in Demand Forecasting

Another approach that has a similar reasoning as the previous one aims to introduce a hybrid deep learning framework for store product demand forecasting that combines CNN and BiLSTM networks. They point out the difficulties faced by retailers in accurately predicting demand in order to prevent problems like overstocking and stock-outs. The suggested model makes use of both the temporal sequence learning skills of BiLSTM and the feature extraction capabilities of CNN that have been tuned with the Lazy Adam optimizer Joseph et al. (2022). Using sales data from 50 products across 10 locations over a five-year period, Kaggle's Store Product Demand Prediction Challenge dataset is utilized in the study Joseph et al. (2022). Several cutting-edge machine learning techniques, such as XGBoost, Random Forest, K-Nearest Neighbors (KNN), Bagging, Support Vector Regressor (SVR), Stochastic Gradient Descent (SGD), Linear Regression, and CNN-LSTM, are compared to the performance of the proposed CNN-BiLSTM model. These major findings are proof of the power of hybrid models. However, we see that different sectors reach more accurate demand forecasts with completely different approaches and interests. This supports our claim that there is no perfect algorithm or technology, and that the success of demand forecasting depends on a good understanding of the business and demographics. We have seen many machine learning methods so far, but of course they are not limited to this. Let's also take a look at the studies on XGBoost and Decision Trees.

2.7 XGBoost and Decision Trees

Clustering and time series models are integrated in a study on the creation of a three-stage XGBoost-based model for sales forecasting in global e-commerce Ji et al. (2019). The suggested C-A-XGBoost model includes XGBoost to handle nonlinearities, an ARIMA model to handle linear patterns, and clustering to handle sales characteristics the goal of this combination is to increase sales projections' precision and effectiveness Ji et al. (2019). Tested on data from a global e-commerce platform, the results are encouraging, demonstrating that the C-A-XGBoost model delivers notable gains in predicting accuracy and beats solo XGBoost and standard ARIMA models Ji et al. (2019). Of course, XGBoost can be used in many areas, not just for demand forecasting. For example, one study approaches our topic in an indirect way. Using the XGBoost method, an effort was made to create a fraud prediction model for supply chains. With an F1 score of 99.31, the model showed better classification abilities than the Logistic Regression and Gaussian Naive Bayes models Zhou et al. (2021).

2.8 Evaluation Metrics for Demand Forecasting

Let's examine the logic behind the usage of metrics such as Mean Absolute Percentage Error and others to assess the accuracy of demand forecasting models. By calculating the average of the squares of the mistakes, the Mean Squared Error (MSE), which measures how much the predicted values differ from the actual values, penalizes significant errors more severely. The square root of the mean square error, or root mean square error (RMSE), makes understanding simpler by providing the average magnitude of the mistakes in the same units as the projected values. The statistic known as Mean Absolute Error (MAE) is easier to comprehend and is less impacted by extreme errors. It is calculated by taking the average of the absolute values of the errors. In contrast, the Mean Absolute Percentage Error (MAPE) calculates the ratio of the difference between the actual and predicted values to the actual value by taking the average of the mistakes and expressing it as a percentage. It makes it simple to compare forecasts at various scales because it is expressed as a percentage. When evaluating the precision of demand forecasting models, these metrics are outstanding. The goal of demand forecasting is to precisely project future consumer demand for goods or services. These metrics are a great way to evaluate how well the forecasts are working. While severe errors can be found and fixed with the aid of MSE and RMSE, the more general and intelligible error rates provided by MAE and MAPE paint a clearer picture of the model's performance. For this reason, these metrics are suitable for evaluating demand forecasting accuracy.

Managing supply chains that facilitate communication between the producer, distributor, and retailer by forecasting demand will ultimately help to prevent excess inventory and waste, as we approach the conclusion of our literature review. This will not only cut expenses but also allow for more sustainable environmental practices and the effective use of resources. It therefore serves as the fundamental foundation for cutting waste and promoting a cleaner environment.

2.9 Conclusion

Studies on demand forecasting for supply chain management encompass a wide range of models, from more contemporary machine learning techniques like LSTM networks to more conventional statistical methods like ARIMA. Because traditional methods like ARIMA and SARIMA can predict linear patterns in historical data, they are commonly utilized for demand forecasting. Nevertheless, they frequently struggle to manage the complex and nonlinear relationships that are more common in contemporary supply chain settings. Machine learning models like LSTM networks, XGBoost, and Decision Trees—which provide greater accuracy and flexibility by automatically learning from enormous datasets—have been the subject of more recent research. These models have also produced hybrid models. Particularly LSTM networks have demonstrated a great deal of promise in identifying nonlinear patterns and long-term dependencies in sales data, which qualifies them for more difficult forecasting jobs.
Though a lot of research has been done to compare these models, little attention has been paid to real-world implementations for various product categories, particularly when it comes to low-demand products. Furthermore, a number of studies offer comprehensive instructions on how to choose the best model in accordance with the features of the data and particular forecasting requirements. Our study closes these gaps by comparing the effectiveness of the ARIMA and LSTM models in various circumstances by a thorough analysis of historical sales data. By concentrating on low-selling products, one can gain important insights into how these models might be used to maximize supply chain efficiency, cut waste, and enhance inventory management. Also, this study offers helpful guidance to supply chain experts, helping them choose the best forecasting model depending on certain data attributes and forecasting requirements.

3 Methodology

This research follows the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, which is quite popular for data mining projects. CRISP-DM makes sure that the data mining process moves forward systematically at different stages and makes the activities carried out easier to understand. Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment are its six steps. This process ensures a thorough analysis, which is what the research does Martínez-Plumed et al. (2021). The exact tasks and actions carried out at each stage of the CRISP-DM process will be thoroughly explained in the sections that follow, along with examples of how this methodology is used to accomplish the research goals.

3.1 Choosing an Appropriate Dataset for Our Study

In supply chain management, stock control is crucial. Achieving optimal stocking levels is essential for efficient operations. We are seeking a dataset suitable for time series analysis to forecast demand accurately. This dataset should contain historical data, making it ideal for our analysis.

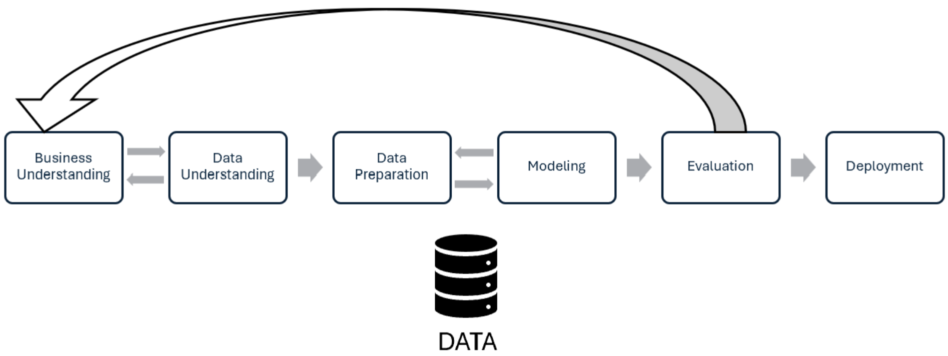

Figure 1: Basics of CRISP-DM Martínez-Plumed et al. (2021)

About the Dataset

For our research, we have chosen a comprehensive dataset that provides detailed insights into US regional sales data across various sales channels. This dataset is publicly available on Kaggle under the CC0: Public Domain license, making it an ideal resource for our analysis[1]. This dataset is particularly suited for our needs as it encompasses a wide range of information essential for analyzing sales patterns, trends, and forecasting demand. The dataset contains 17,992 rows and 15 columns, covering numerous aspects of sales transactions and customer interactions. The key columns relevant to our research include Sales Channel, ProcuredDate, DeliveryDate, Order Quantity, Unit Cost, and Unit Price.

3.2 Data Preprocessing and Feature Engineering

The following stage is data preparation, which comes after understanding the nature of the dataset. This procedure must include cleaning, and any columns that have no bearing on the analysis should be eliminated. To prepare the dataset for machine learning models, data transformation is required after cleaning.

First, the dataset containing US regional sales data is loaded into a pandas DataFrame. The cleaning process starts by dropping unnecessary columns such as ProcuredDate, OrderDate, and ShipDate, as we are going to use only the Delivery Date column, which is retained because it shows the closing of sales. Additionally, columns like WarehouseCode, Unit Cost, and CurrencyCode are also removed to focus only on the columns relevant to our analysis. In order to guarantee correct numerical operations, the Unit Price column is then cleaned up by eliminating any commas and changing the values to a float type. The Order Quantity and the Unit Price are then multiplied to produce a new column called TotalSales. Crucial to our research is the direct measure of sales revenue provided by each order in this column.

Ultimately, the initial rows of the dataset are examined to validate the modifications and identify any missing values to guarantee the accuracy of the data. Resolving missing values is essential to preserving the analysis's integrity. The dataset is cleaned and modified by completing these procedures, preparing it for additional analysis and forecasting.

[1] `https://www.kaggle.com/datasets/talhabu/us-regional-sales-data`

Please refer to the configuration guide document included with this study for further information on these cleaning procedures.

3.3 Applying Models

Modeling is the next step after finishing data preparation. Before applying the model, we need to split the data into training and test sets. This allows the performance of the model to be evaluated on unseen data. To split the data in our research, we start by finding the products with the lowest total sales among the 47 products and separating them from the others. Then, we will use these 5 products as test data according to the monthly order quantities they receive. Accordingly, we need to perform out-of-sample validation. We will compare the last 30 days separately to see how accurately the future values are predicted. In this way, if we can achieve high accuracy, the production of products with low demand will be reduced and waste and excess stock will be avoided. After the data is split, we can apply and run the model. For this study, interactive coding environments for Python were used to clean and model the data.

This study is based on two different machine learning models and requires the models to be modified and run repeatedly for neural networks and other traditional machine learning models. More details about this step can be found in the configuration guide document attached to this study. The models applied to the dataset are as follows:

AutoRegressive Integrated Moving Average

Three parts make up the ARIMA model, which is utilized for time series forecasting:

- **AutoRegressive (AR) component:** Models the time series using historical values by utilizing the dependence between an observation and several lagged (prior) observations.
- **Integrated (I) component:** Entails subtracting the observations in order to remove seasonality and trends, thus making the time series stationary.
- **Moving Average (MA) component:** Applying a moving average model to lagged observations, this component uses the dependence between an observation and a residual error.

In ARIMA models, the notation ARIMA(p, d, q) represents the autoregressive section (p number of lag observations), the integrated part (d number of differences required to make the series stable), and the moving average window size (q moving average part).

Long Short-Term Memory Networks

LSTM networks are a type of recurrent neural network (RNN) used for sequence prediction. They can preserve long-term dependencies in data, making them suitable for time series prediction.

After the data preparation section, the data. Products are sorted by total sales in ascending order to identify the five products with the least total sales. The dataset is filtered to include only these products, and the Delivery Date is converted to date/time format. The data is resampled to a daily frequency, and missing days are filled with zero order quantity. This results in a time series suitable for analysis.

Used Evaluation Metric: Mean Squared Error (MSE)

The method for calculating mean square error (MSE) involves taking the difference between each observation's actual and predicted values, squareing it, adding up all of the squared differences, and then dividing the result by the total number of observations Qi et al. (2020).

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Figure 2: MSE Formula Qi et al. (2020)

- n is the number of observations,
- y_i is the actual value for the i-th observation,
- $\hat{y}_i$ is the predicted value for the i-th observation,
- $(y_i - \hat{y}_i)^2$ is the squared difference between the actual and predicted values for the i-th observation Qi et al. (2020).

4 Design Specification

The framework we have created is designed to positively improve the supply chains of businesses by using both traditional models and neural networks in machine learning models. Figure 3. shows the basic steps of the framework's flow. This work ensures that by adhering to the framework we have constructed, the best application of data science and machine learning to supply chains is achieved and each stage advances the main goal of reducing waste and providing insightful business insights for green logistics.

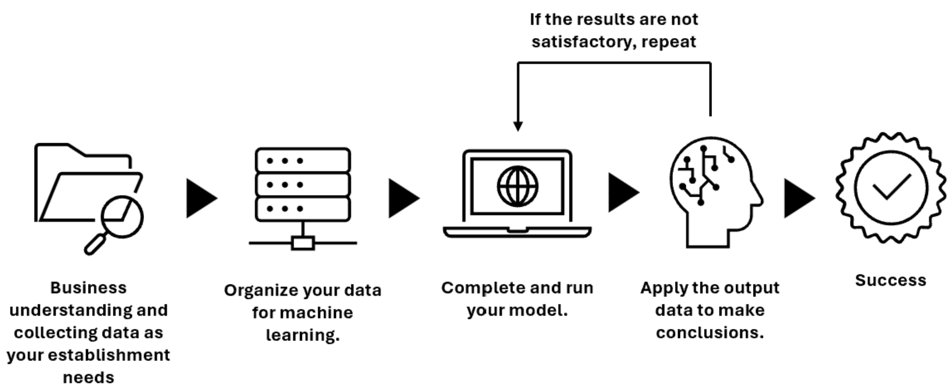

Figure 3: Steps of the framework established for an improved supply chain

5 Implementation

This section details the implementation of ARIMA and LSTM models for forecasting order quantities using the prepared dataset.

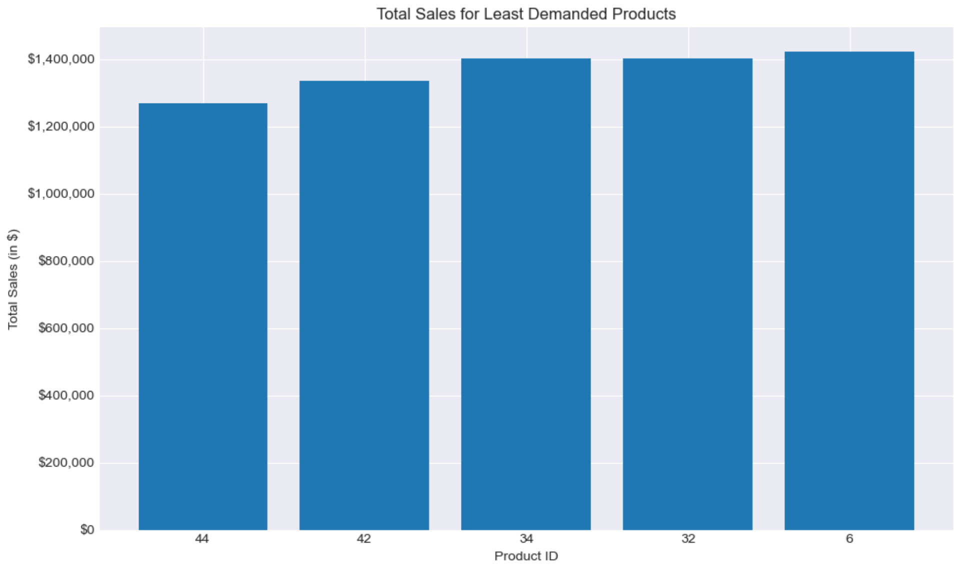

Figure 4: Least Demanded Products

5.1 ARIMA

In this study, the ARIMA model was used to predict the order quantities of one of the least-sold products in the next 30 days. You can see these products in Figure 4. The data is filtered and split into training and test sets. The training set contains all data except the last 30 days, which are reserved for testing.

A rolling forecast approach is used, where the ARIMA model is trained on the historical data and forecasts are generated for the next day. This process is repeated for each day in the test set. The Mean Squared Error (MSE) is calculated to evaluate the accuracy of the predictions. The MSE results for the ARIMA model for various products are as follows in table 1.

Product ID	MSE
32	4.378866337300955
44	3.301619571778478
42	6.709885578182906
34	6.541816452367126
6	0.2991129174585303

Table 1: MSE Results for Least Demanded Products

Figure 5. also provides a visual representation of the data and estimations for product number 44 during the last six months. Table 2. contains the list of estimated and actual values.

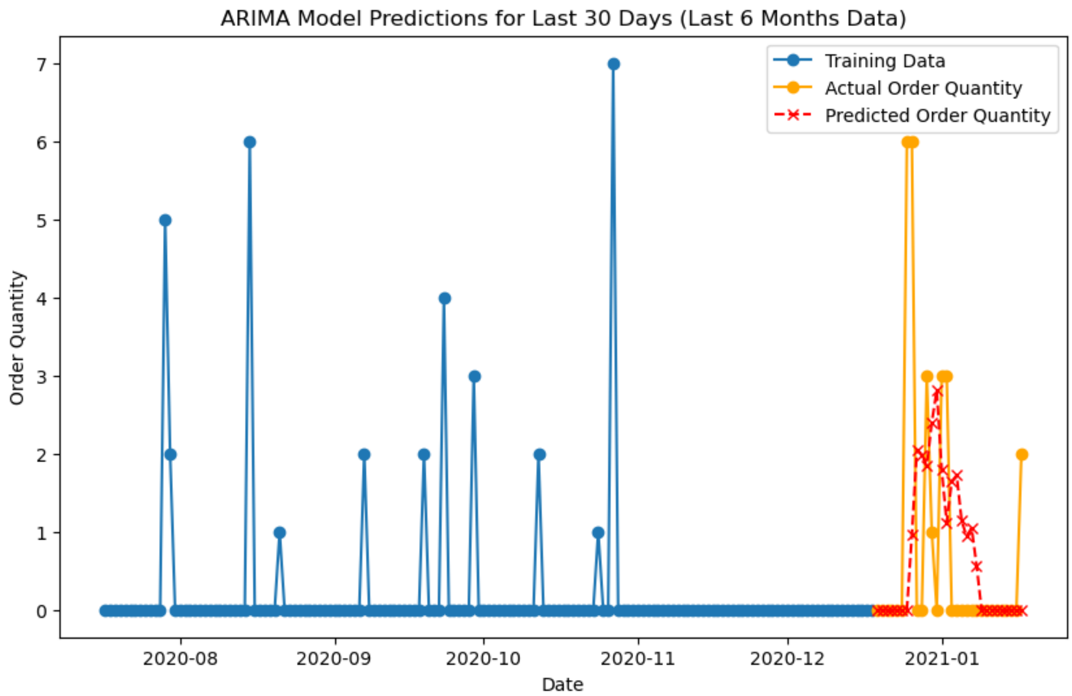

Figure 5: Product 44 Predictions

Delivery Date	Product ID	Real Order Quantity	Predicted Order Quantity	Prediction Error
2020-12-23	44	0	1.689225e-151	-1.689225e-151
2020-12-24	44	0	2.093674e-152	-2.093674e-152
2020-12-25	44	6	-1.777148e-153	6.000000e+00
2020-12-26	44	6	9.581187e-01	5.041881e+00
2020-12-27	44	0	2.052547e+00	-2.052547e+00
2020-12-28	44	0	1.985347e+00	-1.985347e+00
2020-12-29	44	3	1.838623e+00	1.161377e+00
2020-12-30	44	1	2.398183e+00	-1.398183e+00
2020-12-31	44	0	2.809966e+00	-2.809966e+00
2021-01-01	44	3	1.788582e+00	1.211418e+00
2021-01-02	44	3	1.115694e+00	1.884306e+00

Table 2: Real and Predicted Order Quantities

5.2 LSTM Networks

LSTM networks are used for sequence prediction, taking advantage of their ability to maintain long-term dependencies in data. An LSTM model is built using Keras, with two LSTM layers followed by a Dense layer. The model is trained on the training data for 20 epochs. After training, the performance of the model was evaluated on the test set by calculating the MSE between the predicted and actual order quantities. For the LSTM model, the MSE for 5 products is 0.511625269686524. Tables 3. and 4. show us the actual and predicted values of the data obtained from the training.

Table 3: Actual Values for Different Products Over Time

Delivery Date	Actual 6	Actual 32	Actual 34	Actual 42	Actual 44
2020-07-28	0.0	0.0	0.0	0.0	0.0
2020-07-29	0.0	0.0	0.0	4.0	5.0
2020-07-30	0.0	8.0	0.0	0.0	2.0
2020-07-31	0.0	0.0	0.0	0.0	0.0
2020-08-01	0.0	0.0	0.0	6.0	0.0
2021-01-25	0.0	6.0	0.0	0.0	0.0
2021-01-26	0.0	0.0	0.0	0.0	0.0
2021-01-27	0.0	7.0	0.0	0.0	0.0
2021-01-28	0.0	0.0	0.0	0.0	0.0
2021-01-29	3.0	0.0	0.0	0.0	0.0

Table 4: Predicted Values for Different Products Over Time

Delivery Date	Predicted 6	Predicted 32	Predicted 34	Predicted 42
2020-07-28	1.118555	0.762427	1.099517	2.337818
2020-07-29	1.322292	0.920383	1.084734	2.139064
2020-07-30	1.203414	1.044072	1.014103	1.258541
2020-07-31	1.789927	0.606270	1.120349	1.885290
2020-08-01	0.434536	0.247355	0.347052	0.240754
2021-01-25	0.575980	0.540213	0.790076	0.505662
2021-01-26	1.050043	0.258342	0.842316	1.256573
2021-01-27	0.319773	0.131388	0.352148	0.374458
2021-01-28	0.406475	0.039801	0.496412	1.029358
2021-01-29	-0.217089	0.246520	0.287549	0.371810

6 Evaluation

This is the evaluation part, where we discuss how well the models performed in predicting order quantities. The objective is to establish a logical debate environment and provide a clear evaluation of the results.

6.1 ARIMA Model Evaluation

ARIMA model was used to predict future order quantities for the 5 lowest selling products. Data for each product was filtered and divided into training and test sets, with the last 30 days reserved for testing. ARIMA model predictions were compared with actual values and Mean Square Error was calculated to evaluate the model accuracy.

MSE values for each product show the difference between predicted and actual order quantities during the testing period. It is obvious that lower MSE values indicate better model performance. For example, Product ID 6 had the lowest MSE of 0.2991 indicating high accuracy, while Product ID 42 had a higher MSE of 6.7099 indicating lower accuracy. These results highlight the variable performance of the model across different products. This imbalance significantly reduced the confidence in the model, which led us to explore different models.

6.2 LSTM Networks Model Evaluation

LSTM model was also used to predict order quantities by taking advantage of its ability to handle long-term dependencies in the data. The model was built using Keras with two LSTM layers followed by a Dense layer. The model was trained on the training data for 20 epochs and then evaluated on the test set.

The MSE for the LSTM model was calculated as 0.5116, indicating the accuracy of the model in predicting order quantities. Here, instead of predicting the future order figures of 5 products separately, the order quantities of 5 products were predicted together. With these values, it is possible to say that the LSTM model performs better than ARIMA.

6.3 Discussion

The implementation of LSTM and ARIMA models produced results that demonstrated how accurately they predicted order quantities. Different products performed differently according to the ARIMA model, with some achieving lower MSE values and others reaching higher levels. This variation implies that the unique features of each product's sales data may have an impact on how effective the ARIMA model is. Nevertheless, depending on the circumstances of the task at hand, ARIMA can still be utilized even though it is not trusted in this instance. On the other hand, the LSTM model handled the complexity of the dataset better and had a more consistent and lower MSE. It is clear that the LSTM model outperformed ARIMA in terms of maintaining long-term dependencies and detecting non-linear correlations in the data; however, it would be an exaggeration to call either model the best. While each model has advantages, the choice of model may be influenced by the particulars of the data as well as the needs for prediction. While the LSTM model is excellent at capturing intricate, non-linear relationships, the ARIMA model works best with data that has linear patterns.

The findings emphasize how crucial it is to choose the right forecasting model depending on the attributes of the data. In-depth examination and graphic representations offer significant perspectives for enhancing inventory control and demand forecasting within supply chains. Businesses may choose the best model to utilize for a given set of forecasting requirements by knowing the advantages and disadvantages of each model.

7 Conclusion and Future Work

This study successfully applied ARIMA and LSTM models to forecast order quantities, demonstrating their respective strengths and limitations. The key findings indicate that while the LSTM model generally provides more accurate predictions, the ARIMA model may still be valuable for specific scenarios with simpler data patterns.

Future work could explore the integration of additional machine learning models and hybrid approaches to further enhance forecasting accuracy. Additionally, incorporating external factors such as market trends, promotional activities, and economic indicators could provide a more comprehensive demand forecasting framework.

Businesses may optimize inventory levels, minimize waste, and improve overall supply chain efficiency by increasing the accuracy of their demand forecasts. The knowledge gathered from this research assists in the continuous creation of successful demand forecasting plans for the wholesale and retail industries. In order to get the best outcomes, the study also emphasizes the possible advantages of mixing classic and new forecasting models.

References

Abbasimehr, H., Shabani, M. and Yousefi, M. (2020). An optimized model using lstm network for demand forecasting, *Computers & industrial engineering* **143**: 106435.

Ageeli, A. A. et al. (2023). Evaluating and managing sustainability performance of supply chain and business process management: An integrated and applied approach, *Full Length Article* **3**(2): 25–5.

Aguiar-Pérez, J. M. and Pérez-Juárez, M. Á. (2023). An insight of deep learning based demand forecasting in smart grids, *Sensors* **23**(3): 1467.

Cammarano, A., Perano, M., Michelino, F., Del Regno, C. and Caputo, M. (2022). Sdg-oriented supply chains: business practices for procurement and distribution, *Sustainability* **14**(3): 1325.

Dou, Z., Sun, Y., Zhang, Y., Wang, T., Wu, C. and Fan, S. (2021). Regional manufacturing industry demand forecasting: A deep learning approach, *Applied Sciences* **11**(13): 6199.

El Filali, A., Lahmer, E. H. B., El Filali, S., Kasbouya, M., Ajouary, M. A. and Akantous, S. (2022). Machine learning applications in supply chain management: A deep learning model using an optimized lstm network for demand forecasting., *International Journal of Intelligent Engineering & Systems* **15**(2).

Falatouri, T., Darbanian, F., Brandtner, P. and Udokwu, C. (2022). Predictive analytics for demand forecasting–a comparison of sarima and lstm in retail scm, *Procedia Computer Science* **200**: 993–1003.

Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance, *International Journal of Logistics Research and Applications* **25**(2): 119–142.

Ji, S., Wang, X., Zhao, W. and Guo, D. (2019). An application of a three-stage xgboost-based model to sales forecasting of a cross-border e-commerce enterprise, *Mathematical Problems in Engineering* **2019**(1): 8503252.

Joseph, R. V., Mohanty, A., Tyagi, S., Mishra, S., Satapathy, S. K. and Mohanty, S. N. (2022). A hybrid deep learning framework with cnn and bi-directional lstm for store item demand forecasting, *Computers and Electrical Engineering* **103**: 108358.

Khan, M. A., Saqib, S., Alyas, T., Rehman, A. U., Saeed, Y., Zeb, A., Zareei, M. and Mohamed, E. M. (2020). Effective demand forecasting model using business intelligence empowered with machine learning, *IEEE access* **8**: 116013–116023.

Kilimci, Z. H., Akyuz, A. O., Uysal, M., Akyokus, S., Uysal, M. O., Atak Bulbul, B. and Ekmis, M. A. (2019). An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain, *Complexity* **2019**(1): 9067367.

Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Orallo, J. H., Kull, M., Lachiche, N., Quintana, M. J. R. and Flach, P. (2021). Crisp-dm twenty years later: From data mining processes to data science trajectories, *IEEE Transactions on Knowledge and Data Engineering* **33**: 3048–3061.

MK, N. P., Rastogi, S. and Alice, K. (2023). Demand forecasting in supply chain management using cnn-lstm hybrid model, *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, pp. 1–5.

Qi, J., Du, J., Siniscalchi, S. M., Ma, X. and Lee, C.-H. (2020). On mean absolute error for deep neural network based vector-to-vector regression, *IEEE Signal Processing Letters* **27**: 1485–1489.

Terrada, L., El Khaili, M. and Ouajji, H. (2022). Demand forecasting model using deep learning methods for supply chain management 4.0, *International Journal of Advanced Computer Science and Applications* **13**(5).

Van Steenbergen, R. and Mes, M. R. (2020). Forecasting demand profiles of new products, *Decision support systems* **139**: 113401.

Wiyanti, D., Kharisudin, I., Setiawan, A. and Nugroho, A. (2021). Machine-learning algorithm for demand forecasting problem, *Journal of Physics: Conference Series*, Vol. 1918, IOP Publishing, p. 042012.

Zhou, Y., Song, X. and Zhou, M. (2021). Supply chain fraud prediction based on xgboost method, *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, IEEE, pp. 539–542.