Supply Chain Management on Retail Store Inventory Forecasting

Student’s Name

Date

**Table of Contents**

[1. Introduction 3](#_Toc193692445)

[1.1 Overview and Purpose of the Project 3](#_Toc193692446)

[1.2 Research Question, Aim and Objectives 4](#_Toc193692447)

[1.2.1 Research Question 4](#_Toc193692448)

[1.2.2 Aim 4](#_Toc193692449)

[1.2.3 Objectives 5](#_Toc193692450)

[2. Background 5](#_Toc193692451)

[2.1 Introduction to Supply Chain Forecasting 5](#_Toc193692452)

[2.2 Selection Criteria for Papers 5](#_Toc193692453)

[2.3 Critical Analysis of Key Papers 6](#_Toc193692454)

[2.3.1 Paper 1: “AI-driven demand forecasting: Enhancing inventory management and customer satisfaction”. 6](#_Toc193692455)

[2.3.2 Paper 2: “Demand Forecasting in Supply Chain Management for Rossmann Stores using Weather Enhanced Deep Learning Model” 6](#_Toc193692456)

[2.3.3 Paper 3: “Retail forecasting: Research and practice” 7](#_Toc193692457)

[2.3.4 Paper 4: “Unlocking accurate demand forecasting in retail supply chains with AI-driven predictive analytics”. 8](#_Toc193692458)

[2.4 Summary of Literature Review 8](#_Toc193692459)

[3. Methodology 10](#_Toc193692460)

[3.1 Explanation of Features 11](#_Toc193692461)

[3.2 Data Type and Quality 12](#_Toc193692462)

[3.3 Summary of Project Methodology 12](#_Toc193692463)

[References 14](#_Toc193692464)

# 1. Introduction

## 1.1 Overview and Purpose of the Project

An effective supply chain management is essential since it allows optimal inventory levels and customer satisfaction. This project tackles data science approaches in integrating data science techniques into inventory forecasting, one of the crucial supply chain management cores, which facilitates retailers' forecast for future product demand, considering historical data and market trends (Qureshi *et al.*, 2024). This project aims to create a sound learning of how forecasting based on data can help improve the inventory process in a retail environment, achieving improvement in operational efficiency and quality of customer service. Demand planning, or inventory forecasting,, is the process of predicting the required number of stocks to supply customers' future demand. This is not just about figuring out how many sales will be generated; it involves a complete analysis of other things such as historical sales numbers, relevant market trends, seasonality and even external factors like economic conditions or supply chain disruptions (Ingle *et al.*, 2021). By using data science methods, retailers not only allow inventory levels to match market demand but also, instead of overestimating or underestimating, develop a better model of actual consumption and sales.

Theproject's first objective is to show how advanced analytical techniques can streamline the inventory forecasting method. Traditional forecasting techniques tend to take on a purely simplistic perspective and overlook the intricacies of consumer behaviour and market fluctuations. Data science presents a set of tools, from statistical analysis to machine learning algorithms, to work on vast volumes of data and unearthing patterns that could lead to a better line of thought (Mori, 2023). Inventory management is very tricky with balancing the deal with out-of-stock and overstock. A stockout happens with a product which is not available to the customers and smells like a sale leading to dissatisfaction among customers.

On the other hand, overstocking is used up capital using it on unsold goods and can lead to higher storage costs and possible markdowns (Tadayonrad and Ndiaye, 2023). In reducing these risks, effective inventory forecasting provides insights into how to replenish and when stock needs replenishment. This proactively provides a positive effect on customers and improves cash flow balance and improves cash flow balance by diminishing any excess inventory.

The data science integration into the inventory forecasting process equips retailers to see demand more nuancedly. History sales data, coupled with real-time information i.e. current market trends, and consumer behaviour, is used by retailers to anticipate fluctuations in demand. For example, the demand is such that it may spike unpredictably during peak shopping seasons or promotion days, and then retailers are equipped with a strong model to prepare for the surge (Falatouri *et al.*, 2022).

Moreover, modern inventory forecasting systems are also based on more advanced technologies, including artificial intelligence (AI) and machine learning (ML). These technologies allow data to be analysed that traditional methods would have otherwise missed traditional methods would have otherwise missed. For instance, dynamic forecasts, such as can be done by machine learning algorithms, can be dynamically changed due to learning from past sales patterns (Tian, Wang and Erjiang, 2021). With all the retail industry needs varying so rapidly in today's fast-paced retail environment, this adaptability remains essential. The second critical aspect of this project includes investigating different forecasting methods based on the retailer's context-based on the retailer's context. Sales data are used in quantitative approaches to forecasting future demand and qualitative methods use expert judgment and market research iinsights qualitativemethods use expert judgment and market research insights.

Focusing on the good aspects of these methods, a hybrid approach combining both quantitative and qualitative is more reliable in forecasting since it takes advantage of human and statistical trends (Amosu *et al.*, 2024). Bread also holds broader significance in implementing effective inventory forecasting practices to supply chain management as a whole. Better collaboration with suppliers and logistics partners is achieved by having accurate forecasts that enable stock levels to be in alignment with expected demand. Thus, this alignment streamlines the operations, reduces lead times and keeps the costs of emergency replenishments or expedited shipping in check.

## 1.2 Research Question, Aim and Objectives

### 1.2.1 Research Question

* How can machine learning models be used to predict retail store inventory demand through past sales, seasonal trends, and outside factors to manage and maintain supply chain efficiency?

### 1.2.2 Aim

The study aims to analyse managing the supply chain of the retail store by forecasting the inventory.

### 1.2.3 Objectives

* To develop and implement advanced data science techniques that improve the accuracy of inventory forecasts, enabling retail stores to align stock levels with actual customer demand better.
* To analyse the sales and market trends to determine optimal inventory levels, reducing stockouts and overstock situations while maximising sales opportunities.
* To identify key factors influencing inventory turnover and supply chain efficiency, facilitating improved collaboration with suppliers and logistics partners for timely replenishment and reduced operational costs.

# 2. Background

## 2.1 Introduction to Supply Chain Forecasting

A supply chain encompasses all entities directly or indirectly engaged in satisfying a consumer request or demand (Syntetos *et al.*, 2016). A 'party' refers to any decision-making entity inside the supply chain. It may refer to an organisation or a business unit inside an enterprise. The supply chain encompasses the eventual client, various retailers, wholesalers, and distributors and stretches back to manufacturers along with their component and raw material suppliers. The chain includes the movement of commodities, goods, information, and finances. This article emphasises the movements of resources, goods, and information, notwithstanding the undeniable significance of money flows.

Incorporating financial projections into an organisation's planning framework is beyond the purview of this analysis (Muthukalyani, 2023). The consumer's desire initiates the whole supply chain. It formulates a strategy for retail businesses to address such demand by ensuring the availability of requisite items and services to meet client needs. These eventually include the creation of requests or demand at the following level upstream in the supply chain: wholesalers or distributors, who then react by issuing requests to manufacturers, and so forth. The upstream flow of requests represents the transfer of information between supply chain members supply chain members. This information flow is augmented by a downstream flow of materials/products in the supply chain to fulfil these requirements (Syntetos *et al.*, 2016).

## 2.2 Selection Criteria for Papers

Closely defined selection criteria were used to include to include papers in this sto toto provide relevance and quality of articles. The research was only focused on peer-reviewed articles published in reputable journals as they are credible. Secondly, the papers needed to concentrate on data science applications in the domain of inventory forecasting or supply chain management in the retail sector (Fildes, Ma and Kolassa, 2022). Moreover, at the forefront were studies with or without empirical evidence or case studies that provide practical application. Vanity, a survey and book published within the last five years, was favoured to select recent publications in order to ensure that the findings are current and easily translated to present-day retail issues.

## 2.3 Critical Analysis of Key Papers

### 2.3.1 Paper 1: “AI-driven demand forecasting: Enhancing inventory management and customer satisfaction”.

Amosu et al. (2024) evaluate the importance of using AI-based demand forecasting for supplying and providing better customer satisfaction instead of its absence. According to Amosu et al. (2024), classical forecasting fails to predict consumer demand precisely, causing undermining results, i.e. excess inventories or stockouts. The study shows that resorting to advanced AI algorithms and machine learning models allows for improving forecasting accuracy directly in proportion to enhancing operational efficiency and customer satisfaction. A particularly good thing about the paper is that it is empirical: It shows the effect of adding AI to the inventory system by automating replenishment processes and aligning its stock levels with what it expects to need. Amosu et al. (2024) highlighted that the neural network models outperform in the sense of lower Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), thus arguing the need for encompassing external factors such as seasonality and promotions in forecasting models.

Nonetheless, the study might be enriched by discussing possible challenges for implementing AI, such as data quality problems and a high initial investment. From a practitioner's point of view, this study highlights the potential power of AI in retail inventory management to lessen supply chain operations. It offers valuable knowledge on effectively utilising AI's benefits in this area (Amosu *et al.*, 2024).

### 2.3.2 Paper 2: “Demand Forecasting in Supply Chain Management for Rossmann Stores using Weather Enhanced Deep Learning Model”

Qureshi et al. (2024) present an enhanced deep-learning model to solve the demand forecasting problem for Rossmann stores. The addition of external variables, such as weather conditions, makes incorporating these variables into forecasting models a critical point. It lays out an in traditional demand forecasting, which frequently dismisses environmental considerations regarding consumer behaviour. This model is empirically validated, and this decisive point in the study is provided by applying this model to a massive dataset of 1,115 Rossmann stores across Europe (Qureshi *et al.*, 2024). This results in a considerable improvement in forecasting performance from conventional methods, providing evidence of the effectiveness of brain-deep learning methods in retail inventory management. However, since the study could benefit further from elaborating on the model's limitations and possible biases that heavily relies on weather data, this is a relatively short read.

In addition, as the integration of deep learning is promising, Qureshi et al. (2024) need to discuss the computational complexity and resource needs of such models, which may be undesirable for smaller retailers. This study provides vital ideas on the ways to include the effectiveness of demand forecast methodologies in supply chain management that allow for more responsive and efficient retail operations.

### 2.3.3 Paper 3: “Retail forecasting: Research and practice”

Fildes, Ma, and Kolassa (2022) provide a comprehensive review of retail forecasting, encompassing research accomplishments and applications in real-world settings, specifically duringamid dynamic market changes due to the shaking of the norm by the COVID-19 pandemic. Kolassa (2022) succeeded in effectively synhesisg literature effectively and highlighting the evofacethat retailers face, such as increasingonline comfortition and the need for faster and more agile forecasting methods. One strength of the paper is that it integrates demand forecasting with machine learning techniques, which are taking the modern trend, and provides hints on how to strengthen accuracy.

The study does, however have some limitations to discuss. It highlights the advantages of using advanced forecasting methods but fails to adequately convey the extent of the barriers to implementation, such as problems with data quality and an extremely high level of investment required in technology and training. Additionally, Fildes, Ma, and Kolassa (2022) acknowledge improved forecasting accuracy and enhancedoperational performance, but their work does not continually enhance operational performance. If oversight is not exercised in this respect, there may be misconceptions about the direct benefits of adopting new forecasting technologies. In addition, more empirical case studies of successful applications of their recommendations in real-life settings would add value to the paper. The review is interesting, and some refinement of challenges and practical implications could add more excellent value to the contribution to academic research and practice in retail (Fildes, Ma and Kolassa, 2022).

### 2.3.4 Paper 4: “Unlocking accurate demand forecasting in retail supply chains with AI-driven predictive analytics”.

Muthukalyani (2023) explores the transformative potential of AI-driven predictive analytics in improving demand forecasting for retail supply chains. Muthukalyani (2023) effectively proves that traditional forecasting methods usually cannot reflect the intricacy of consumer behaviour and market dynamics to reduce inventory management inefficiency. The study integrates advanced AI algorithms like machine learning and deep learning, by which the retailers will analyse the big datasets to fetch the patterns and trends, thereby enhancing forecasting accuracy.

One key strength of this work is its practicality, given embedded case studies that demonstrate the practical benefits of having Al (increase in inventory turnover and customer satisfaction). Nevertheless, one point worth mentioning as a limitation is the absence of details on the difficulties involved in applying AI-based solutions in the retail space. Muthukalyani's (2023) study acknowledges data quality problems and continuous model refinement, but not the organisational barriers that may face retailers, constraints like not wanting to change or needing skilled staff. Therefore, there may be an underestimation of the complexities associated with the operationalisation of these advanced technologies in real-world settings.

## 2.4 Summary of Literature Review

The literature study on the demand forecast in the retail supply chain showed that the technology is being incorporated into the increasing demand forecast to improve its accuracy, Above all, artificial intelligence (AI) and machine learning. While those methods are also insufficient, researchers have identified more traditional methods as inappropriate in dealing with consumer behaviour and market dynamics (Falatouri *et al.*, 2022). AI-driven predictive analytics use their studies to show the accuracy of AI-driven predictive analytics, which analyse massive data sets and incorporate outside factors like weather and economic conditions to improve demand prediction. Furthermore, the literature underlines the requisite empirical validation through case studies where the application of these technologies has been realised in the practical environment. Nevertheless, those challenges exist, such as a lack of data quality issues and the requirement of experienced personnel to deploy sophisticated models (Tadayonrad and Ndiaye, 2023).

A critical gap identified is that these technologies are not road-tested in discussing how retailers deal with the associated organisational barriers. While AI and machine learning promise significant improvement for retail forecasting, these transformational changes depend on a nuanced assessment of implementation challenges for prospective practitioners seeking to improve their supply chain operations (Babai, Boylan and Rostami-Tabar, 2022). The synthesis of these findings served as a reasonable basis for further studies of effective strategies to overcome these obstacles.

# 3. Methodology

I followed a complete data science approach to predict retail store inventory needs by developing several machine learning models and carefully recording all steps. I started my work by adding Pandas' data handling to my project, then added Numpy mathematical tools alongside visualisation libraries Matplotlib and seaborn, plus Scikit-learn to implement and evaluate ML models. I first loaded "Warehouse\_and\_Retail\_Sales.csv" into a Pandas DataFrame and then checked its initial data rows using the head() function to verify the data import. I showed yearly trends in this format by summing retail sales numbers annually through a visual bar plot. I selected the 'Viridis' palette since its colour gradient works well to show differences among all years.

I processed data by filling out the Supplier and Item Type categories with Unknown and setting the missing Retail Sales values to 0. Model training results would produce errors when we did not normalise retail sales data before starting. I transformed categorical supplier and item type data into numeric representation using pandas' get\_dummies() to detect how each supplier type and item category affects the model without assuming a specific order among them. I trimmed unneeded columns from the dataset, including ITEM CODE and ITEM DESCRIPTION, because these data points could reduce the effectiveness and accuracy of our data analysis. I adjusted all numeric columns YEAR, MONTH, RETAIL SALES, RETAIL TRANSFERS, and WAREHOUSE SALES through StandardScaler to reach standard values of 0 mean and one standard deviation. This standardisation method became key for ensuring proper algorithm performances because these algorithms react strongly to input feature sizes.

After cleaning my dataset, I divided the features into X and assigned RETAIL SALES as the target variable Y. To create training and testing segments from the data, I used Scikit-learn's train\_test\_split function with an 80-20 split and a fixed random state for testing results reproduction. Following the data split, I rescaled all training and test features using the same normalisation technique before beginning training. The first linear regression model served as my baseline to start the modelling. My selection of Linear Regression stands because it helps forecast future values directly and offers a basis for comparing more advanced forecasting methods. I trained the model on training data and checked its prediction accuracy on test data using both Mean Squared Error (MSE) and R² score. I used RMSE to measure prediction errors because this statistic reflects the target variable units.

To lower the overfitting risk, I switched to Ridge Regression, which adds L2 regularisation to the linear regression model. I trained GridSearchCV to adjust the alpha value so the model reached its best performance between bias and variance. The adjusted Ridge model showed better results for test set MSE and R square values since it handled multicollinearity and limited model complexity.

I launched my analysis by testing a Decision Tree Regressor in my dataset. The structure of decision trees enables them to detect varied data patterns, but they tend to align too closely with specific observations. I determined the model's performance for Decision Trees through the same evaluation criteria. The next step was to create predictive models using various team members. The Gradient Boosting Regressor adds weak learners’ step by step to optimise the remaining errors in the dataset. This model works well to avoid overfitting problems while finding hidden connexions in data. I applied LightGBM and XGBoost regressors to analyse regression data because these methods are fast and handle regression tasks effectively. XGBoost and LightGBM achieved accurate results in regression tasks with lower overall error scores than other models.

I used CatBoost Regressor for its exceptional handling of categorial data and its automated pre-processing functions. To test model accuracy, I used R² scores to determine variance capture and MSE and RMSE to measure prediction error values during all project activities. I presented model performances on a bar chart that showed how well each model worked by displaying their percentage accuracy results excluding linear regression model.

## 3.1 Explanation of Features

The dataset contains of 9 features.

| **Feature** | **Description** |
| --- | --- |
| **YEAR** | The year in which the data was recorded. |
| **MONTH** | The month in which the data was recorded. |
| **RETAIL SALES** | The total sales made in retail stores (Target Variable). |
| **RETAIL TRANSFERS** | The number of items transferred between retail stores. |
| **WAREHOUSE SALES** | The total sales made in warehouses. |
| **SUPPLIER** | The supplier of the items. |
| **ITEM TYPE** | The type of item sold. |
| **ITEM CODE** | A unique identifier for each item. |
| **ITEM DESCRIPTION** | A textual description of the item. |

## 3.2 Data Type and Quality

| **Feature** | **Data Type (Pre-cleaning)** | **No. Missing Values** | **Cleaning Action with Justification** |
| --- | --- | --- | --- |
| **YEAR** | int64 | 0 |  |
| **MONTH** | int64 | 0 |  |
| **RETAIL SALES** | object | 167 | Converted to numerical and missing values filled with 0 to prevent errors in modeling. |
| **RETAIL TRANSFERS** | object | 0 |  |
| **WAREHOUSE SALES** | object | 0 |  |
| **SUPPLIER** | object | 1 | Missing values filled with 'Unknown', then one-hot encoded for machine learning. |
| **ITEM TYPE** | float64 | 3 | Missing values filled with 'Unknown', then one-hot encoded to convert categorical data into numerical format. |
| **ITEM CODE** | float64 | 0 |  |
| **ITEM DESCRIPTION** | float64 | 0 |  |

## 3.3 Summary of Project Methodology

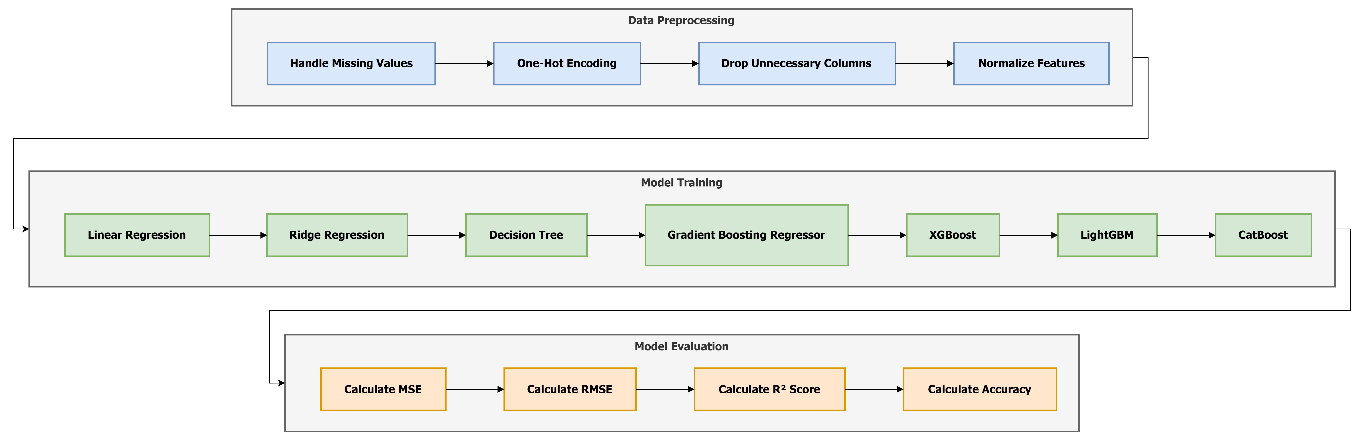


Figure 1: Project Methodology

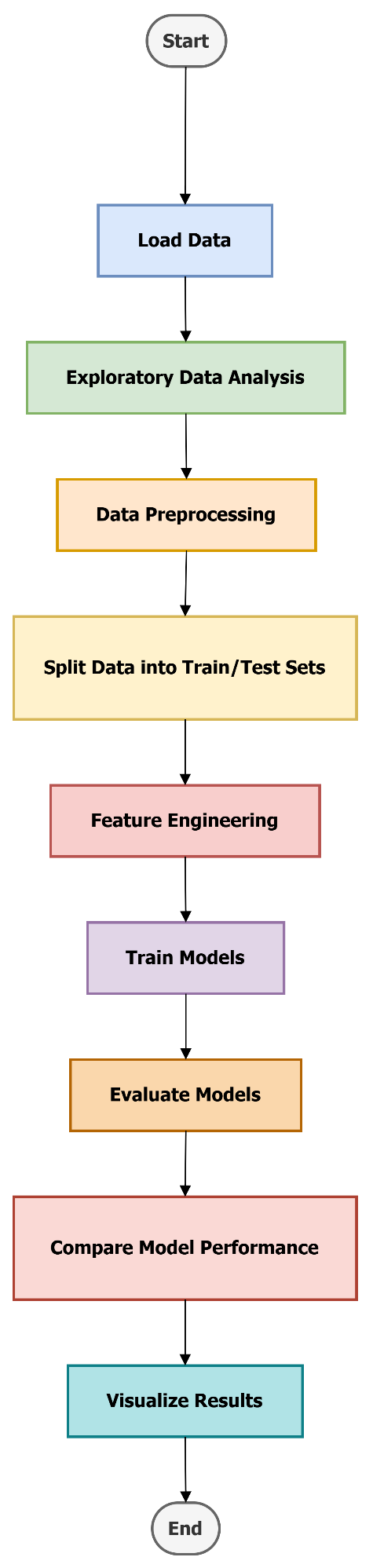


Figure 2: Project Methodology Flow Chart

# References

Amosu, O.R., Kumar, P., Ogunsuji, Y.M., Oni, S. and Faworaja, O. (2024) ‘AI-driven demand forecasting: Enhancing inventory management and customer satisfaction’, *World Journal of Advanced Research and Reviews*, 23(2), pp. 100–110.

Babai, M.Z., Boylan, J.E. and Rostami-Tabar, B. (2022) ‘Demand forecasting in supply chains: a review of aggregation and hierarchical approaches’, *International Journal of Production Research*, 60(1), pp. 324–348. Available at: <https://doi.org/10.1080/00207543.2021.2005268>.

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, pp. 993–1003.

Fildes, R., Ma, S. and Kolassa, S. (2022) ‘Retail forecasting: Research and practice’, *International Journal of Forecasting*, 38(4), pp. 1283–1318.

Ingle, C., Bakliwal, D., Jain, J., Singh, P., Kale, P. and Chhajed, V. (2021) 'Demand forecasting: A literature review on various methodologies', in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, pp. 1–7. Available at: <https://ieeexplore.ieee.org/abstract/document/9580139/> (Accessed: 23 February 2025).

Mori, C. (2023) *Mastering Supply Chain Optimization with Inventory Forecasting*. Available at: <http://projectverte-8211470.hs-sites.com/hs-web-interactive-8211470-171805697384> (Accessed: 23 February 2025).

Muthukalyani, A.R. (2023) ‘Unlocking accurate demand forecasting in retail supply chains with AI-driven predictive analytics’, *Information Technology and Management*, 14(2), pp. 48–57.

Qureshi, N.U.H., Javed, S., Javed, K., Naqvi, S.M.R., Raza, A. and Saeed, Z. (2024) ‘Demand Forecasting in Supply Chain Management for Rossmann Stores using Weather Enhanced Deep Learning Model’, *IEEE Access* [Preprint]. Available at: <https://ieeexplore.ieee.org/abstract/document/10703040/> (Accessed: 23 February 2025).

Syntetos, A.A., Babai, Z., Boylan, J.E., Kolassa, S. and Nikolopoulos, K. (2016) ‘Supply chain forecasting: Theory, practice, their gap and the future’, *European journal of operational research*, 252(1), pp. 1–26.

Tadayonrad, Y. and Ndiaye, A.B. (2023) ‘A new key performance indicator model for demand forecasting in inventory management considering supply chain reliability and seasonality’, *Supply Chain Analytics*, 3, p. 100026.

Tian, X., Wang, H. and Erjiang, E. (2021) ‘Forecasting intermittent demand for inventory management by retailers: A new approach’, *Journal of Retailing and Consumer Services*, 62, p. 102662.