Computer and Decision Making

An International Journal

www.comdem.org

eISSN: 3008-1416

FORECASTING EMERGING PRODUCT TRENDS IN SMART SUPPLY CHAINS

Jianing Mao[1], Wenqing Hu[2,*], Xin Wen[1]

[1] Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

[2] Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Article history:

Received 11 September 2024

Received in revised form 29 September 2024

Accepted 10 October 2024

Available online 10 October 2024

Keywords:

Product trends; supply chain; time series analysis; deep learning

ABSTRACT

This paper aims to develop innovative solutions using multi-modal data to predict emerging product trends in the market. By combining data flow, historical product information, and the latest market open data, our solution empowers companies to proactively adjust their supply chain and grasp emerging market trends, leading to supply chain optimization and enhanced consumer satisfaction. The proposed solution involves the establishment of machine learning models for big data analytics, focusing on product selling, consumer behaviors, and emerging markets. The first part of the solution is applying various quantitative methods to forecast the sales of the products available in the dataset. Then, we can process and analyze unstructured data like product reviews and social media posts by applying natural language processing (NLP) and time series analysis, providing insights into psychological and cultural factors. We use data dashboards to present Integrated Analysis, Sales Forecast, and Customer Review Analysis. The generated Data dashboards can be used for effective supply chain and e-commerce data management. Forecasting product trends allows companies to optimize supply chain management and adjust strategies based on accurate predictions. The approach offers valuable customers and market insights for informed decision-making.

* *Corresponding author.*

E-mail address: wenqing.hu@connect.polyu.hk

Doi: https://doi.org/10.59543/comdem.v1i.10699

1. Introduction

In the past, failure to incorporate multi-modal data into trend prediction can result in critical issues such as trend analysis inaccuracy, decreased competitiveness, and financial impact [15]. A reliance solely on traditional data sources or a single type of data can yield an incomplete or distorted view of market trends, leading to strategic missteps. Offerings might not align with current consumer preferences, resulting in diminished customer satisfaction and loyalty [6]. Without insights from multi-modal data, companies may only react once trends become apparent, falling behind proactive competitors. Inaccurate trend analysis leads to reduced competitiveness, which can have negative financial impacts. Inefficiencies in managing supply chains and inventories can inflate costs due to emergency orders, surplus stock, or wasted products. A disconnect with prevailing market trends can lead to a forfeiture of sales opportunities and a contraction of revenue.

Predicting emerging product trends is very important for businesses to utilize the resources effectively, visualize business performance, arrange the emerging products and service launch time reasonably, and review management decisions. By developing an advanced prediction analysis model, the solution can accurately identify the pattern and potential product trends in the market. Based on the history and latest market data, machine learning algorithms are used to predict the popularity of specific goods to help business insight into trends in the market competition. By analyzing the online and offline consumer behavior data in depth, the solution can help businesses gain insight into consumer preferences and buying patterns. The solution will assess the sentiment of consumers for specific products by sentiment analysis technology to help businesses adjust to market needs. Then, the solution will provide practical supply chain adjustment advice based on the predicted trends. Through the optimization of the whole process of warehouse, production, and distribution, the solution can empower the business's ability to meet the needs of emerging products, reduce stock risks, and improve the efficiency and flexibility of the supply chain.

The rest of the paper is structured as follows: Section 2 is a review of related works. In Section 3 we compared the several models for predicting time-series sales data and the model used to understand the consumer behavior data. In Section 4 we mainly discuss how to adjust the supply chain based on our models, the limitations of our current model, and the future directions. Section 5 concludes the whole paper.

2. Literature Review

In the past years, there has been a lot of research work on sales prediction and forecasting supply chain demand. They focus on different aspects and use different methods. In this section, we mainly consider the following aspects: 1. Sales prediction using ordinary machine learning methods 2. Sales prediction with deep learning, for example, Long Short-Term Memory (LSTM). 3. Sales prediction with other forms of data source, not limited to pure sales data, such as review data and other user-generated content.

2.1 Traditional Machine Learning Method

Sales prediction [17, 22] is closely linked to forecasting emerging product trends in the supply chain. Several previous studies have explored sales prediction using different models and methods. Kohli *et al.* [19] applied linear regression and K-Nearest Neighborhood (KNN) regression for sales prediction. These models are simple, easy to implement, and yield fair results on their dataset. Makridakis *et al.* [25], in their book, proposed a range of methods, including simple regression, multivariate regression, the AutoRegressive Integrated Moving Average (ARIMA) algorithm for time series analysis, and combining machine learning algorithms with econometric models to achieve better results. However, linear models can only capture the linear relationships between factors, which may not be sufficient for modern, complex sales models [35].

2.2 Sales Prediction with Deep Learning

The advancement of neural networks and deep learning has led to the use of deep learning methods for sales prediction. Kechyn *et al.* [42] used WaveNet for time series prediction in a Kaggle competition and achieved relatively good results compared to other models. Bandara *et al.* [1] highlighted that, in modern e-commerce, employing cross-series data from related products could be beneficial. They proposed using LSTM networks to capture the non-linear relationships between related products within the same category. While neural networks often outperform other models, their complex structures can make them difficult to interpret, limiting the insights they provide in some cases [35].

2.3 Other Data Sources for Sales Prediction

In recent years, sales prediction has expanded beyond pure sales data to include other sources, such as user-generated content and social media influence. Several studies have explored this area. Ye *et al.* [37] highlighted the impact of positive online reviews on the hospitality industry. Yu *et al.* [38] further examined the content of reviews and found a relationship between review sentiment and sales at the cinema box office. Huang *et al.* [41] used sentiment analysis to explore the connection between the topic distribution of product reviews and sales prediction. In addition to post-purchase reviews, attention has also shifted to the impact of social media on pre-purchase decisions. Zhang *et al.* [39] studied the correlation between trending topics on Twitter and the sales of related products on eBay, suggesting that such correlations could be used for predictive tasks. For instance, if a topic is trending on Twitter, sellers may take note and adjust their supply chains accordingly [39].

In conclusion, the current research on sales prediction developed from traditional machine learning method to deep learning, and the source data also have developed to multi-modal data: combing the sales data and other user-generated data. However, most of these research works focus on a single aspect, such as pure sales data or pure review data, this may be impractical for modern complex online sales situation. In this paper, we apply the multi-modal concept in data collection and propose the model to combine the possible factors that may affect the sales prediction to provide a better result.

3. Methodology and Results

Our basic idea is to establish machine learning models for big data analytics. The entire model includes three parts: product selling, consumer behaviors, and emerging markets. This structure satisfies the problem description of general market prediction and the specified conditions for emerging products.

3.1 Dataset Collection

The effectiveness of the proposed predictive models hinges on the careful selection and integration of diverse datasets [11], which collectively offer a comprehensive view of market dynamics, consumer behavior, and emerging product trends. The datasets used in this project are sourced from both historical and real-time data streams to ensure robust and accurate trend forecasting.

The historical datasets primarily include transactional data from e-commerce platforms and consumer behavior logs. One key source is the Online Retail II dataset, which encompasses all transactions from a UK-based online retail company over a two-year period. This dataset includes features such as invoice numbers, product codes, quantities, invoice dates, and customer identifiers, offering detailed insights into historical sales patterns. Another significant historical dataset comes from a comprehensive database of online shopping intentions, containing over 12,000 user sessions. This dataset provides attributes related to product specifics, ratings, and consumer interaction data, which are crucial for understanding past consumer behavior and product performance.

In addition to historical data, real-time datasets are critical for capturing current market trends and consumer sentiment [33]. These datasets are aggregated from multiple sources, including search engine trends and social media platforms such as Instagram and Xiaohongshu. By leveraging hashtag tracking [24, 36] and user interaction data from these platforms, the system can monitor emerging trends and consumer preferences in near real-time. The integration of real-time data with historical datasets allows for a dynamic understanding of market conditions, enabling the models to predict shifts in consumer behavior and product trends more accurately [31]. Additionally, data from Amazon Reviews, includes user reviews, item metadata, and interaction data. Review data provide valuable insights into consumer satisfaction and product reception, further enriching the predictive capabilities of the system [7].

This multi-modal approach to dataset collection, combining historical and real-time data, is essential for developing a predictive model that is both accurate and adaptable to the rapidly changing landscape of consumer behavior and market trends.

3.2 Forecasting based on historical selling data

The first part of the solution is applying various quantitative methods, such as building traditional time series models and linear regression, to forecast the sales of the products.

As a key innovation in our solution, we introduce data abstraction before modeling, a method that deviates from the conventional use of historical data as a feature. Instead, we try different combinations to test potential correlation relationships, as shown in Figure 1, for example, incorporating the historical data's rolling mean, maximum, and minimum in a specific window as one additional dimension. We also use the lag for different durations and integrate derivatives for certain features as one dimension. This approach allows us to create more characteristics for the data description, revealing the significant role of some overlooked dimensions in the final model.

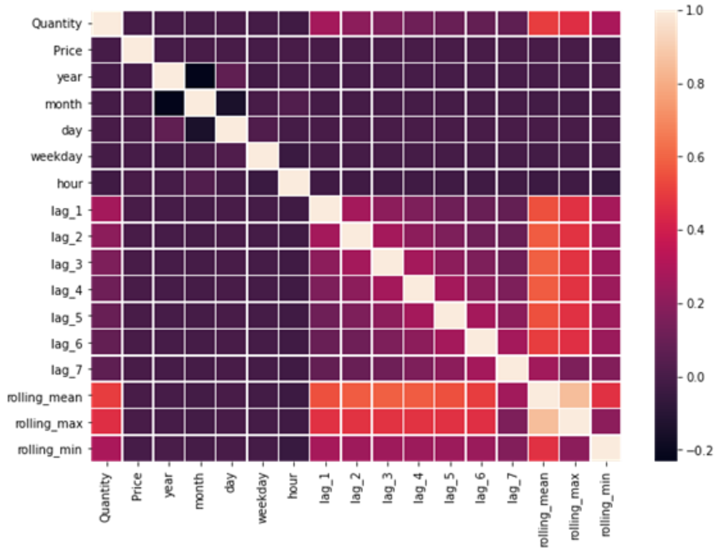

Fig. 1. Correlation Map

Our analysis begins with linear regression. Since product selling often has seasonal patterns, we delve into time series modeling, starting with the practical seasonal naive model. This method considers seasonal patterns by referencing the same period from the previous year. For instance, if we predict sales for December 2017, the seasonal naive method will assume the same number of sales as in December 2016. We set it as a baseline to compare with more sophisticated models. We then explore the ARIMA model [25] to adapt to the data variance. ARIMA can capture different standard temporal structures in time series data. It can be beneficial for understanding sales trends and product seasonality. Finally, we test the Holt Winter triple exponential smoothing model, which combines value, trend, and seasonality influences to predict current or future values. It assigns exponentially decreasing weights over time, making it practical for predicting future values based on past and present ones. Prediction values and errors from different models are shown in Figure 2 below.

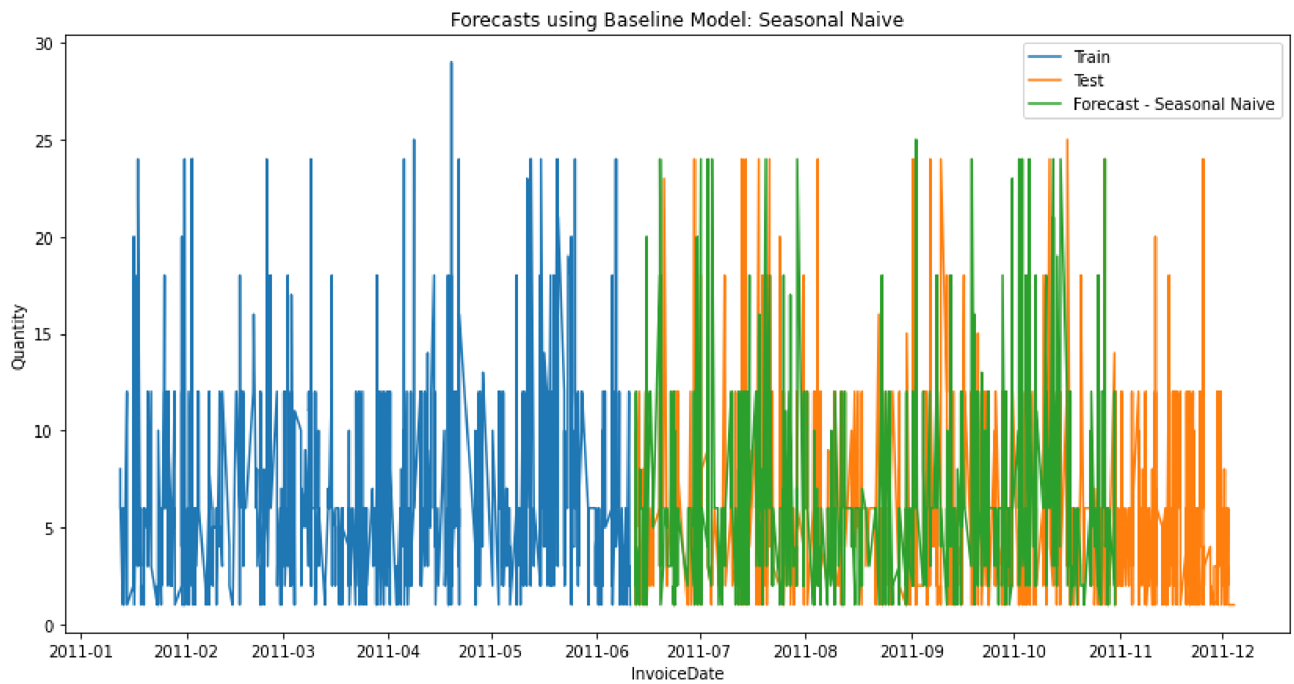

Fig. 2. Prediction, Actual Value, and Error Visualization

Following the training phase, we evaluate the forecasts by comparing the actual and forecasted sales values. We employ several standard evaluation metrics, including the mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE). These metrics provide a comprehensive view of the model's performance. In our comparison, the Holt Winter triple exponential model is the most accurate, instilling confidence in its reliability, and the relative error can be controlled within five percent. The comparison between different models is presented in Table 1 below.

Table 1
Model Evaluation Result

Model	**MAE**	**RMSE**	**MAPE**
Linear Regression	6.743671	75.854784	3288.938781
Seasonal naive	5.056452	6.934137	60.497877
SARIMA	4.773979	6.085521	11.792152
Holt-Winters	3.440474	4.652401	9.826132
Without combining feature	8.741493	86.179229	2649.601202

By combining these algorithms, we can create a robust time series analysis model capable of handling complex trends, seasonal influences, and other factors that affect product sales. This will enable us to predict future product trends more accurately, providing more precise information for supply chain management.

3.3 Understanding Consumer Behavior Factors

Consumer behavior [8] is a complex and multifaceted process influenced by various factors that can significantly impact purchasing decisions. Understanding these factors is crucial to the success of our predictive model, as it can enhance the accuracy of our predictions and provide more nuanced insights into market trends [40].

Internet Behavior Factors: This involves the analysis of online shopping behaviors, such as browsing patterns, time spent on different pages, click-through rates, and abandonment rates, which are most accessible to quantify. For example, understanding their online behavior could predict the probability of a consumer completing a transaction. Web analytics tools can be used to collect this type of data and deep learning methods can be used to analyze the patterns and predict the likelihood of a purchase. For instance, if consumers frequently visit a specific product page and spend significant time there, they might be more likely to purchase that product. Similarly, if consumers abandon their shopping cart often, they might be less likely to complete a transaction. By incorporating these behavioral data into our predictive model, we can better understand the online shopping behaviors that lead to purchases and tailor our marketing strategies accordingly.

Table 2
Features we collected and used to predict the likelihood of leading to a transaction

Feature	Description
Administrative	The count of account management-related pages visited by the user.
Administrative duration	The total duration (in seconds) the user spent on account management-related pages.
Informational	The count of pages the user visits about the shopping site's website, communication, and address information.
Informational duration	The total duration (in seconds) the user spent on informational pages of the shopping site.
Product-related	The count of product-related pages visited by the user.
Product-related duration	The total duration (in seconds) the user spent on product-related pages.
Bounce rate	The average bounce rate for the pages the user visits.
Exit rate	The average exit rate for the pages visited by the user.
Page value	The average value of the pages visited by the user.
Special day	The user's site visit time is near a particular day.
Region	The geographic region where the user initiated the session.
Traffic Type	The source of traffic that directed the user to the website (e.g., banner, SMS, direct).
Visitor Type	The categorization of the user as 'New Visitor,' 'Returning Visitor,' or 'Other.'
Weekend	A binary value indicating whether the user's visit occurred on a weekend.
Month	The month in which the user's visit occurred.

Based on these features, we build the model, which consists of 5 hidden layers (200, 200, 200, 200, 100) and uses ReLU as an activation function. We optimize the model with Adam for 500 epochs at most. The evaluation for our model is shown in Table 3 below.

Table 3
Evaluations of our deep multi-layer perception model.

Accuracy	True-positive rate	True-negative rate	F1 Score
90.3%	0.58	0.95	0.58

3.4 Reviews, Ratings, and Social Media Comments

Social media, like Instagram, TikTok, and Xiaohongshu, sometimes play a vital role in determining trending products [3, 29]. For example, an influencer or KOL's recommendation may become the next trend. There are two different marketing strategies on social media: direct marketing and indirect marketing. Direct marketing means directly advertising to the customers [2, 4, 20], and indirect marketing is usually done via an intermediary [21].

In our analysis, we will mainly focus on indirect marketing strategies. Indirect marketing with influencers or KOL usually gains more visibility on social media, as well as more likes and comments. Likes or comments significantly increase the speed at which products are promoted, making them more likely to become trending.

Our evaluation is based on the statistics of KOL or the influencer who promotes this product, including the number of their followers, the statistics of the post, including the number of likes and comments, and also the reputation of the influencer in specific domains, such as influencer specialized in technological products, fitness or cooking. Their reputation and impact in specialized areas also greatly influence the trending products in certain areas [14, 30].

KOL's advertising is one factor; comments people post on social media can also be an essential factor affecting sales. Comments on social media can affect the sales in the following ways:

Reviews and ratings can help sellers to get feedback from buyers. Sellers can assess customer satisfaction and receive direct feedback on their products. This allows them to understand what customers like and dislike, helping to improve product quality and service and helping with quality control and product improvement. Advanced sellers might use language models for sentiment analysis to quantify and categorize comments at scale, providing a more data-driven approach to understanding customer feedback.

Product reviews are in the form of unstructured data [26]. NLP is pivotal in analyzing unstructured data [18]. It can help extract sentiments, identify key topics, and even detect subtle cultural nuances from text data, providing rich insights into psychological and cultural factors. We provide a keyword cloud, including the keywords extracted from numerous reviews posted by the buyers. These views can reflect the actual feelings of buyers on the product. Sellers can use the data for simple inspection or use LLMs GPT model from OpenAI or other sources for sentiment analysis.

In summary, we can build a more robust and accurate predictive model by integrating these diverse data sources and applying appropriate data mining and NLP techniques. This comprehensive approach will allow us to anticipate emerging product trends more effectively and assist businesses in making informed decisions about supply chain adjustments and market strategies [9].

4 Discussion

4.1 Generalized Analysis

The proposed solution is aimed to deal with the challenge of predicting the emerging product trends in smart supply chains based on using programming-based methods. Based on multi-modal data, applying predictive analysis algorithms to analyze consumer behaviors and provide practical suggestions can help businesses equipped with tools to adjust supply chains and optimize operations. The proposed solution makes the businesses grasp the emerging product trends, meet consumer needs, and gain strategic advantages in the fierce competition in the market.

This solution analyzes sales data to predict product demand, allowing for better optimization of production and distribution across stores. By doing so, it reduces surplus inventory and guarantees that stores are stocked with the appropriate products when needed. Insights derived from data empower the company to adapt its inventory and operations in response to current trends and customer preferences. Analyzing products in social media can help to predict the potential demand for certain products. If one product is promoted by many influencers on social media, some people may also want to try this. Even if the influencers' audience only buys this product once, this could probably cause a sudden increase in demand. Through the combination of social media trends and other analysis results, the firm can adjust the supply chain coordinately to achieve higher efficiency. By comparing the predicted data and real data from the supply chain, the model can be further improved to make the prediction more accurate.

This approach is practical as it allows for the extraction of important customer and market insights via data analysis, which can then be converted into informed business decisions. By adapting to emerging trends quickly and refining operations according to demand predictions, a competitive edge can be established that fosters success and growth.

4.2 Adjustments to the supply chain based on forecast results

The forecast of product trends is quite important for the enterprise. By forecasting sales change and seasonal demand fluctuation, the enterprise can enhance supply chain management to boost efficiency and lower costs based on meeting customer needs [13, 28]. When the enterprise can accurately predict product trends consisting of the change of market trends and customer preferences, it can adjust supply chain strategies to adopt these trends and changes to occupy a favorable position in the market competition.

First, the forecast of product trends can give insight into the future market trends and demand situation so that enterprises can conduct demand-orientation resource planning. By analyzing market trends, the enterprise can better understand the possible future sales situation. Optimizing production or restocking plans and logistics arrangements can ensure the timely availability of required products and materials [12]. Predictive analysis can help improve inventory management, ensuring that companies maintain appropriate supply levels, and reducing the costs of excessive inventory storage or potential customer loss due to stockouts [23, 34]. By adjusting inventory strategies based on historical product information and customer behavior patterns, enterprises can be prepared for anticipated demand peaks or downturns and seize opportunities based on customer needs, preferences, and purchase intentions. Inventory optimization through data analysis can help calculate delivery cycles. By combining delivery cycles with current sales data, enterprises can estimate safety stock and calculate the reorder point to notify retailers when to place reordering requests. As a result, inventory turnover is improved, helping to accelerate cash flow, improve capital utilization, and enhance the operational efficiency of the supply chain.

Product trend prediction can facilitate coordination and collaboration among various stages of the supply chain [16, 27, 10], guiding companies to establish strategic partnerships with suppliers. If an enterprise predicts future demand growth for a particular product, it can establish cooperative relationships with potential suppliers to ensure that the supply chain can meet future demand. Through close collaboration with suppliers, manufacturers, and logistics partners, supply chain elements can respond promptly to changes in product trends, minimizing delays and mismatches. This helps enhance the overall reliability, flexibility, and efficiency of the supply chain while reducing supply risks.

Before predictive analysis, enterprises could only rely on experts and past data to make business decisions. In contrast, predictive analysis uses various data and real-time trends to prepare models for multiple scenarios and identify potential risks that may disrupt the supply chain. This significantly improves risk management capabilities. By accurately predicting sales trends and adjusting supply chain strategies accordingly, companies can improve supply chain efficiency, flexibility, and customer satisfaction.

The company can adopt appropriate inventory and cooperation strategies based on the results of the sales forecast [32]. When sales forecasts are low, the company doesn't need to hold too much inventory, avoiding excess stock. However, when sales forecasts are high, the company needs to increase inventory in advance to prepare for the peak. The company can work closely with suppliers based on sales forecasts to ensure that every aspect of the supply chain responds promptly to changes in product trends, enhancing reliability and flexibility. Based on the forecast data, the company can develop a comprehensive supply chain strategy. When adjusting the supply chain based on forecast data, the company should pay attention to fluctuations in monthly sales quantities. For example, if there is a significant increase in a certain month. By analyzing whether this pattern is consistent year after year, conclusions can be drawn about whether there is seasonal growth in demand during this period. Considering the fluctuations in monthly sales quantities, managing inventory levels to meet demand without excessive stock is crucial. Using historical data to forecast future months' inventory requirements and avoiding excess inventory reflects the principles of Just-in-Time (JIT). The company can negotiate better terms with suppliers using the forecasted demand data, such as bulk discounts for months with expected high sales. If lower sales are forecasted for certain months, the company can consider increasing marketing or promotional activities to stimulate demand. Supply chain flexibility is also essential. The enterprises should be prepared to quickly adjust the supply to respond to unexpected increases or decreases in demand. This may involve diversifying the supplier base and improving logistics flexibility.

Sales forecasts can be specific to a particular day of the month. Based on the fluctuations in sales, we can adjust the safety stock levels and reorder points. By considering the changes in sales, we can collaborate with suppliers to adjust the supply quantities based on demand. We can optimize the logistics network based on the variations in sales. For example, if sales are higher on a particular day of the month, we can adjust the logistics resources to ensure timely delivery. Sales forecasts that are specific to a particular day can help enterprises adapt to sales variations, optimize the efficiency and flexibility of the supply chain, improve customer satisfaction, reduce costs, and maintain a competitive advantage.

From an environmental perspective, accurate forecasting of product trends can optimize supply chain and inventory management, reducing excess inventory and waste generation. This will help reduce energy consumption, carbon emissions, and resource waste, thereby mitigating negative impacts on the environment. From a social perspective, considering consumer behavior and market trends in the forecasting model can help enterprises better meet market demands and consumer preferences, providing the required products and services. This contributes to improving customer satisfaction, brand reputation, and market share, thus creating social value for the enterprises. From a governance perspective, this model utilizes advanced technological tools to enhance supply chain management and forecasting capabilities, promoting technological innovation and continuous improvement for the enterprise's ability to address market competition and changes. At the same time, we have implemented robust measures to strengthen data privacy and security, ensuring compliance with the forecasting model and dashboard. We have established monitoring and evaluation mechanisms to regularly review the accuracy of forecasting results and continuously update the database with new data, thereby improving the model over time.

4.3 Limitations and Future Work

The proposed system has several limitations that could impact its effectiveness and broader applicability. One major limitation is the dependency on the quality and availability of data; the accuracy of the predictive models is largely contingent upon how complete and consistent the data collected from various sources, such as historical sales, consumer behavior, and social media. Inadequate or biased data may result in inaccurate predictions, and restricted access to real-time data due to privacy concerns or proprietary restrictions could further diminish the system's reliability. Additionally, the complexity of advanced machine learning models, particularly those involving deep learning techniques like LSTM and NLP, presents challenges in terms of interpretability. Decision-makers may struggle to understand the reasoning behind certain predictions, which could hinder the adoption of these models in business contexts where clear and actionable insights are crucial. Furthermore, the system faces scalability challenges, especially for small to medium-sized enterprises that may lack the necessary computational infrastructure. The models are also vulnerable to sudden market changes or external factors, such as economic crises, which may not be fully captured by historical data, potentially affecting the accuracy of predictions. Lastly, the solution's initial design for specific regions or market segments may limit its generalizability, requiring substantial adjustments for use in different geographical locations or industries.

To address these limitations, future work should focus on several key areas. Enhancing data integration and real-time processing capabilities is essential to ensure the system can handle high-quality, timely data inputs from multiple sources. This could involve developing more sophisticated data pipelines and securing partnerships with data providers. Improving model interpretability is another crucial step, as more user-friendly tools and methods for explaining predictions will help bridge the gap between model complexity and business usability. Research should also focus on optimizing the system's architecture for scalability, potentially exploring cloud-based solutions or creating lightweight model versions suitable for SMEs. Additionally, incorporating adaptive learning mechanisms will be vital for maintaining model accuracy in the face of dynamic market conditions, allowing the system to update continuously with new data. Expanding the solution's applicability to diverse markets will require developing flexible frameworks that can be customized for different regions and industries. Finally, addressing ethical considerations and ensuring compliance with global data protection regulations will be critical as the system evolves, helping to build trust and promote responsible use of predictive analytics in business settings.

In the future, we should also consider other more complex factors that could influence the prediction. Psychological factors, such as perception, motivation, learning, beliefs, and attitudes, influence purchasing decisions, with product reviews and ratings data providing insights that can be analyzed through text mining and NLP techniques to capture sentiments. Personal factors, including age, occupation, lifestyle, personality, and economic status, can be inferred from user profiles and transaction histories, enabling trend predictions using traditional analysis methods like classification and regression. Social factors, like reference groups, family, and social roles, affect decisions and can be analyzed through social media data using network analysis to predict trend diffusion. Cultural factors, encompassing culture, subculture, and social class, are often reflected in location data, language, and purchasing behaviors, especially around cultural events.

5. Conclusion

In this study, we have developed models to forecast sales for facilitating supply chain adjustments. Our focus has centered on three key areas: analyzing time series data, studying consumer behavior, and evaluating consumer feedback and social media responses regarding products. We have also proposed practical strategies for aligning the supply chain based on the predictive outcomes derived from our models. When use our model and follow the methods we suggested for adjusting the supply chain, potential users, such as retail companies, can better manage their supply chain, in turn increasing their service quality and lowering the total cost.

In future research, there is potential to expand on our current work. Firstly, researchers could examine the quality of our data input, a vital factor for enhancing model performance. Secondly, enhancing model interpretability could unlock future business opportunities. Lastly, scalability is crucial, warranting consideration for future development prospects.

Author Contributions

Conceptualization, methodology, Mao Jianing.; methodology, software, Hu Wenqing.; formal analysis, Wen Xin. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

This research was not funded by any grant. The authors sincerely thank the editor Prof. Muhammet Deveci for his valuable advice on improving the paper.

References

[1] Bandara, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q., & Seaman, B. (2019). Sales Demand Forecast in E-commerce Using a Long Short-Term Memory Neural Network Methodology. In T. Gedeon, K. W. Wong, & M. Lee (Ed.), *Neural Information Processing* (pp. 462–474). Cham: Springer International Publishing.

[2] Bauer, C. L., & Miglautsch, J. (1992). A conceptual definition of direct marketing. *Journal of Direct Marketing, 6(2)*, 7-17. https://doi.org/10.1002/dir.4000060204

[3] Baum, D., Spann, M., Füller, J., & Thürridl, C. (2019). The impact of social media campaigns on the success of new product introductions. *Journal of Retailing and Consumer Services, 50*, 289-297. https://doi.org/10.1016/j.jretconser.2018.07.003

[4] Bearden, W. O., Ingram, T. N., & LaForge, R. W. (2007). *Marketing: principles and perspectives* (5th ed. ed.). New York: McGraw-Hill/Irwin.

[5] Boone, T., Ganeshan, R., Jain, A., & Sanders, N. R. (2019). Forecasting sales in the supply chain: Consumer analytics in the big data era. *International Journal of Forecasting, 35(1)*, 170–180. https://doi.org/10.1016/j.ijforecast.2018.09.003

[6] Cai, W., Song, Y., & Wei, Z. (2021). Multimodal Data Guided Spatial Feature Fusion and Grouping Strategy for E-Commerce Commodity Demand Forecasting. *Mobile Information Systems, 2021(1)*, 5568208. https://doi.org/10.1155/2021/5568208

[7] Chen, Y., & Xie, J. (2008). Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix. *Management Science, 54(3)*, 477–491. https://doi.org/10.1287/mnsc.1070.0810

[8] Cummins, S., W. Peltier, J., A. Schibrowsky, J., & Nill, A. (2014). Consumer behavior in the online context. (D. Angela Hausman, Ed.) *Journal of Research in Interactive Marketing, 8(3)*, 169–202. https://doi.org/10.1108/JRIM-04-2013-0019

[9] Fatouros, G., Soldatos, J., Kouroumali, K., Makridis, G., & Kyriazis, D. (2023). Transforming sentiment analysis in the financial domain with ChatGPT. *Machine Learning with Applications, 14*, 100508. https://doi.org/10.1016/j.mlwa.2023.100508

[10] Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications, 25(2)*, 119–142. https://doi.org/10.1080/13675567.2020.1803246

[11] Fenza, G., Gallo, M., Loia, V., Orciuoli, F., & Herrera-Viedma, E. (2021). Data set quality in Machine Learning: Consistency measure based on Group Decision Making. *Applied Soft Computing, 106*, 107366. https://doi.org/10.1016/j.asoc.2021.107366

[12] Ferreira, K. J., Lee, B. H., & Simchi-Levi, D. (2015). Analytics for an Online Retailer: Demand Forecasting and Price Optimization. *Manufacturing & Service Operations Management, 18(1)*, 69–88. https://doi.org/10.1287/msom.2015.0561

[13] Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting, 25(1)*, 3–23. https://doi.org/10.1016/j.ijforecast.2008.11.010

[14] Flynn, L. R., Goldsmith, R. E., & Eastman, J. K. (1996). Opinion leaders and opinion seekers: Two new measurement scales. *Journal of the Academy of Marketing Science, 24(2)*, 137–147. https://doi.org/10.1177/0092070396242004

[15] Gaw, N., Yousefi, S., & Gahrooei, M. R. (2022). Multimodal data fusion for systems improvement: A review. *IISE Transactions, 54(11)*, 1098–1116. https://doi.org/10.1080/24725854.2021.1987593

[16] Helms, M. M., Ettkin, L. P., & Chapman, S. (2000). Supply chain forecasting –Collaborative forecasting supports supply chain management. *Business Process Management Journal, 6(5)*, 392–407. https://doi.org/10.1108/14637150010352408

[17] Hsieh, P.-H. (2019). A Study of Models for Forecasting E-Commerce Sales During a Price War in the Medical Product

Industry. In F. F.-H. Nah, & K. Siau (Ed.), *HCI in Business, Government and Organizations. eCommerce and Consumer Behavior* (pp. 3–21). Cham: Springer International Publishing.

[18] Ireland, R., & Liu, A. (2018). Application of data analytics for product design: Sentiment analysis of online product reviews. *CIRP Journal of Manufacturing Science and Technology, 23*, 128-144. https://doi.org/10.1016/j.cirpj.2018.06.003

[19] Kohli, S., Godwin, G. T., & Urolagin, S. (2021). Sales Prediction Using Linear and KNN Regression. In S. Patnaik, X.-S. Yang, & I. K. Sethi (Ed.), *Advances in Machine Learning and Computational Intelligence* (pp. 321–329). Singapore: Springer Singapore.

[20] Kotler, P., Brady, M., Goodman, M., Hansen, T., & Keller, K. (2010). *Marketing Management.* Welwyn Garden City, UNITED KINGDOM: Pearson Education UK. Retrieved from http://ebookcentral.proquest.com/lib/polyu-ebooks/detail.action?docID=5136334

[21] Chang, K. F. & Shu, W. C. (2009). An efficiency comparison of direct and indirect channels in Taiwan insurance marketing. *Direct Marketing: An International Journal, 3(4)*, 343–359. https://doi.org/10.1108/17505930911000900

[22] Li, M., Ji, S., & Liu, G. (2018). Forecasting of Chinese E-Commerce Sales: An Empirical Comparison of ARIMA, Nonlinear Autoregressive Neural Network, and a Combined ARIMA-NARNN Model. *Mathematical Problems in Engineering, 2018(1)*, 6924960. https://doi.org/10.1155/2018/6924960

[23] Liang, Y., Niu, D., Cao, Y., & Hong, W. C. (2016). Analysis and Modeling for China's Electricity Demand Forecasting Using a Hybrid Method Based on Multiple Regression and Extreme Learning Machine: A View from Carbon Emission. *Energies, 9(11)*. https://doi.org/10.3390/en9110941

[24] Lorenz-Spreen, P., Wolf, F., Braun, J., Ghoshal, G., Djurdjevac Conrad, N., & Hövel, P. (2018). Tracking online topics over time: understanding dynamic hashtag communities. *Computational Social Networks, 5(1)*, 9. https://doi.org/10.1186/s40649-018-0058-6

[25] Makridakis, S., Wheelwright, S., & Hyndman, R. (1984). Forecasting: Methods and Applications. https://doi.org/10.2307/2581936

[26] McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review, 90(10)*, 60—6, 68, 128. Retrieved from http://europepmc.org/abstract/MED/23074865

[27] McCarthy, T. M., & Golicic, S. L. (2002). Implementing collaborative forecasting to improve supply chain performance. *International Journal of Physical Distribution & Logistics Management, 32(6)*, 431–454. https://doi.org/10.1108/09600030210437960

[28] Moon, M. A., Mentzer, J. T., & Smith, C. D. (2003). Conducting a sales forecasting audit. *International Journal of Forecasting, 19(1)*, 5–25. https://doi.org/10.1016/S0169-2070(02)00032-8

[29] Piller, F., Vossen, A., & Ihl, C. (2011). From Social Media to Social Product Development: The Impact of Social Media on Co-Creation of Innovation. *Die Unternehmung, 66*. https://doi.org/10.5771/0042-059X-2012-1-7

[30] Shoham, A., & Ruvio, A. (2008). Opinion leaders and followers: A replication and extension. *Psychology & Marketing, 25(3)*, 280-297. https://doi.org/10.1002/mar.20209

[31] Shubho, O. Q., Tumpa, Z. N., Dipto, W. I., & Alam, M. R. (2022). Real-Time Data Visualization Using Business Intelligence Techniques in Small and Medium Enterprises for Making a Faster Decision on Sales Data. In P. M. Jeyanthi, T. Choudhury, D. Hack-Polay, T. P. Singh, & S. Abujar (Eds.), *Decision Intelligence Analytics and the Implementation of Strategic Business Management* (pp. 189–198). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-82763-2_17

[32] Taghipour, A. (2021). Demand Forecasting and Order Planning in Supply Chains and Humanitarian Logistics. *Demand Forecasting and Order Planning in Supply Chains and Humanitarian Logistics*, 1–313. Hershey, PA, USA: IGI Global. https://doi.org/10.4018/978-1-7998-3805-0

[33] Tuarob, S., Wettayakorn, P., Phetchai, P., Traivijitkhun, S., Lim, S., Noraset, T., & Thaipisutikul, T. (2021). DAViS: a unified solution for data collection, analyzation, and visualization in real-time stock market prediction. *Financial Innovation, 7(1)*, 56. https://doi.org/10.1186/s40854-021-00269-7

[34] Wang, J., Liu, G. Q., & Liu, L. (2019). A Selection of Advanced Technologies for Demand Forecasting in the Retail Industry. *2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)*, (pp. 317–320). https://doi.org/10.1109/ICBDA.2019.8713196

[35] Weng, T., Liu, W., & Xiao, J. (2020). Supply chain sales forecasting based on lightGBM and LSTM combination model. *Industrial Management & Data Systems, 120(2)*, 265–279. https://doi.org/10.1108/IMDS-03-2019-0170

[36] Xu, X., Li, L., Zhang, J., Wang, J., & He, S. (2018). Content tracking by leveraging hashtag and time information in Twitter social media. *Web Intelligence, 16(2)*, 113–122. https://doi.org/10.3233/WEB-180377

[37] Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management, 28(1)*, 180–182. https://doi.org/10.1016/j.ijhm.2008.06.011

[38] Yu, X., Liu, Y., Huang, X., & An, A. (2012). Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain. *IEEE Transactions on Knowledge and Data Engineering, 24(4)*, 720–734. https://doi.org/10.1109/TKDE.2010.269

[39] Zhang, H., Parikh, N., Singh, G., & Sundaresan, N. (2013). Chelsea won, and you bought a T-shirt: Characterizing the interplay between Twitter and e-commerce. *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)*, (pp. 829-836). https://doi.org/10.1145/2492517.2500302

[40] Zhao, X., & Keikhosrokiani, P. (2022, January). Sales Prediction and Product Recommendation Model Through User Behavior Analytics. *Computers, Materials and Continua, 70*, 3855–3874. https://doi.org/10.32604/cmc.2022.019750

[41] Huang, L., Dou, Z., Hu, Y., & Huang, R. (2019). Online Sales Prediction: An Analysis With Dependency SCOR-Topic Sentiment Model. *IEEE Access*, *7*, 79791–79797. https://doi.org/10.1109/ACCESS.2019.2919734

[42] Kechyn, G., Yu, L., Zang, Y., & Kechyn, S. (2018). Sales forecasting using WaveNet within the framework of the Kaggle competition. CoRR, abs/1803.04037. https://doi.org/10.48550/arXiv.1803.04037