Enhancing Retail Forecast Accuracy with AI

By Dr. Alexey Karpov

Associate Professor of Artificial Intelligence, National Research University – Higher School of Economics (HSE), Russia

1. Introduction

In today's consumer-centric retail market, where innovation and digitization are rapidly transforming traditional operating norms, correctly balancing inventory levels and supply chain requirements with constantly moving consumer demand for a better anytime, anywhere shopping experience is fast becoming a cornerstone for business supremacy, profitability, and even survival and stock valuation. Product unavailability is costing U.S. retailers $223 billion annually. To address this problem and to better communicate between buyers and suppliers, there is a need to capture the future requirements of the retailer accurately. In essence, the forecasting method of the retailer holds the key to forecasting accuracy and the success of almost all forecasting algorithms.

The prime goal of this study is to address the above challenge and to enable retail forecasting technology to keep pace with AI technologies. The objective of this paper is to show how the presence of an expert AI retailer layer, positioned atop a planning system, can keep the system's enhanced forecasting tool at the cutting edge of forecasting science. Worldwide, the artificial intelligence market is being propelled at an unprecedented pace; the AI market is projected to grow at a significant CAGR of 61.5% between 2017 and 2022 and become a $16.06 billion market by 2022. In this paper, we address a new area of application pertaining to the deployment of artificial intelligence technologies in retail, principally forecasting, and espouse its major transformational benefits and challenges. The annotated bibliography summarizes recent contributions to forecasting, AI for retail, demand and sales forecasting, inventory forecasting, and data exchange under the six clear sub-themes as shown in the mind map - The Evolution of Forecasting and Data Exchange.

1.1. Background and Importance of Forecast Accuracy in Retail

Forecast accuracy is a critical factor in a successful retail operation. With accurate forecasts, retailers can better plan inventory levels and navigate the supply chain, make better-informed decisions about promotional events, pricing, and marketing strategies, and avoid unnecessary stockouts or excess inventory at the store level. This leads to savings throughout the entire retail value chain, especially in minimizing the consequences of overstocks and minimizing lost sales due to stockouts. Retail demand is influenced by a multitude of factors including seasonality and other cyclical effects, changes in the economy, and consumer behavior. In order to handle these demand-shaping factors, retailers need accurate forecasts to anticipate how much of what products they will sell. They must also be able to adjust these plans swiftly to continue to thrive in an ever-evolving retail market. Making such capacity decisions with the goal of accurately forecasting demand has been a longstanding challenge for retailers. Historically, forecasts have been determined based on expert judgment, historical data, demographic data, and/or the retailer's own data. Although such quantitative forecasting methods involved simple adjustments to account for current trends, they were not robust against complex phenomena in such data. As retail demand feedback signals have become more intricate, forecasting has become increasingly difficult with the added complexities and mixture of autoregressive terms necessary to predict today's outcome based on a time series. These feedback signals are problematic for simpler forecasting techniques, and thus, newer, more flexible techniques enable a retailer to handle this problem to create more accurate forecasts.

2. Fundamentals of Retail Forecasting

In retail operations, accurately anticipating customer demand is key. Forecasting, in its basic form, minimizes total costs by supporting prediction-based decisions regarding staffing, capacity, and inventory quantities. By securing the right quantity of inventory at the correct location and time – while reducing the risk of excess stock – a retailer is able to avoid running out of stock. Forecasting the level of customer demand in a retail supply chain is only the starting point in terms of operational processes. A forecasting model must produce as its output the level and uncertainty (both timing and quantity) in customer demand. The product of forecast accuracy and the amount of stock results in the retailer's level of uncertainty regarding the exact timing of a customer purchase.

In a general setting, forecasting models may require one or more of the following data as input:

Feature/Series: Generally, models require (or make use of) the past sales data of a given product or group of products. It is also crucial for the models to be sensitive to wear and tear of products. Market/Trend Data: External market data could be used for forecasting, like inflation, which would have an impact on retail. Consumer Behavior Data: For instance, types of promotions due to their impact on sales and their computation. Weather Data: This could be one of the auxiliary data that might be introduced. Incorporating weather data in retail forecasting and topology design. Short-term forecasts can guide decisions concerning ordering and stocking from a distributor; whereas long-term forecasts can influence supply chain strategy, including the decision concerning domestic production and the logistics of global sourcing.

In the context of retail supply chain forecasting, the notion of adopting the "one-number forecast game" within supply chain partners has become the consensus, where there is alignment of forecasting planning among the partners of the chain. Given competitive markets, this kind of collaborative decision-making that is modeled as lead time, order quantity, or capacity by supply chain partners is to be determined in response to forecast contributions, which may also inhibit the potential for increased sharing of information and more accurate forecasting and planning. In the last 10 years, with a dramatic increase in the power of computers, food and perishable supply chains have relied heavily on quantitative forecasting methods. More recently, however, the development of methods has provided new hope regarding human decision-making behavior and forecasting, as it hopefully does not resolve everything in forecasting.

2.1. Traditional Methods vs. AI-Driven Approaches

For several decades, retail forecasting's goal has been to quantify uncertainties around sales predictions and find ways to reduce them by using historical data and quantitative forecasting methodologies. The development of demand forecasting has evolved remarkably, and future trends and technologies must be considered carefully. By contrast, traditional forecasting methods often rely heavily on historical data, such as past sales figures, and lack the tools provided by machine learning algorithms to engage in predictive analytics more effectively. In addition, these traditional methods tend to be handled manually rather than by algorithms. In forecasting, machine learning and deep learning models have been shown to have a higher magnitude of memory processing and to be better at handling non-stationary data than many traditional time series approaches. They allow the storage of large structured or unstructured databases and data distributed over multiple machines. Furthermore, these deep learning models can be continuously recalibrated with real-time data to detect new trends, while the traditional methods are unable to cope with the vast information available nowadays, such as social media, sound, image, and video data. Overall, this shows how an automated machine learning forecast can assist in improving forecast accuracy by identifying patterns and correlations in big data that are difficult to uncover manually. In particular, this forecast aids in planning and may contribute to reducing waste and empty stocks, as well as enhancing inventory turnover significantly.

3. Machine Learning Models for Demand Prediction

Introduction Several different machine learning models have proven to be beneficial when used for demand prediction in retail environments due to their capacity to analyze complicated datasets. They can identify subtle patterns, interrelationships, and exclude those features that are irrelevant. Furthermore, decision-makers can easily interpret the insights provided by these models and act upon them. These models aim to minimize the mean squared error and price sensitivity; as such, they allow proper adjustments in price and stocks. Moreover, the advantage of these neural network models is their capacity to predict a complicated, dynamic, and irregular system, with the potential to learn through experience and improve their predictions over time. It raises the forecast accuracy in contrast to traditional statistical methods. These algorithms belong to the category of supervised approaches, which make forecasts based on historical data. The aim of this research is to speculate on which desired future event will take place. The approaches discussed in this paper optimize solutions and adjust parameters so that a desirable end result is obtained.

These models have learned from historical data in a manner that the higher the price, the less the desire to purchase certain products. The models then will use recent data processing that is collected every hour in the supply chain in conjunction with rewards and penalties feedback. For instance, when an item is out of stock, the models expect a penalty from the customer in the following weeks. The results show a two-digit increase in forecast accuracy. Nonetheless, due to the complexity of using these algorithms in project supply chain, the forecasting models will run consecutively each hour, giving results within a 20-minute time frame. Demand prediction should be accurate and reliable as it affects essential processes in a company, such as planning customer service levels and inventory management, as automatic replenishment systems in retailers continue to grow. It is, however, an aggressive competitive advantage for the company that is able to offer its customers the right product at the right time, effortlessly requiring no action.

3.1. Types of Machine Learning Models in Retail Forecasting

Machine learning models are increasingly leveraged to improve forecast accuracy in the retail sector. Broadly, two categories of models are prevalent: supervised learning models and unsupervised learning models using clustering techniques. The choice of model has significant implications for forecast accuracy as well as computational efficiency. When data is sufficiently large and diverse, models are generally moved to an ensemble setting in order to capture the advantages of individual models while mitigating the limitations of each model. Ensemble methods use a combination of learning algorithms to perform better, capturing more patterns without compromising computational time. Such models are capable of modeling unique forecasts and meaningful aspects of the data.

Other than mean absolute percentage error, different types of organizational goals play a crucial role in model selection. For example, linear regression models ought to be selected when interpretability is required; in contrast, the more complex network models are chosen to incorporate higher forecasting accuracies. Other considerations include computational time, the volume of data available, and the ability to employ advanced computing resources. For lower volumes of data, a relatively easy technique such as linear regression models could be employed. In general, such selection results in any retail organization looking to adopt AI-based forecasting gaining insights based on the pertinent characteristics of the forecasting problem.

4. Case Studies and Applications

This section illustrates the practical applications of AI in enhancing forecasting accuracy in the retail market. The various case studies help to furnish important conference pillars from different industry and local perspectives.

4.1 Introduction and Objectives This section presents the retail applications of AI in enhancing forecasting accuracy of different specialized demand requirements in practice. It highlights the AI-enhanced case experiences from four major retail sectors: (1) Run for Good campaign in shoe retailing, (2) consumer electronic products in retail, (3) social commerce in e-commerce, and (4) fashion and luxury retailing. The presented study cases span wide geographical and retail sectors to capture versatile demand characteristics from diverse practitioner lenses. For each practitioner case, the AI implementation process has been revealed with the gross benefits along with a discussion of the successes and various limitations. Metrics on the technical improvements attributed to AI framework implementation on baseline forecasting accuracies have been presented in each application.

4.3 Key Enablers and Success Factors - The AI model provides additional sales by identifying products with a 30% chance of stock-out. - Our teams now rely on AI insights and have started to use AI modules and scenarios for now-casting what if the scenario launching timing is given. Once developed, we realized that this gave us the agility to quickly divert some of our media buys to slow-moving lines. - The foot traffic boost for a special launch of a running shoe comes from multiple activations.

4.1. Real-World Examples of AI Implementation in Retail Forecasting

Until now, the argument for AI in demand forecasting has been analyzed from a more general perspective, focusing on the results AI can achieve and comparing it to more traditional models. In the following, we want to give insights into AI for forecasting in the retail sector in terms of concrete use cases of companies active in this area. In this first part, the AI application, especially in terms of technological aspects, is presented. We will highlight the challenges that have been specifically addressed with the introduction of AI technology, as well as its expected effect on day-to-day business processes.

AI in Retail Forecasting: Keller Logistics Group is a 3PL and transportation services company. They have been active for more than 40 years, and until a few years ago, forecasting demand for customers was done manually. Given the size of a typical retailer's product range and the difficulty of calculating sales for new items, only the most popular items were forecasted as a matter of course. This is only 30-40% of sales on average. Not forecasting 60%-70% led to severe under-stocking of these items. Keller Logistics Group could have embarked on an ambitious project to upgrade its ERP systems, train planners, or even hire a heavyweight consulting firm to remedy historical demand forecasting failures. Instead, we turned to AI to give us accurate baseline sales data to anchor our demand forecasts and got it for 90% of our items in a matter of months.

5. Challenges and Future Directions

Challenges Data privacy and security concerns. One of the main concerns for retailers using AI to access external support in their forecasting processes is the fear of leakage of private and sensitive information. Furthermore, retailers must comply with the constraints imposed by data protection regulations whenever personal, sensitive, or identifying information is involved. Integration with existing solutions and obsolete, buggy, and not suitably programmed solutions are also pressing concerns. The same issues could indeed haunt existing systems whenever a possible integration is considered. Costs and ease of use. Although AI solutions for time series forecasting are now actively available, some costs could still be high, and technicians usually need specific know-how. Moreover, the availability of these experts is not widespread. These solutions indeed have to be proposed by individuals who have strong expertise in the field.

Data-related problems. Collecting a vast quantity of data is not a guarantee of prosperity. To properly use artificial neural networks and further machine learning algorithms, data quality is a critical issue. Data collected in improper ways, or that is affected by noise, can indeed provide dubious help to forecasters. Data trustworthiness must be guaranteed. Future directions. A great deal of AI potential for retail forecasting has still to be explored. Data availability and computational power have indeed paved the way to the so-called big data paradigm, which is expected to make significant progress in retail forecasting. The moderate role of domain knowledge is also considered an excellent environment to automate the forecasting process. Recently, however, alternative strategies such as the hybrid approach have been proposed. The choice of partnership and construction plays a significant role in a successful, or not, forecasting application. Further research in this sense should help researchers and practitioners understand a comprehensive set of insights useful to enrich the application of machine learning technologies to retail forecasting.

5.1. Key Challenges in Implementing AI for Retail Forecasting

Independent retailers are facing specific challenges when implementing AI applications for forecasting use cases. The highly disparate IT landscape, along with a lack of data standards, renders many old data sources inappropriate for analysis. As a consequence, retailers are incurring high efforts in terms of data engineering to gain access to integrated databases. Additionally, the existing technology stack, based on traditional in-house and vendor-developed systems, is not capable of executing AI-based applications. Introducing these applications on top of the existing architecture requires major changes in terms of system performance, user competencies, or hardware investments. In some cases, the total cost of adopting AI is significantly higher than for greenfield scenarios, as new or tighter defined regulations require implementing additional data protection measures, process integrations, or ethical councils.

One of the toughest challenges in implementing AI is aligning staff to implement the new tools and aligning stakeholders to support them. Regulatory environments will require monitoring AI decisions after implementation, leaning on retailers' judgment process. From an organizational point of view, obtaining a binding commitment from stakeholders in managing and using AI-based forecasts is an essential condition for the successful scaling of AI in forecast processes. Thus, significant project investments have to be made in hiring external consultancy services or in training campaigns. Data engineers need to attend specialized programming courses to set up and run AI or to benefit from external consultancy services. Since organizations can't upscale and formalize the system design on a regular basis, the forecast has to rely on manual expertise. Converging stakeholders' opinions for a go/no-go decision on each blueprint of the implementation scenario is another point proving the challenge. While big retailers still have financial resources and scale mechanisms to take the challenge and implement disruptive changes that leverage technology to its full potential, many others are limited by existing systems, immediate financial requirements, or the capability to train and educate their workforce.

During the conduct of our project, we were continuously looking for strategies and approaches to overcome these challenges and benefit from the insights of the best-in-class retailers who strive to use AI to improve forecasting.

5.2. Potential Future Developments and Trends

This section will provide insight and understanding for the potential future developments and trends in AI and their implications for forecasting in retail.

The future of small data is moving towards advanced analytics and predictive modeling with automation. New autonomous forecasting models are already being developed, reducing the need for human intervention when using machine learning. Automation, along with anticipation of consumer behavior and company expectations, is expected to be a powerful driver that will change the way forecasts are applied now. The prediction models that use current data in real-time can be produced and evaluated, which will reduce errors in the delivery of goods and services. Further developments in machine learning algorithms are expected to lead to more accurate and faster predictions about requirements.

The integration of IoT will also be a trend. This trend assumes that retailers and customers can more accurately understand actual demand. It has been found that there is very little use of actual data from the IoT to make predictions in real time. This has a negative impact on the personalization aspect of retail. Evidence shows that the future of forecasting is moving towards a society where human intuition can coexist with artificial intelligence. There is also a general movement towards a more consumer-centric approach. Artificial intelligence can already be used to infer consumer behavior that is clearly relevant for forecasting. This can make a major change to the way the retailer is currently forecasting. The way we see these innovations is that technology is evolving all the time, and the way we will forecast is moving very agile.

6. Conclusion

This paper has sought to demonstrate the importance of effective forecasting for retail, along with the increasing levels of complex, univariate time series data gathered and utilized in that context. It has argued that the use of AI forecasting methods, which are able to detect and leverage time-varying phenomena from these data sources, can best facilitate the calculation of robust forecasting signals and help to improve the matching of supply and demand dynamics. The current state of retailing in the modern omnichannel context has been summarized to demonstrate the potential value of improved forecasting techniques. Regardless of the validity or scope of the arguments that have been presented herein, it is clear that some organizations have resistance to the deployment of these sophisticated methodologies within operations, with many preferring to rely upon high ROI, low complexity, time-proven solutions. Organizations looking to invest in AI methodologies for forecasting are encouraged to invest in the capabilities that can reduce risk, reduce project duration, and provide insight into the technologies, or develop them within realistic financial and organizational constraints. Technologies that offer solid but efficient experimentation design capabilities can be implemented into an organization's corporate strategy and used in tandem with existing forecasting processes. These long-term forecasting research agendas should be continuously informed by pragmatic methodologies and established methodology; the role of RP can shift as the processes become sufficiently understood. Every retail executive knows that abrupt changes in customer behavior can have a significant impact on the retailer's operations. A better awareness of future sales and roadblocks could help minimize these impacts and also help increase profitability. Artificial intelligence can help retailers quickly adjust to these challenges and remain competitive. With real-time forecasting, using sophisticated AI methodologies to capture the problem's time-varying features as effectively as possible, retailers will be prepared for whatever their customers demand. This should be the ultimate objective.

Reference:

1. S. Kumari, "AI-Driven Cloud Transformation for Product Management: Optimizing Resource Allocation, Cost Management, and Market Adaptation in Digital Products ", *IoT and Edge Comp. J*, vol. 2, no. 1, pp. 29–54, Jun. 2022

2. Tamanampudi, Venkata Mohit. "Automating CI/CD Pipelines with Machine Learning Algorithms: Optimizing Build and Deployment Processes in DevOps Ecosystems." *Distributed Learning and Broad Applications in Scientific Research* 5 (2019): 810-849.

3. Machireddy, Jeshwanth Reddy. "Data-Driven Insights: Analyzing the Effects of Underutilized HRAs and HSAs on Healthcare Spending and Insurance Efficiency." Journal of Bioinformatics and Artificial Intelligence 1.1 (2021): 450-470.

4. J. Singh, "The Future of Autonomous Driving: Vision-Based Systems vs. LiDAR and the Benefits of Combining Both for Fully Autonomous Vehicles ", *J. of Artificial Int. Research and App.*, vol. 1, no. 2, pp. 333–376, Jul. 2021

5. Tamanampudi, Venkata Mohit. "AI and DevOps: Enhancing Pipeline Automation with Deep Learning Models for Predictive Resource Scaling and Fault Tolerance." *Distributed Learning and Broad Applications in Scientific Research* 7 (2021): 38-77.

6. Ahmed Qureshi, Hamza, et al. "The Promising Role of Artificial Intelligence in Navigating Lung Cancer Prognosis." International Journal for Multidisciplinary Research, vol. 6, no. 4, 14 Aug. 2024, pp. 1–21.

7. Singh, Jaswinder. "Deepfakes: The Threat to Data Authenticity and Public Trust in the Age of AI-Driven Manipulation of Visual and Audio Content." *Journal of AI-Assisted Scientific Discovery* 2.1 (2022): 428-467.

8. Tamanampudi, Venkata Mohit. "Autonomous AI Agents for Continuous Deployment Pipelines: Using Machine Learning for Automated Code Testing and Release Management in DevOps." *Australian Journal of Machine Learning Research & Applications* 3.1 (2023): 557-600.