

A Review of Machine Learning Applications in Business Forecasting

Chenxi Qiu

*School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong, China
Aqiu0508@gmail.com*

Abstract. In the context of the digital economy, consumer demand, platform traffic, and supply chain environments are constantly changing. Accordingly, business forecasting has laid a critical foundation for enterprise inventory scheduling, marketing resource allocation and operational risk management. Compared with traditional statistical forecasting methods, such as ARIMA and exponential smoothing, machine learning can process multi-source, high-dimensional, and nonlinear information through big data and algorithms. It has gradually become an important research direction in business forecasting. This paper adopts systematic literature review and typical case analysis, to compare machine learning with traditional forecasting methods, explore its practical value in sales prediction, customer churn early warning and demand fluctuation evaluation, and sort out corresponding operational, compliance and technical risks as well as standardized governance requirements. The findings show that machine learning can improve forecasting accuracy and dynamic responsiveness while further supporting inventory and supply chain coordination, marketing decision-making, customer segmentation, and personalized recommendation. However, its effectiveness is also constrained by data quality, model interpretability, privacy compliance, and technical security. To solve the above problems, enterprises need to build a complete governance system covering standardized data management, scientific model supervision and manual human oversight. This paper contributes to understanding the transition of business forecasting from the traditional statistical paradigm to the intelligent forecasting paradigm and provides practical guidance for enterprises seeking to balance technological application, operational optimization, and risk control.

Keywords: Machine Learning, Business Forecasting, Demand Forecasting, Decision Optimization, Risk Governance

1. Introduction

In the digital economy, consumer demand, platform traffic, the competitive environment, and supply chain conditions are constantly changing. Business forecasting has therefore become an essential basis for enterprises to manage inventory, marketing activities, and capital allocation. Early international studies mainly relied on statistical methods such as ARIMA and exponential smoothing

to identify trends and fluctuation patterns from historical data [1, 2]. These methods have clear structures and strong interpretability, but their adaptability is limited in high-dimensional, nonlinear, and rapidly changing scenarios. With the development of big data, research has gradually shifted toward machine learning, which uses algorithms to identify patterns in complex data and improve out-of-sample forecasting performance [3]. Although related research in China started relatively late, it has expanded into management, marketing, finance, and supply chain fields. However, limitations remain in its application to local contexts, in explaining how it works, and in translating predictions into decisions [4].

Based on this background, this paper addresses three questions: What forecasting advantages does machine learning offer compared with traditional methods? What value does it provide in applications such as sales forecasting, customer churn prediction, and demand fluctuation assessment? What risks and governance requirements arise during implementation? This study adopts a combination of literature review and case analysis. The literature review examines the development of existing research and identifies its limitations, while the case analysis evaluates representative business practices. The study aims to explain the transition of business forecasting from a statistical paradigm to an intelligent forecasting paradigm and to analyze how forecasting results are incorporated into inventory management, marketing, customer management, and risk control processes. It therefore provides a reference for enterprises seeking to strengthen forecasting capabilities, improve operational decision-making, and enhance risk governance. This paper does not aim to prove that machine learning is better than traditional methods in every situation; instead, it focuses on the application boundaries of different approaches and their alignment with specific business needs.

2. Differences between machine learning and traditional business forecasting approaches

Traditional business forecasting begins by specifying variable relationships, model structures, and parameters based on theory and experience, followed by estimation and testing using historical data. In contrast, machine learning places greater emphasis on allowing algorithms to automatically identify patterns from large volumes of data and improve forecasting results through training and iterative optimization [3]. Therefore, the difference between the two approaches lies not only in the tools employed but also in their underlying forecasting logic. Traditional methods follow the logic of "specifying the model first and then validating it with data", whereas machine learning focuses on "learning patterns from data".

From the perspective of data foundations, traditional methods mainly rely on structured historical data, such as sales, prices, costs, and time cycles. Machine learning, by contrast, can further integrate multi-source information, including user behavior, review text, platform traffic, promotional activities, and weather conditions, to identify complex nonlinear relationships. In terms of research objectives, traditional methods emphasize explaining relationships between variables and model parameters, whereas machine learning focuses on predictive performance on unseen samples. Regarding operational mechanisms, traditional models often require researchers to make periodic adjustments, while machine learning models can be iteratively updated with new data and evaluated through cross-validation and rolling-window testing to assess their actual forecasting performance [2].

The two approaches should not be viewed as simple substitutes for one another. Traditional models offer advantages such as clear structures, interpretable parameters, and relatively low implementation costs. Machine learning performs better in high-dimensional data processing, dynamic iteration and complex rule identification. However, it also has obvious limitations

including black-box prediction, high data dependence and increased governance risks and costs. Therefore, machine learning is more appropriately regarded as an extension of traditional business forecasting approaches. When selecting forecasting methods, enterprises should combine different approaches according to data volume, business stability, interpretability requirements, and implementation costs, rather than simply pursuing more complex algorithms.

2.1. Classical approaches and limitations of traditional business forecasting

Traditional business forecasting mainly relies on historical data analysis and empirical managerial experience. Common methods include regression analysis, moving averages, exponential smoothing, and ARIMA. When market conditions are relatively stable and the data have relatively few dimensions, these methods can provide fundamental support for sales forecasting, inventory planning, and operational planning [1, 2]. However, their limitations are mainly reflected in the following three aspects.

Managerial judgment can easily introduce subjective bias. Enterprises usually depend on manual forecast adjustment in scenarios such as insufficient historical data, new product launches and sudden market fluctuations. Fildes and Goodwin examined more than 60,000 forecasting records from four companies and found that approximately 75% of statistical forecasts had been manually adjusted, although some of these adjustments did not improve forecasting accuracy [5]. Biases such as optimism, anchoring, and overreaction may also lead managers to misinterpret short-term growth as a long-term trend, resulting in excessive inventory and inefficient resource allocation.

Limited adaptability to sudden market changes. Traditional methods generally extrapolate future outcomes from historical trends and seasonal patterns. When promotional campaigns, platform rules, or consumer behavior change abruptly, previously identified patterns may quickly become invalid. Alibaba's research team found that during the "6.18" shopping festival, the online AUC of a conversion-rate prediction model trained under normal conditions declined by up to 10%. At one point, the ratio of predicted values to actual values fell to only about one-third. This finding proves that data distribution shifts caused by promotional activities greatly reduce the prediction performance of traditional statistical models [6].

Limited capacity to process multi-source, high-dimensional data. Modern retail demand is simultaneously influenced by factors such as prices, promotions, page traffic, brands, and user visits, making e-commerce forecasting a problem shaped by multiple variables. Traditional low-dimensional models struggle to fully integrate such information and cannot promptly capture complex changes in demand. Overall, traditional business forecasting remains suitable for scenarios involving small data volumes and stable market relationships. However, it is vulnerable to managerial bias, sudden environmental changes, and increasing data dimensionality, and this provides a reason for enterprises to adopt machine learning.

2.2. Technical characteristics and advantages of machine learning

Machine learning has three technical advantages. First, it can process high-dimensional, multi-source, and nonlinear data. Second, it emphasizes out-of-sample predictive performance and can be improved through training, validation, and iterative optimization. Third, it can integrate external variables, such as promotional activities, platform traffic, customer reviews, and weather conditions, into a unified framework, making forecasting more closely aligned with real business environments [3].

In practical applications, machine learning aims not only to improve forecasting accuracy but also to support faster responses and more efficient resource allocation. The authoritative M5 forecasting competition adopted 42,840 hierarchical Walmart sales datasets to test the performance of complex retail forecasting models. The results showed that some machine learning and hybrid methods outperformed established statistical approaches in specific scenarios [7]. This indicates that the advantages of machine learning lie in the combined efficiency of forecasting, model updating, and operational execution, rather than simply increasing model complexity. However, these advantages depend on a stable data foundation, continuous maintenance, and appropriate evaluation metrics. When data quality is poor, complex models may still produce substantial forecasting errors.

Furthermore, machine learning has changed the way forecasting work is organized. Traditional forecasting usually requires analysts to periodically consolidate data and manually update reports. In contrast, machine learning systems can connect data integration, feature calculation, model updating, and result delivery, thereby shortening the time between market changes and business responses. However, automation does not eliminate the need for human management. Enterprises need to formulate unified data standards, standardized anomaly handling procedures and reasonable model update cycles, so as to avoid error accumulation and amplification in automated forecasting processes.

2.3. Practical validation of business forecasting applications

The feasibility of machine learning in business forecasting has been demonstrated across multiple scenarios. First, in sales forecasting, machine learning can integrate historical sales, prices, promotional activities, holidays, and platform traffic. This enables the model to effectively capture nonlinear demand fluctuations in complex retail scenarios. Related studies indicate that these methods can support inventory management, pricing, and production planning [7].

Second, in customer churn prediction, machine learning can screen potential lost users according to multi-dimensional user characteristics including consumption frequency, platform activity, complaint records, payment habits and service interaction data. Systematic reviews show that these methods have been widely applied in telecommunications, banking, e-commerce, and subscription services, helping enterprises implement customer-retention measures and differentiated operations in advance [8].

Finally, in demand fluctuation forecasting, machine learning can incorporate weather conditions, promotional activities, platform traffic, customer reviews, and external environmental factors to identify early signals of rising or declining demand. Relevant research confirms that such algorithm-based methods adapt well to volatile supply chain environments with complex variable interactions. They can help enterprises adjust procurement, inventory, and production plans while reducing the risks of stockouts and excess inventory [9].

Overall, machine learning has evolved from a theoretical approach into a practical business tool. Its value lies not only in improving forecasting accuracy but also in translating predictive results into actions related to inventory optimization, customer retention, and resource allocation. The above three practical scenarios further illustrate that the practical value of machine learning forecasting depends on its integration degree with real business processes, instead of merely relying on offline experimental accuracy.

3. The enabling role of machine learning in business forecasting

Machine learning enhances business forecasting primarily at three levels: operational efficiency, decision quality, and differentiated services. By integrating multi-source data and identifying complex relationships, it shifts forecasting from the extrapolation of historical trends toward dynamic learning. It also embeds model outputs into inventory management, marketing, risk management, and customer service processes. The three dimensions are mutually correlated and mutually supportive. Optimized inventory and supply chain management lay a solid operational foundation; precise marketing and risk assessment realize efficient resource allocation; refined customer segmentation and recommendation systems further expand the service boundary of forecasting models.

3.1. Improvement in operational efficiency

The core role of machine learning is to improve inventory allocation and supply chain coordination. Enterprises can integrate sales, prices, promotions, seasonal changes, and external conditions to generate more detailed demand forecasts for different products. These results can then support procurement, replenishment, safety-stock management, warehousing, and distribution. Compared with approaches that rely on historical averages and managerial experience, this model helps reduce stockouts, excess inventory, and inefficient inventory transfers.

Alibaba applied deep learning-based demand forecasting and inventory optimization models to its retail operations, creating an integrated process linking demand forecasting, replenishment, and inventory allocation. Relevant research data show that this intelligent system can save about 42 million US dollars in annual inventory costs and reduce product waste [10]. This demonstrates that the operational value of machine learning lies not only in improving forecasting accuracy but also in translating forecasts into inventory-control and supply-chain execution plans, thereby shifting operations from reactive adjustment to proactive optimization. When forecasting results are integrated with procurement cycles, warehouse capacity, and suppliers' delivery capabilities, enterprises can also reduce emergency inventory transfers and costly urgent purchases, thereby strengthening supply chain resilience.

3.2. Optimization of decision quality

Machine learning can improve the accuracy of marketing resource allocation and business risk assessment. By analyzing customer behavior, purchase intentions, channel feedback, and risk characteristics, models can provide a data-driven basis for selecting target audiences, allocating budgets, and formulating risk responses.

In marketing, Sift Lab uses predictive analytics and customer segmentation to identify high-value customers, repeat purchasers, and customers at risk of churn, allowing marketing resources to be directed toward segments that are more likely to generate returns. The system increased the speed of analysis and campaign setup tenfold, improved return on advertising spend by 40%–50%, and raised annual customer revenue by 5%–7% [11].

In risk assessment, Alonso and Carbó compared methods including Logit, random forests, and XGBoost using anonymized data from a major Spanish bank. The results showed that machine learning models generally achieved better default-prediction performance, although greater model complexity did not necessarily produce better results [12]. Therefore, machine learning can improve marketing and risk-management decisions, but interpretability and model validation must also be

considered. Especially in core business links such as marketing budget allocation and credit risk evaluation, enterprises need to comprehensively consider prediction accuracy, misclassification loss and actual decision consequences, instead of relying merely on a single evaluation indicator for model selection.

3.3. Development of differentiated service capabilities

Machine learning can strengthen differentiated service capabilities via customer segmentation and personalized recommendations. Enterprises can classify customers into high-value, potential-churn, or specific-needs groups based on purchase frequency, spending levels, activity, lifecycle stage, and preferences. They can then adjust service priorities and communication strategies accordingly.

In personalized recommendation systems, machine learning continuously analyzes customers' historical interactions and real-time behavior to predict which products are most likely to interest them. After deploying the Amazon Personalize algorithm system, beauty retail brand MECCA produced over 10 million personalized recommendations every week. This practice increased the click-through rate of personalized emails by 65% and brought a 76.4% growth in corresponding email-driven revenue [13]. This case demonstrates that customer segmentation identifies differences among customers, while personalized recommendations translate predictive results into specific services, thereby improving customer experience, conversion performance, and customer relationship management. However, enterprises must also control recommendation frequency and establish clear boundaries for information use to avoid excessive contact that may cause customer dissatisfaction.

4. Risks and governance of machine learning in business forecasting

While machine learning improves forecasting capabilities, it also introduces new forms of uncertainty, which can be broadly classified into business, compliance, and technical risks. Business risks mainly arise when data or model bias affects credit decisions, inventory management, supply chains, and customer management. Compliance risks mainly focus on personal privacy protection, algorithmic fairness and reasonable responsibility attribution in model application. Technical risks include black-box models, data poisoning, model drift, and dependence on third-party systems [14].

Accordingly, enterprises should not merely pursue model prediction accuracy. Full-process risk management should be implemented throughout data collection, model development, result deployment and long-term dynamic monitoring. These three categories of risk may also reinforce one another: data problems can lead to model misjudgments, which may subsequently result in business losses, disputes over consumer rights, and regulatory liability.

4.1. Business risk forecasting capabilities

Machine learning can integrate multidimensional data to provide earlier and more detailed forecasts of key business risks. In loan default prediction, models can combine credit history, income, liabilities, transaction records, and repayment performance to estimate default probabilities and support credit approval, pricing, and post-loan early warning [12]. In supply chain disruption forecasting, enterprises can use supplier performance, delivery cycles, quality fluctuations, inventory levels, and logistics information to adjust procurement plans in advance, allocate safety stock, and assess alternative suppliers.

When forecasting declines in market demand, data such as sales, prices, promotions, traffic, customer reviews, and macroeconomic conditions can be used to identify signs of weakening demand, helping enterprises adjust production, inventory, and marketing strategies [9]. In customer churn prediction, features such as purchase frequency, activity levels, complaints, and payment anomalies can help identify high-risk customers and support targeted retention measures [8].

These four types of risk forecasting differ in their objectives. Loan default prediction focuses on identifying financial loss probability; supply chain disruption forecasting aims to secure response time; demand decline forecasting emphasizes controlling inventory and production capacity; and customer churn prediction seeks intervention before the relationship ends. Therefore, enterprises should not evaluate all risks using a single threshold. Instead, they should define different model indicators and action rules according to the scale of potential losses, the required warning lead time, and the cost of intervention.

Machine learning therefore shifts risk management from reactive response toward proactive early warning. However, predictive results must still be applied in combination with business context and human judgment. Enterprises should also establish different warning thresholds according to the severity of each risk, preventing frequent false alarms from disrupting normal operations while ensuring that high-risk signals are not overlooked.

4.2. Ethical and privacy compliance challenges

Business forecasting often involves transaction records, browsing behavior, location data, and platform interaction data. When enterprises fail to provide adequate information about the scope, purpose, and methods of data collection and processing, users' rights to be informed and to make choices may be undermined. The large data dependency of machine learning may induce enterprises to collect excessive user data, repurpose data beyond original intentions, and conduct unauthorized secondary usage. When predictive results are applied to targeted marketing, customer segmentation, or risk identification, insufficient model transparency and lack of manual review may trigger algorithmic discrimination, improper user profiling and ambiguous responsibility allocation.

Therefore, enterprises should protect users' rights to be informed, to provide consent, to withdraw consent, to access data, and to request corrections. They should also follow the principles of purpose limitation and data minimization, while establishing transparent explanations, human review, and appeal mechanisms for high-impact applications [15]. Compliance requirements should not remain confined to privacy policies but should be implemented at the levels of data fields, model purposes, and business processes. For high-impact applications such as customer segmentation, personalized pricing, and automated risk assessment, enterprises should also explain the main processing logic and potential consequences to reduce information asymmetry.

4.3. Potential risks in technical applications

Machine learning also faces risks related to data poisoning and systemic failure. Contaminated or maliciously manipulated training data may cause systematic bias in model outputs, potentially leading to inventory imbalances, misdirected marketing expenditure, or inaccurate credit decisions. Model drift may also cause previously effective forecasting models to gradually lose validity as market conditions change [14].

When many enterprises rely on similar data sources, algorithms, and third-party platforms, model errors, technical failures, or misinterpretations of market signals may be amplified simultaneously, resulting in convergent decisions and cascading risks. Enterprises should therefore pay attention to

data-source security, anomaly detection, supplier dependence, and emergency alternatives to prevent forecasting systems from becoming new single points of failure. Critical models should also include rollback mechanisms and human takeover procedures, enabling essential business operations to continue when data anomalies or service disruptions occur.

4.4. Risk governance framework

To address the risks discussed above, enterprises should establish a three-layer framework consisting of data governance, model governance, and human supervision. First, data governance. Enterprises should clearly define data sources, purposes, retention periods, and access permissions while adhering to the principles of purpose limitation and data minimization. Chinese e-commerce and retail enterprises should classify and manage membership information, transaction records, browsing histories, and location data. They should not collect excessive data merely to improve model performance. In personalized recommendation and commercial marketing activities, enterprises should explain how data are used and provide users with convenient options to disable personalized recommendations or refuse marketing communications [15].

Second, model governance. Enterprises should conduct accuracy, fairness, and security testing before model deployment. After model deployment, enterprises need to dynamically monitor prediction deviation, data drift and abnormal outputs, and completely retain training datasets, model versions and decision logs for traceability. In applications such as pricing, customer segmentation, and risk scoring, enterprises should examine whether models create unreasonable differential treatment for particular groups. They may also apply the NIST Artificial Intelligence Risk Management Framework to embed risk governance throughout model design, measurement, monitoring, and response processes [14].

Third, human supervision. Decisions that may significantly affect individual rights and interests, such as credit rejection, account suspension, differentiated pricing, and high-value customer identification, should be subject to human review, user appeals, and model-correction mechanisms. Retailers should refrain from generating unreasonable price discrimination based solely on user profiling, and ensure fair and transparent transaction rules for consumers under equivalent trading conditions [15].

Overall, governance should focus on establishing a full-lifecycle mechanism in which data are traceable, models are continuously monitored, and decisions are reviewable, thereby advancing technological innovation and accountability simultaneously. The three layers of governance should not operate independently. Instead, they should be jointly implemented by business, technology, legal, and audit departments, with different measures applied according to the purpose and potential impact of each model.

5. Conclusion and future outlook

Based on systematic literature review and typical case analysis, this study confirms that machine learning has transformed traditional business forecasting from simple historical trend extrapolation to a multi-dimensional, data-driven analytical paradigm. It has considerable practical value in areas such as sales forecasting, demand forecasting, customer churn prediction, and risk early warning. Its contribution lies not only in improving forecasting accuracy but also in optimizing inventory and supply chains, enhancing marketing and risk-management decisions, and supporting customer segmentation and personalized recommendations. At the same time, machine learning introduces

risks related to data bias, privacy compliance, black-box models, data poisoning, and systemic dependence. Therefore, governance capabilities are as important as forecasting capabilities.

This study has several limitations. First, it does not use primary enterprise data for model training and empirical testing, and its comparison of model performance relies mainly on existing literature and publicly available cases. Second, some cases were disclosed by technology service providers or enterprises and may emphasize successful outcomes while providing limited information about implementation costs and failures. Third, data foundations vary considerably across industries and enterprises, meaning that the conclusions may not apply equally to all small and medium-sized enterprises. In addition, technological, regulatory, and market conditions continue to change, so the governance recommendations in this study may not remain valid over time.

Future research could use primary operational data from Chinese enterprises to conduct empirical comparisons between traditional forecasting models and machine learning models. It should also include more failure cases, cases involving small and medium-sized enterprises, and long-term performance evaluations. Further research could examine how different governance measures affect model trustworthiness, consumer acceptance, and enterprise performance. As multi-source data integration, explainable methods, and governance mechanisms continue to improve, machine learning is expected to develop beyond a forecasting tool into a trustworthy decision-support system. Enterprises that achieve competitive advantage will not be those with only more complex algorithms, but those capable of balancing technological innovation, practical implementation, and risk governance.

References

- [1] Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.
- [2] Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. <https://otexts.com/fpp3/>
- [3] Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199-231. <https://doi.org/10.1214/ss/1009213726>
- [4] Liu, J., Zheng, C., & Hong, Y. (2023). How can machine learning empower management research? A review of international frontiers and future prospects. *Management World*, 39(9), 191-216.
- [5] Fildes, R. A., & Goodwin, P. (2007). Good and bad judgment in forecasting: Lessons from four companies. *Foresight: The International Journal of Applied Forecasting*, 8, 5-10.
- [6] Chan, Z., Zhang, Y., Han, S., Bai, Y., Sheng, X.-R., Lou, S., Hu, J., Liu, B., Jiang, Y., Xu, J., & Zheng, B. (2023). Capturing conversion rate fluctuation during sales promotions: A novel historical data reuse approach. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 3774-3784). Association for Computing Machinery. <https://doi.org/10.1145/3580305.3599788>
- [7] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38(4), 1346-1364. <https://doi.org/10.1016/j.ijforecast.2021.11.013>
- [8] Imani, M., Joudaki, M., Beikmohammadi, A., & Arabnia, H. R. (2025). Customer churn prediction: A systematic review of recent advances, trends, and challenges in machine learning and deep learning. *Machine Learning and Knowledge Extraction*, 7(3), Article 105. <https://doi.org/10.3390/make7030105>
- [9] Douaioui, K., Oucheikh, R., Benmoussa, O., & Mabrouki, C. (2024). Machine learning and deep learning models for demand forecasting in supply chain management: A critical review. *Applied System Innovation*, 7(5), Article 93. <https://doi.org/10.3390/asi7050093>
- [10] Deng, Y., Zhang, X., Wang, T., Wang, L., Zhang, Y., Wang, X., Zhao, S., Qi, Y., Yang, G., & Peng, X. (2023). Alibaba realizes millions in cost savings through integrated demand forecasting, inventory management, price optimization, and product recommendations. *INFORMS Journal on Applied Analytics*, 53(1), 32-46. <https://doi.org/10.1287/inte.2022.1145>
- [11] Google Cloud. (n.d.). Sift Lab turns complex data into real-time insights and marketing actions with AI. <https://cloud.google.com/customers/sift-lab>

- [12] Alonso, A., & Carbó, J. M. (2021). Understanding the performance of machine learning models to predict credit default: A novel approach for supervisory evaluation (Working Paper No. 2105). Banco de España. <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/21/Files/dt2105e.pdf>
- [13] Amazon Web Services. (n.d.). MECCA: Using machine learning to deliver a highly tailored customer experience. <https://aws.amazon.com/solutions/case-studies/mecca/>
- [14] National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
- [15] Standing Committee of the National People's Congress. (2021). Personal Information Protection Law of the People's Republic of China.