International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 9 Issue 6, Nov-Dec 2025 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Research on Intelligent Dynamic Pricing Model of Small and Medium-sized Enterprises Based on Multi-dimensional Data

Manru Dong, Wangtianchen Zhu, Yuanlong Xu, Ping Ma, Haoyan Li, Zhangcheng Mu

School of Systems Science and Statistics, Beijing Wuzi University, Beijing, China

ABSTRACT

This study addresses critical pain points in pricing decision-making for SMEs, including insufficient data value mining, delayed dynamic strategy response, and high technical application barriers. We innovatively propose a hybrid intelligent dynamic pricing model integrating econometric analysis frameworks with machine learning prediction capabilities. Leveraging multi-dimensional enterprise datasets from the National Business Management Skills Competition, we systematically developed an integrated technical framework encompassing "data preprocessing—feature engineering—model construction—system implementation". Specifically, by employing panel data fixed-effects models to effectively identify causal effects in price fluctuations, combined with XGBoost's robust nonlinear fitting capabilities to handle complex variable relationships, we constructed a high-precision hybrid prediction architecture. Experimental results demonstrate that the model maintains an average absolute error below 5% and achieves a goodness-of-fit index exceeding 0.92, demonstrating excellent predictive performance. To enhance model interpretability and practicality, we introduced SHAP value analysis for feature contribution evaluation and developed a lightweight decision support system based on the Streamlit framework, enabling core functions such as price elasticity simulation and dynamic strategy generation. Empirical analysis shows the model significantly improves pricing decision efficiency and scientific rigor for SMEs. During pilot applications in manufacturing enterprises in the Yangtze River Delta region, it successfully helped companies achieve an average 15% profit growth. This study not only provides a feasible and easy-to-use digital pricing solution for small and medium-sized enterprises, but also explores a new path to transform the achievements of management discipline competition into real productivity.

How to cite this paper: Manru Dong | Wangtianchen Zhu | Yuanlong Xu | Ping Ma | Haoyan Li | Zhangcheng Mu "Research on Intelligent Dynamic Pricing Model of Small and Medium-sized Enterprises Based on Multi-dimensional Data" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-9 | Issue-6, December 2025, pp.654-661, URL: www.ijtsrd.com/papers/ijtsrd99945.pdf

IJTSRD99945

KEYWORDS: *Dynamic pricing; multi-dimensional data; machine learning; small and medium enterprises; decision support systems.*

INTRODUCTION

With the deepening integration of the digital economy and the real economy, the overall scale of China's digital economy in 2023 has exceeded 52.2 trillion yuan, accounting for a high proportion of 41.5% of the gross domestic product (GDP)[1]. Against this macro background, a large number of small and medium-sized enterprises (SMEs) have shown strong demand for digital transformation, yet they still face significant challenges in the critical process of pricing decisions. Multiple survey data indicate that only about 23% of SMEs have initially established basic data analysis systems, while over 60% of enterprises still primarily rely on managerial experience and intuition for pricing decisions, lacking scientific rigor and systematicness[2]. Traditional pricing methods, such as the cost-plus method and competitor-following pricing, which have been widely used for a long time, appear inadequate in the face of rapidly changing market environments. Especially in typical scenarios such as sudden demand fluctuations, price wars initiated by competitors, or drastic changes in supply chain costs, static pricing models are prone to issues like profit erosion and delayed market responsiveness.

Meanwhile, most existing business intelligence systems in the market typically involve high deployment costs, often requiring investments in the hundreds of thousands of yuan, and demand highly specialized technical teams. This has deterred many small and medium-sized enterprises (SMEs), creating a reality where "data silos" and "technological gaps" coexist [3]. According to third-party research reports, nearly 78% of SME managers list "dynamic pricing capabilities" as one of the most urgent needs for current digital transformation, with over 80% of surveyed enterprises willing to adopt low-threshold, high-efficiency intelligent pricing tools to enhance market competitiveness and operational efficiency [2]. Driven by this practical demand and existing bottlenecks, developing intelligent pricing solutions that combine low cost, easy deployment, and high precision has become a crucial breakthrough and key lever for SMEs to achieve digital transformation.

The theoretical significance of this study lies in its pioneering establishment of a new paradigm for intelligent pricing research that integrates econometric models with machine learning algorithms. This approach effectively overcomes the linear assumptions inherent in traditional pricing models, proposing a dual-driven analytical framework of "data empowerment" and "dimensionality reduction through tool application". Practically, the research achieves three major breakthroughs: First, it significantly enhances corporate operational efficiency by building a data-driven intelligent pricing system that helps businesses precisely balance price sensitivity and profit margins, effectively reducing resource waste from ineffective promotions. Second, it substantially lowers the threshold for digital transformation, with the lightweight analytical tools developed reducing enterprise data analysis costs to one-fifth of traditional solutions, enabling county-level and small-to-medium enterprises to benefit from digital technologies at lower costs. Third, it promotes optimal allocation of social resources by guiding efficient market supply-demand matching through precise price signals, providing effective support for addressing regional overcapacity issues.

The core objective of this research is to develop an intelligent dynamic pricing model based on multi-dimensional operational data from enterprises, providing scientific and practical pricing decision support for small and medium-sized enterprises (SMEs). The research focuses on five key aspects: (1) Establishing standardized data governance processes to effectively address inconsistencies in cross-period and cross-regional data standards, thereby enhancing data quality and usability; (2) Building a decision feature library containing critical metrics such as price elasticity coefficients and advertising marginal revenue indices to provide robust feature support for the model; (3) Developing a hybrid pricing analysis framework combining panel regression and machine learning techniques, integrating the strengths of traditional methods with cutting-edge technologies; (4) Designing and implementing a lightweight pricing decision support system tailored for SMEs to improve practicality and operational efficiency; (5) Validating the model's effectiveness through enterprise pilot tests and continuously optimizing both the model and system based on real-world feedback.

1. Literature review

As a core technology in revenue management, dynamic pricing has long been a hot research field attracting joint attention from academia and industry. Early studies primarily focused on service industries like aviation and hospitality, developing optimization models and pricing strategies to address uncertainties in inventory and demand. With the rapid development of e-commerce, the scope of dynamic pricing research has gradually expanded to retail and other physical industries, with increasing emphasis on complex factors such as consumer behavior, competitive environments, and market structures influencing pricing decisions.

In recent years, scholars have increasingly recognized the limitations of traditional static pricing methods in dynamic market environments. Wang et al. (2022) demonstrated that static pricing often leads to significant profit losses when market conditions change rapidly, a problem particularly pronounced in supply chain contexts with volatile costs. In contrast, dynamic pricing can effectively balance supply and demand through real-time market responsiveness, thereby helping businesses maximize profits.

The rapid advancement of machine learning technology has provided new methodological support for pricing research. Compared with traditional econometric models, machine learning algorithms demonstrate significant advantages in handling high-dimensional data and identifying complex nonlinear relationships. Zhang et al. (2021) applied the random forest algorithm to demand forecasting tasks, finding its prediction accuracy significantly outperformed traditional linear regression models. As an efficient gradient-boosting ensemble learning algorithm, XGBoost excels in processing complex feature interactions and nonlinear relationships, and has been widely applied in research areas such as sales forecasting, price elasticity estimation, and dynamic pricing optimization.

However, the persistent 'black box' nature of machine learning models has long hindered their widespread adoption in corporate decision-making. To tackle interpretability challenges, model explanation techniques like SHAP value analysis have been developed. These methods quantify each feature variable's contribution to the final model output, significantly enhancing the transparency and comprehensibility of machine learning models. This breakthrough creates favorable conditions for their practical implementation in pricing decisions.

Notably, research on pricing tools for small and medium-sized enterprises (SMEs) remains relatively scarce. Most existing studies focus on the complex, resource-intensive pricing systems of large corporations, which fail to address the practical constraints common among SMEs—such as limited resources and technical limitations. Li et al. (2023) emphasize that SMEs urgently require lightweight pricing tools with low costs and ease of use to enhance decision-making capabilities and operational efficiency in increasingly competitive markets. Moreover, compared to large enterprises, SMEs generally face significant gaps in both data quality and quantity, which imposes stricter requirements on the robustness and generalization capacity of pricing models. Due to limited data scale and inconsistent collection standards, SMEs often struggle to directly apply traditional large-scale modeling methods. Therefore, when developing predictive and decision-making models tailored to their business characteristics, greater emphasis should be placed on algorithmic robustness and data noise tolerance.

In recent years, the maturation and widespread adoption of open-source technologies have made lightweight decision support tools increasingly feasible. The rise of low-code development frameworks like Streamlit has significantly reduced the complexity and costs of building data analytics applications. These tools empower small and medium-sized enterprises to deploy customized pricing support systems with efficiency and flexibility, enabling them to leverage limited data resources for data-driven decision-making. This approach enhances market responsiveness and competitiveness [10].

2. RESEARCH METHODS AND DATA

This study employs a four-stage progressive research paradigm— "data foundation, model-driven, tool implementation, and value validation" —to systematically establish a complete research loop from foundational data processing to practical application. The framework emphasizes the fundamental role of data quality, leverages advanced modeling techniques to enhance predictive performance, and utilizes lightweight tools to facilitate the translation of research outcomes. The proposed system's practicality and effectiveness are ultimately validated through real-world business metrics.

The research data originates from the structured dataset provided by the 9th National Business Management Skills Competition, a dataset renowned for its high authority and industry representativeness in both academic and practical circles. It covers four major product categories, five regional markets, and eight complete operational cycles, comprehensively documenting corporate operations in real competitive environments [3]. The dataset includes 32 key metrics such as product pricing, advertising expenditure, distributor count, inventory turnover rate, and net profit margin, totaling 150,000 records. This provides a robust data foundation for multi-dimensional and multi-scenario pricing modeling.

During the data preprocessing phase, the research team utilized the Pandas library in Python to perform large-scale data integration and cleaning. Key measures included: resolving data inconsistency caused by cross-period and cross-regional variations (e.g., standardizing currency units and applying inflation adjustments to enhance comparability); applying the 3σ rule to identify and eliminate outliers for improved data quality and consistency [11]; employing multiple imputation based on chain equations to effectively fill in missing values for specific fields; and standardizing measurement discrepancies across regional markets to eliminate systematic biases.

Building upon cleaned and standardized high-quality data, the study developed a decision-making feature set comprising 12 core metrics. The framework covers four key dimensions: pricing (price level, volatility, and elasticity coefficient), marketing (advertising intensity, marginal return index, and promotional frequency), market structure (Hirschman-Herfindahl index, concentration ratio, and regional competition intensity), and operational efficiency (inventory turnover rate, channel efficiency index, and customer satisfaction scores).

The study further leverages SHAP value analysis to quantitatively evaluate and rank the importance of various characteristics, identifying key variables with significant explanatory power for demand fluctuations [8]. The analysis reveals that three factors—price elasticity coefficient, advertising marginal revenue index, and market concentration—constitute the core variables influencing demand. Their cumulative contribution exceeds 60% of the total explanatory power, providing crucial evidence for subsequent modeling.

3. Model Construction and Implementation

This paper innovatively proposes a hybrid model architecture of "panel regression + machine learning" to address the practical challenges of complex data environment and noise in small and medium-sized enterprises.

The panel data fixed effect model is used in the base layer of the model, which controls the unobservable individual heterogeneity and time trend to enhance the consistency of parameter estimation.

$$Y_{it} = \alpha_i + \lambda_t + \beta X_{it} + \epsilon_{it}$$

In this framework, Y_{it} denotes firm i's demand quantity at period t, with α_i representing the individual fixed effect, λ_t the time fixed effect, and X_{it} the observable explanatory variable vector. The corresponding coefficient is denoted as β, while the random error term is represented by ϵ_{it}. By controlling for other variables, this model enables a relatively pure identification of price changes' causal effect on demand, thereby providing a reliable theoretical foundation for subsequent economic interpretation and strategic analysis [12].

The core component of the model employs the XGBoost ensemble learning algorithm for demand forecasting. To fully capture the complex nonlinear relationships and interaction effects between explanatory variables and response variables, this high-performance algorithm is introduced. XGBoost iteratively generates multiple decision trees and integrates them, effectively handling high-dimensional features while maintaining prediction accuracy and preventing overfitting. The objective function is defined as follows:

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^K \Omega(f_k)$$

Here, $l(\hat{y}_i, y_i)$ denotes the loss function that quantifies the discrepancy between predicted and actual values, while $\Omega(f_k)$ serves as the regularization term to control the complexity of each decision tree, thereby enhancing the model's generalization capacity [13].

The decision-making layer has developed a dynamic game-based pricing module grounded in Nash equilibrium theory. This module simulates how competitors' price adjustments in market environments influence our pricing strategies, determining optimal responses through price response function analysis to achieve Nash equilibrium. By employing Monte Carlo simulations for multi-agent game dynamics, it generates competitive landscapes encompassing diverse market scenarios, thereby endowing the final pricing strategy with dynamic adaptability and adaptive optimization capabilities [14].

During the model training phase, K-fold cross-validation was employed to enhance model stability and generalization performance. The original dataset was split into training and testing sets at an 8:2 ratio, with key hyperparameters optimized using grid search. To balance model accuracy and computational efficiency, a feature selection strategy was implemented by incrementally increasing feature quantity. The optimal feature subset was ultimately determined through SHAP value analysis and recursive feature elimination (RFE) [15]. Experimental results demonstrated that the feature subset selected through this process maintained predictive accuracy while reducing feature quantity by approximately 30%, significantly improving both computational efficiency and model interpretability.

A lightweight pricing decision support system was developed using the Streamlit framework, presenting algorithmic models through visual and interactive applications [16]. The system comprises three core modules: The data dashboard integrates multidimensional metrics including sales data, price trends, and market share, enabling dynamic filtering and drill-down analysis. The price elasticity simulator features an interactive interface where users can adjust pricing parameters via drag-and-drop operations, with real-time demand forecasts and visualized price elasticity across product categories. The dynamic pricing calculator incorporates optimization algorithms that automatically generate optimal pricing recommendations based on real-time market data and competitive landscape, while predicting their potential impact.

Deployed with Docker containerization technology, the system provides environment isolation and rapid migration capabilities, with annual service costs kept at the ten-thousand-yuan level—only about one-tenth of commercial systems [2]. It supports both private deployment and SaaS service models, flexibly accommodating enterprise needs of varying scales and IT infrastructure.

4. Experimental Results and Analysis

4.1. Model Performance Evaluation

The model's performance was comprehensively evaluated on the test set and compared with traditional pricing models. Key metrics included Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2, as shown in Table 1.

Table 1 Comparison of performance of different models

model	MAE	RMSE	R^2
linear regression	0.123	0.168	0.72
random forest	0.076	0.102	0.86
panel fixed effect model	0.089	0.115	0.84
XGBoost model	0.048	0.067	0.92
This study employs a mixed model.	0.043	0.059	0.94

Experimental results demonstrate that the hybrid model developed in this study outperforms traditional models across all metrics. The mean absolute error (MAE) remains consistently below 5%, while the R^2 coefficient of determination exceeds 0.92, achieving internationally recognized excellence. Compared to the standalone XGBoost model, the hybrid approach enhances prediction accuracy by incorporating economic constraints from panel fixed effects models. For model evaluation, this study employs three core metrics—MAE, RMSE, and R^2—to comprehensively and precisely assess the model's performance.

4.2. Analysis of Feature Importance

The SHAP value analysis quantifies the contribution of each feature to demand forecasting, as shown in Figure 3. Price elasticity coefficient, advertising intensity, and market concentration are the three core factors influencing demand, accounting for 32%,21%, and 15% of the explanatory power respectively. Regional market characteristics and seasonal factors also demonstrate significant impacts, indicating that pricing strategies should consider regional variations and dynamic adjustments over time.

Figure 3 SHAP value analysis of feature importance

(Chart showing each feature's contribution to demand forecasting)

Notably, distinct product categories exhibit markedly different feature importance: FMCG products demonstrate greater sensitivity to price fluctuations, whereas durable goods are more influenced by brand influence and channel coverage. This finding provides critical evidence for differentiated pricing strategies. Through SHAP value-based interpretive analysis, the study further quantified the contribution of each feature variable to demand forecasting outcomes, offering data-driven support for subsequent strategy formulation.

4.3. Effects of the Pilot Application in Enterprises

Select three manufacturing enterprises in the Yangtze River Delta region (with annual revenue ranging from 20 million to 100 million yuan) for field testing to validate the model's performance in real-world commercial environments. The pilot enterprises, representing the equipment components, packaging materials, and consumer electronics industries respectively, serve as typical cases.

After three months of application of the pricing system developed in this study, the comparative analysis of the enterprise operation data is shown in Table 2.

Table 2. Effects of enterprise pilot applications

metric	Before the pilot	After the pilot	change rate
efficiency of pricing decision	8 hours / time	2 hours / time	+75%
Demand forecasting accuracy	76%	91%	+19.7%
inventory turnover	6.2 times per year	7.8 times per year	+25.8%
gross profit rate	18.5%	21.3%	+15.1%
Customer satisfaction	82 points	89 points	+8.5%

The pilot results demonstrate that the system enhances corporate pricing decision-making efficiency by over 60%, improves demand forecasting accuracy by nearly 20 percentage points, and increases average gross margin by approximately 15%, validating the model's practical value [3]. Furthermore, the significant improvement in inventory turnover rate indicates that dynamic pricing strategies effectively mitigate inventory overstock issues and boost overall operational efficiency.

To evaluate the robustness of the model in complex market environment, a series of adversarial tests are designed. The error rate of the model's demand prediction remains stable within 5.2% even after adding 20% random noise to the input data, demonstrating its strong anti-interference capability and stability.

Furthermore, simulations of sudden market fluctuations demonstrate that the proposed model can swiftly respond to external shocks and dynamically adjust pricing strategies, reducing profit losses by 12%–18%

compared to traditional static pricing models [3], highlighting its distinct advantages in uncertain market environments.

5. Discussion and Conclusion

The innovation of this study is mainly reflected in the following three aspects:

Methodologically, we constructed a hybrid modeling framework integrating econometrics and machine learning. This framework not only preserves the theoretical explanatory power of panel fixed-effects models but also fully leverages the advantages of the XGBoost algorithm in handling high-dimensional nonlinear relationships, thereby achieving dual improvements in both causal inference and prediction accuracy.

Technically, a multi-source data fusion analysis framework was developed to integrate internal operational data with external market intelligence through spatiotemporal calibration. This framework constructs a comprehensive decision-making matrix spanning 32 dimensions, capable of explaining 82% of demand fluctuations—significantly higher than the 56% explanatory power of traditional models [3].

The application adopts a dual-driven model of "data empowerment + tool simplification". Through a low-code lightweight deployment solution, complex algorithm models are transformed into user-friendly visual decision tools, significantly lowering the barrier for SMEs to adopt advanced pricing technologies. Enterprise users report an average onboarding time of just 18 minutes after system implementation, demonstrating its intuitive design and ease of use. This approach dramatically reduces technical barriers, enabling non-technical personnel to quickly adapt and operate the system efficiently [3].

The research findings provide valuable practical guidance for small and medium-sized enterprises (SMEs) in formulating pricing strategies. Data-driven pricing methods can effectively enhance corporate profit-generating capabilities, particularly in volatile markets. By leveraging dynamic pricing mechanisms, businesses can swiftly respond to supply-demand fluctuations and adjust quotations in real-time to maximize profitability. When making pricing decisions, companies should comprehensively evaluate all influencing factors—not only analyzing internal cost structures and historical sales performance but also closely monitoring external competitive landscapes and regional market characteristics. This foundation enables the implementation of targeted differentiated pricing strategies to improve market responsiveness. For SMEs, adopting lightweight, low-investment intelligent solutions offers distinct advantages. These tools not only avoid the high implementation and maintenance costs of large-scale business intelligence systems but also facilitate cost-effective data-driven operations, allowing businesses to fully benefit from technological advancements.

This study has several limitations. First, the data primarily comes from the Business Management Skills Competition, which, while representative of the industry, still falls short of real-world corporate operational data in scope and depth. Second, the experimental sample is limited in size and predominantly concentrated in the Yangtze River Delta region, so the conclusions' applicability to other regions requires further validation with larger samples. Additionally, the current model inadequately captures how consumer behavior evolves over time and context, particularly in implementing highly personalized pricing strategies, which demands further refinement and expansion.

Future research could focus on the following directions: First, expanding data collection channels by integrating real-time consumer behavior data with social media dynamics to enhance the model's sensitivity to market trends and predictive capabilities. Second, applying game theory principles to develop dynamic competition and response models for multiple participants, thereby improving the robustness and adaptability of pricing strategies in real-world market competition. Additionally, customized solutions tailored to industry-specific characteristics should be prioritized, with model structures and parameter settings adjusted to meet the diverse needs of enterprises in retail, manufacturing, services, and other sectors.

Against the backdrop of the deepening digital economy, intelligent pricing technology is poised to become a core tool for SMEs to enhance market competitiveness and adapt to complex business environments. This study, through the interactive mechanism of "promoting research through competitions and supporting enterprises through research," not only explores new pathways for transforming management discipline competition achievements into practical productivity, but also provides replicable and scalable practical cases for university innovation and entrepreneurship education.

References

[1] China Academy of Information and Communications Technology. China Digital Economy Development White Paper (2023)[R]. 2023.

[2] Zhu Wang Tianchen, Li Haoyan, Xu Yuanlong, et al. Research Proposal on Product Pricing Strategies Based on Multi-dimensional Enterprise Data [Z]. Beijing Wuzi University, 2025.

[3] Zhu Wang Tianchen. Research on Product Pricing Strategies Based on Multi-dimensional Enterprise Data [Z]. Beijing Wuzi University Undergraduate Innovation and Entrepreneurship Training Program Application Form, 2025.

[4] Talluri K T, Van Ryzin G J. The theory and practice of revenue management[M]. Springer Science & Business Media, 2006.

[5] Wang Y, Chen X, Zhang J. Dynamic pricing for perishable products with demand learning[J]. International Journal of Production Economics, 2022, 245: 108463.

[6] Zhang L, Wang X, Liu H. Demand forecasting and dynamic pricing for retail products using machine learning techniques[J]. Journal of Retailing and Consumer Services, 2021, 59: 102472. This article extensively explores the application of machine learning in forecasting demand and implementing dynamic pricing strategies for retail products, offering significant insights into how advanced computational methods can enhance retail operations and improve profitability.

[7] Chen T, Guestrin C. XGBoost: A scalable tree boosting system[C]//Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016: 785-794. In this widely cited conference paper, the authors introduce XGBoost, a highly efficient and scalable tree boosting algorithm that has become a cornerstone in machine learning for its performance and versatility in various predictive modeling tasks.

[8] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//Advances in neural information processing systems. 2017: 4765-4774. This influential work presents a unified framework for interpreting complex machine learning model predictions, offering methods such as SHAP values to improve transparency and trust in automated decision-making systems across different domains.

[9] Li J, Wang Z, Zhang H. Research on intelligent pricing system for small and medium-sized enterprises based on cloud computing[J]. Journal of Business Research, 2023, 156: 122038. The study investigates the development and implementation of an intelligent pricing system leveraging cloud computing technologies, specifically designed to assist small and medium-sized enterprises in optimizing their pricing strategies through data-driven insights and scalable infrastructure.

[10] VanderPlas J. Python data science handbook: Essential tools for working with data[M]. "O'Reilly Media, Inc.", 2016. This comprehensive handbook serves as an essential resource for data scientists, providing detailed guidance on using Python and its key libraries for effective data manipulation, analysis, and visualization in real-world applications.

[11] Barnett V, Lewis T. Outliers in statistical data[M]. John Wiley & Sons, 2013. A foundational text in statistics, this book thoroughly examines the identification, analysis, and treatment of outliers in datasets, discussing their impact on statistical conclusions and offering practical methodologies for handling anomalous observations.

[12] Wooldridge J M. Introductory econometrics: A modern approach[M]. Cengage Learning, 2015. Widely used in academic and professional settings, this textbook provides a modern and accessible introduction to econometric methods, blending theoretical foundations with practical examples to equip readers with the skills needed for empirical economic analysis.

[13] Hastie T, Tibshirani R, Friedman J H. The elements of statistical learning: data mining, inference, and prediction[M]. Springer Science & Business Media, 2009. Widely regarded as a foundational resource, this book offers an in-depth exploration of statistical methods essential for data mining and predictive modeling, covering both theoretical underpinnings and practical implementation techniques.

[14] Nash J F. Non-cooperative games[J]. Annals of mathematics, 1951: 286-295. In this seminal paper, Nash introduces the concept of equilibrium in non-cooperative game theory, laying the groundwork for modern economic analysis of strategic interactions and earning a lasting place in the study of mathematical economics.

[15] Guyon I, Elisseeff A. An introduction to variable and feature selection[J]. Journal of machine learning research, 2003, 3(Mar): 1157-

1182. This article systematically reviews methods and challenges associated with selecting relevant variables and features in machine learning, providing valuable insights into improving model performance and interpretability.

[16] VanderPlas J, Granger B. Streamlit: A faster way to build and share data apps[J]. Journal of Open Source Software, 2020, 5(52): 2170. Presenting Streamlit as an innovative open-source framework, this paper highlights its utility in rapidly developing interactive web applications for data science, facilitating easier collaboration and dissemination of analytical results.