Applied Science and Innovative Research
ISSN 2474-4972 (Print) ISSN 2474-4980 (Online)
Vol. 10, No. 1, 2026
www.scholink.org/ojs/index.php/asir

Original Paper

Application of Data Mining Techniques in Customer Segmentation for Commercial Banks

Wankun Shang
Nke Safety Apparel, 16616 Chandler RD East Lansing, MI 48823, United States

Received: December 22, 2025 Accepted: January 19, 2026 Online Published: January 22, 2026
doi:10.22158/asir.v10n1p45 URL: http://dx.doi.org/10.22158/asir.v10n1p45

Abstract

Commercial banks are able to use Data Mining Technology (DMT) to improve their capability to segment customers or gain more thorough insight into their value. By developing DMT research methodologies that incorporate a variety of techniques including clustering methods, classification algorithms and other techniques to analyze multiple dimensions (e.g., customer transaction history and consumer behavior), conclusions can be drawn from the resultant output. Results of the research show that DMT can uncover previously unrecognized patterns of demand among customers, allowing for the design of individualized products and providing banks with a scientific framework for determining the optimal allocation of marketing resources.

Keywords

Data mining, Commercial banks, Customer segmentation, Cluster analysis, Precision marketing

1. Introduction

Commercial banks are confronting two practical realities of diversified customer demands and intensified competition in the marketplace. As a result, traditional segmentation approaches are becoming increasingly difficult to adapt to dynamic changing environments. As a solution, the use of data mining technology allows for the analysis of large amounts of customer data to identify patterns inherent to the data. The use of data mining technology provides banks with a new way to approach segmentation work. In this paper, the author discusses the specific application of this technology within bank customer management by providing a framework through which banks can utilize data-driven decision-making to improve their focus on serving customers' needs, and therefore increase bank customer retention and ultimately their business growth potential.

2. Theoretical Foundations of Data Mining and Customer Segmentation

2.1 Overview of Data Mining Techniques and Common Algorithms

Data mining technology aims to automatically discover hidden patterns and trends in complex business datasets through the use of algorithms to uncover knowledge. A Clustering Algorithm may be used to separate clients into various categories based on the attributes of like-ness, for example, multiple customers having different transaction volumes and types of products they carry. Classification Algorithms use historical behaviour, and therefore allows for the construction of prediction-based models, predicting which category, or trend will fall under new clients. Algorithms process withdrawal and deposit activity and the consultation activity of financial institutions to merge the many disparate points into a consolidated picture of groups that share similar characteristics. Building models with the use of Algorithms requires the careful balancing of two opposing forces - complexity and interpretability - so that the resulting model can support the decisions and actions of the business user (Fu Guojin, 2007).

2.2 Core Concepts and Value of Customer Segmentation in Commercial Banks

Customer segmentation is reflected in the systematic classification implemented by banks based on the differences in customers' intrinsic characteristics, with the aim of identifying groups with similar financial behaviors and needs. Banks develop differentiated service strategies based on clear customer classifications, such as optimizing transfer processes for customer groups that value liquidity. Without this structured classification method, the service push and product design of banks are prone to deviate from the actual expectations of various customers. A detailed customer profile enables banks to optimize internal operational priorities under limited resources, focusing service efforts on responding to the core demands of key customer groups.

2.3 Theoretical Framework for Applying Data Mining to Customer Segmentation

Applying data mining techniques to customer segmentation follows a closed-loop logical framework from data to decision-making. This framework starts with the collection and cleaning of multi-source customer data, aiming to build a reliable data foundation for analysis. The data analysis stage identifies customer groups with similar characteristics from the processed data based on algorithms. The business department designs targeted products and communication strategies based on these clustering results, and applies them to actual marketing and service scenarios. The customer feedback and market performance data generated by strategy execution will be fed back into the data collection end to drive continuous optimization of the model. The complete process can be summarized as the theoretical framework shown in Figure 1.

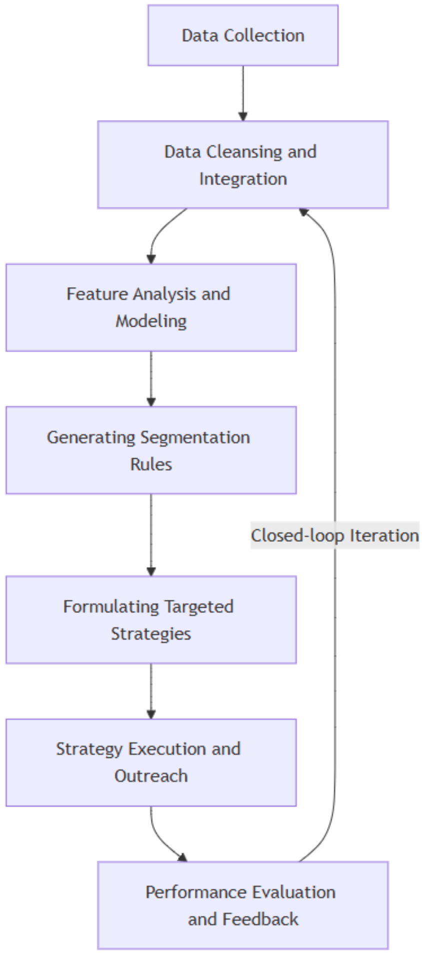

Figure 1. Theoretical Framework for Data Mining Applied to Customer Segmentation

3. Current Status and Challenges of Data Mining Applications in Commercial Bank Customer Segmentation

3.1 Incomplete Data Infrastructure and Data Quality Challenges

The primary obstacle faced by commercial banks in integrating customer information is that data is scattered across multiple independent systems, and the data caliber of different business lines is not consistent, making it difficult to form a unified view. The basic information, transaction records, and behavioral preferences of customers are stored in different databases, and the differences in standard formats make cross channel data fusion technically difficult. Due to human negligence in the data entry process or legacy issues with system integration, some customer information may be missing or incorrect, such as invalid contact information or conflicting transaction records. Low quality raw data directly affects the accuracy of subsequent modeling and analysis, and customer groups based on incorrect information segmentation often deviate from real needs. The dynamic changes in the banking business environment require real-time data updates, but some manual or batch processing update mechanisms have delays, making it difficult for analysis results to capture the latest market trends.

3.2 Superficial Technology Application and Insufficient Model Accuracy

At present, some commercial banks need to strengthen the application depth of data mining technology. Their customer segmentation models often focus on easily obtainable static indicators such as asset size and transaction frequency, while the integration and analysis of dynamic behavioral characteristics such as customer lifecycle stage and multi-channel interaction preferences are not sufficient. This relatively superficial variable selection makes it difficult for the model to accurately depict the internal demand structure and future evolution trend of the customer segmentation output, and the model parameters gradually lag behind market changes due to the lack of a continuous iterative optimization mechanism. The consequence is that there is sometimes a distance between the analytical conclusions and the intuitive cognition formed by front-line customer managers based on daily services, which affects the confidence of business departments in data-driven decision-making. The lack of precision will eventually manifest in marketing practice, as the unified product information pushed based on existing segmentation results often fails to effectively reach the deep needs of different customer groups, thereby restricting the actual effectiveness of customer value mining and service experience improvement (Liu Hongshen & Xu Yabin, 2008).

3.3 Shallow Integration of Segmentation Results with Application Scenarios

In the current banking system, the customer segmentation reports produced by analysis teams sometimes have a presentation format that is too biased towards statistical dimensions, making it difficult to directly guide the daily service actions of customer managers. There is a gap in understanding between the two types of teams in terms of information interpretation and application. When formulating specific marketing plans, the decision-making process of the business department may not have fully integrated the deep dependence on segmented labels, and traditional experience still holds a certain decision-making weight. The product innovation and annual marketing planning of banks usually follow established fixed cycles and approval processes, making it difficult to respond in a timely manner to the dynamic and subtle customer demand shifts revealed by segmented insights. This fusion gap prevents the potential service opportunities and customer value growth points identified by data work from being systematically transformed into perceptible product optimization and precise outreach strategies, thereby weakening the business growth potential and customer experience improvement that data analysis should bring.

3.4 Shortage of Specialized Talent and Lack of Process Mechanisms

The reserve of composite talents who are proficient in both data analysis and retail banking within banking institutions is relatively limited. There are often cognitive differences between business departments and technology teams in setting project goals and expected results, which directly lead to model construction deviating from the real business scenario requirements. In the process of advancing customer segmentation projects, there is a lack of standardized process specifications that run through data preparation, model development, strategy formulation, and frontline implementation. Departmental collaboration often relies on temporary communication rather than stable institutional design, which increases the uncertainty of project coordination costs and results. The lack of an effective closed-loop feedback mechanism makes it difficult to systematically evaluate and iterate the performance of the model in real business scenarios, and the practical experience of frontline personnel cannot be smoothly fed back to the technical team for optimization, resulting in the model gradually lagging behind market changes. On a deeper level, organizational structures often lack a dedicated position or team that can coordinate business goals and technology implementation, resulting in a less clear path for converting data insights into business value (Zhao Dan & Li Yadai, 2017).

3.5 Customer Privacy and Data Security Risks

Banking institutions continuously collect sensitive data such as customer identity information and transaction records in their daily operations, and the security measures for data storage and transmission are directly related to the customer trust foundation. In some business scenarios, there are blind spots in permission management for data calling and sharing processes, and internal employees or partners may have access to information beyond the necessary scope. The regulatory environment's requirements for personal information protection are constantly increasing, and the innovation speed of banking business sometimes exceeds the improvement process of compliance framework. The awareness of customers' right to know and choose their own data usage is increasing, and vague data usage terms may cause doubts and disputes. The negligence of security protection mechanisms may not only directly lead to data leakage incidents, but also erode the long-term trust relationship established between banks and customers.

4. Optimized Application Strategies for Data Mining in Commercial Bank Customer Segmentation

4.1 Strengthening Data Foundations and Enhancing Data Governance

The primary task for banks to consolidate their data foundation is to integrate data sources scattered across dozens of independent systems such as core systems, credit management, and customer relationship management, and build a unified data platform covering the entire business to eliminate "data silos". Jiangnan Rural Commercial Bank has integrated over 660 business tables and more than 35000 fields at once through the construction of a unified regulatory marketplace, providing complete raw materials for subsequent analysis. After collecting data, banks need to establish unified data standards and governance rules throughout the bank. For example, Qilu Bank has released over 9000 basic data standards and ensured that the failure rate of the new system reaches over 93%. This fundamentally solves the problem of different departments having different data standards and opinions. Data quality needs to be ensured through a governance process that spans the entire lifecycle. Banks will establish a multi-layered verification system from input, processing to application. For example, Jiangnan Rural Commercial Bank uses automated rule verification to increase the regulatory indicator verification pass rate from 75% in the early stages to 100% in the end. After completing high-quality data integration and governance, its business value will quickly emerge. The practice of Su Shang Bank shows that an intelligent risk control system built on a solid data foundation can significantly shorten the approval time for small and micro enterprise loans from the traditional 5 days to 3 minutes, with an automated approval rate of 88%.

4.2 Selecting and Optimizing Data Mining Models and Methods

The selection and tuning of models should be closely aligned with specific business objectives, such as improving cross-selling success rates or identifying potential churn customers, rather than solely pursuing technical complexity. The analytics team must collaborate with retail banking and wealth management departments early in the project to define core business metrics for measuring segmentation effectiveness, ensuring model training aligns with actual needs. For preliminary exploration of customer segments, clustering algorithms can group individuals with similar financial behaviors and preferences based on natural characteristics, providing foundational profiles for subsequent precision services. When business goals involve predicting specific customer behaviors, classification algorithms like decision trees can build predictive models using historical data to help identify high-potential groups or early warning signals. The model development process must incorporate continuous validation and iteration mechanisms, regularly testing predictive stability with newly occurring business data and adjusting feature variables and parameter settings based on market feedback. Model interpretability is critical for business applications, requiring analysis reports to clearly present key characteristics and drivers of different customer groups, enabling frontline teams to understand and translate them into communication strategies. The final evaluation of model utility should return to real-world testing within business scenarios, observing whether marketing plans based on its recommendations truly trigger the desired customer responses and value enhancement (Guo-Jin F. U., 2007).

4.3 Deepening Closed-Loop Integration of Segmentation Outcomes and Precision Marketing

The core of closed-loop design lies in establishing a continuous cycle from insights to actions and then to feedback. Its starting point is transforming the granular rules output by models into a customer tagging system that frontline business personnel can understand and apply. This system should be business-readable, such as descriptions like "high-frequency trading - low risk preference" or "growth-stage family - significant education demand," directly corresponding to product and service solutions. Based on this tagging system, the marketing department must pre-design matching product recommendation combinations, communication scripts, and channel selection strategies for each granular category, embedding these strategies into the customer manager's operational platform. During the pilot validation phase, representative customer groups should be selected for small-scale marketing campaigns, closely tracking the actual changes in key metrics such as customer response rate, conversion rate, and satisfaction. Market feedback data needs to be systematically fed back to the data analysis team to evaluate the predictive accuracy of segmentation tags and the effectiveness of strategies, enabling adjustments to the next round of customer segmentation rules and marketing actions. This cyclical mechanism ensures that customer segmentation is no longer a one-time study but a dynamic core engine driving the continuous optimization of marketing resource allocation.

4.4 Cultivating Multidisciplinary Talent and Refining Application Workflows

Banking organizations should implement talent development as part of their human resources strategy and clearly define the competencies needed for prospective staff with business intellects and analytical foundations on which they create business strategies. The internal training program must have modularized course offerings that include topics related to data analysis, customer behavior and banking methodologies. These courses must be augmented with the opportunity to engage in practical, hands-on training based upon real historical data examples. Cross-functional teams made up of individuals from the organization who have either managed or been on either side of the business and technology functions should work collaboratively to share knowledge and experiences in the area of specific customer segmentation initiatives to close the communication gap between the two types of professionals. Banks also need to develop standardized processes and deliverables to help define the roles, responsibilities, deliverables and key decision points throughout the life cycle of a data mining project, starting from the identification of the business need to the deployment of a solution model. Additionally, the process of executing a data mining project must incorporate methods for performing project assessments and regular reviews of project results, routinely gathering and evaluating the achievement and bottlenecks encountered in the execution of a data mining project so that the organization can create a comprehensive repository of organizational best practices based upon individual and collective learning achieved from project engagement. Continuous investment at the organizational level can transform talent and process development into reusable institutional assets, supporting banks to maintain agile response and innovation capabilities in the face of constantly changing customer demands (Shi-Chao L. I., Fan-Jin M., & Zhen W., 2009).

4.5 Establishing Customer Privacy Protection and Security Management Systems

To maintain consumer privacy, commercial banks should create data privacy protective policies from the time an individual enters the bank as well as define clearly the parameters of why the bank collects customer information. In the Bank, criminal history information is classified by level of sensitivity and each classification has a different level of protection through the control of access to that information and keeping a complete log of who accessed that information. At the Technical Level, banks should use encryption and anonymization of data when data is processed for internal analysis and testing and banks should only use processed sample data from the bank when conducting internal analyses and tests. When it comes to Employee Training in Regulatory Compliance, it is important for banks to have training that focuses on being compliant by combining regulatory compliance requirements with actual business practices, and enhancing employee awareness of protecting consumer privacy in their everyday employment. Banks should create and conduct a pre-privacy impact assessment as part of the product development process to ensure that all innovative methods and new products will comply with the laws and regulations. Lastly, banks should conduct routine independent audits of the effectiveness of their information security plans and quickly fix any weaknesses or vulnerabilities that may be detected. Providing customers with ways to verify the accuracy of their own personal information is an important way to show customers the transparency and trustworthiness of how banks are using their information.

5. Application Effectiveness Evaluation and Future Outlook

5.1 Evaluation Dimensions and Methods for Customer Segmentation Outcomes

Evaluating the effectiveness of customer segmentation needs to be based on a quantifiable real business chain, with the core being tracking the entire process from data insights to final value growth. The practice of Sichuan Weiyuan Rural Commercial Bank provides specific examples for this, and its effectiveness can be systematically summarized into several key dimensions (as shown in the Table 1). The bank constructed customer profiles by integrating data and accurately identified 5749 "dormant" customers. This step directly optimized the visit targets of customer managers, significantly increasing their monthly average visit volume from 28 to 71, reflecting a significant improvement in operational efficiency. Subsequently, targeted services successfully activated 328 of them and directly brought in 52 million yuan in new credit usage, which verified the practical promotion of the strategy on business growth. Throughout the process, the precise conversion of hundreds of households from thousands of target customer groups also reflects a high level of accuracy in achieving the goals. This series of closed-loop data from "recognition" to "transformation" constitutes a solid basis for evaluating the success of segmentation strategies.

Table 1. Dimensions of Customer Segmentation Effectiveness Evaluation Based on the Weiyuan Rural Commercial Bank Case

Evaluation Dimensions	Evaluation Focus	Specific Case Performance / Metrics
Improvement in Operational Efficiency	Whether data tools have optimized customer engagement processes	The average monthly client visits per account manager increased from 28 to 71
Verification of Business Growth	Whether strategies directly translate into core business metrics	Precision activation of 328 clients generated RMB 52 million in new credit utilization
Accuracy in Goal Achievement	Whether identification and engagement of target audiences are accurate and effective	From a target list of 5,749 clients, 328 were successfully converted and activated

5.2 Key Implementation Challenges and Optimization Directions

The core challenges in implementing customer segmentation projects often stem from two levels: cross departmental collaboration and dynamic market adaptation. Differences in goals between business departments and technical teams can easily lead to analysis results being difficult to translate into frontline action guidelines. The rapid changes in the market environment may render static models based on historical data ineffective and unable to accurately capture the latest trends in customer demand. To address these challenges, banks need to optimize their organization and mechanisms by establishing a normalized cross functional collaboration team with clear responsibilities and decision-making authority. The team is responsible for regularly reviewing and updating customer segmentation rules based on the latest market feedback and business data, ensuring that the model evolves in sync with market reality. Simultaneously establish a fast feedback loop from strategy execution to effectiveness evaluation, enabling frontline experience to timely validate and revise the output assumptions of the analytical model. This dynamic optimization mechanism can help banks balance the flexibility and timeliness of business applications while maintaining the scientificity of segmentation strategies (Wan Dingsheng, Liu Cong, Liu Yuansheng, 2009).

5.3 Technological Trends and Future Application Prospects

Technological evolution will drive the continuous development of customer segmentation methods from static profiling to dynamic intelligence, with the core being the improvement of integration and interpretation capabilities for real-time data streams. In the future, banking institutions will be able to more effectively integrate real-time interactive information from customers on mobile devices, social media, and offline branches, in order to capture their real-time financial service needs and emotional changes. Machine learning models will be increasingly applied to predict potential demand turning points for customers at different stages of their lifecycle, providing forward-looking guidance for precise service interventions. The advancement of interpretable artificial intelligence technology helps business personnel understand the segmentation logic derived from complex models, enhancing decision-making confidence and collaborative efficiency. The maturity of privacy computing and other technologies enables banks to conduct more comprehensive joint modeling with compliance partners and improve customer profiles while ensuring data security. These trends collectively point to a more refined and agile customer insight system, whose application will not be limited to precision marketing, but will expand to core areas such as full journey service optimization, personalized risk pricing, and product innovation (Hu Hongyi, 2015).

6. Conclusion

Through data mining technology, commercial banks have enhanced their data-driven transformation and strategy optimization processes. The data-driven transformation process has created a more robust link from model development to business delivery. As technology develops, data mining technologies will continue to advance segmentation methodologies toward dynamic intelligence, thus banks must invest in establishing data governance systems and developing talent across disciplines to secure the success of their implementations. Implementing these technologies will help the banking industry move into its next stage of customer relationship management, thereby increasing the ability for customized and sophisticated relationships with customers.

References

Fu Guojin. (2007). Application of Data Mining Technology in Customer Segmentation of Commercial Banks. *Journal of Jinhua Vocational and Technical College*, *7*(4), 5.

Guo-Jin, F. U. (2007). Application of Data Mining Technology in Customer Segmentation of

Commercial Banks. *Journal of Jinhua College of Profession and Technology*, *2007*, 12.

Hu Hongyi. (2015). *Design of a Data Mining-Based Personal Customer Segmentation System for Commercial Banks*. Hunan University, 2015, 11.

Liu Hongshen, & Xu Yabin. (2008). Application of Data Mining Technology in Customer Segmentation. *Fujian Computer*, *24*(10), 2.

Shi-Chao, L. I., Fan-Jin, M., & Zhen, W. (2009). Study of Customer Segmentation Based on the Technology of Data Mining. *Computer Knowledge and Technology*, *2009*, 9-10.

Wan Dingsheng, Liu Cong, & Liu Yuansheng. (2009). Application of Data Mining Technology in Bank Customer Segmentation. *Microcomputer Applications*, *25*(2), 31-34.

Zhao Dan, & Li Yadai. (2017). Application of Data Mining Technology in Bank Customer Segmentation. *Market Weekly: Theory Edition*, *2017*(18), 1.