

How to Develop Targeted Marketing Strategies in the Retail Market? Combining Customer Profile Clustering Analysis

Shu Ren*

Abstract

This paper focuses on the banking retail market, using the retail market of Bank A as a case study. It established an amount-combined recency, frequency, and monetary (ARFM) model by combining the RFM model with the total number of transactions and collected customer data from Bank A based on this model. The clustering performance of three methods, including k-means, agglomerative nesting (AGNES), and density-based spatial clustering of applications with noise (DBSCAN) was compared on the customer data of Bank A. The data used covered the period from January to June 2024. After removing outliers and missing values, a final sample of 24,587 customers was analyzed. The effectiveness of each clustering algorithm was evaluated using the Silhouette coefficient (SC), Calinski-Harabasz Index (CHI), and Davies-Bouldin Index (DBI). The results showed that the k-means algorithm achieved the best clustering performance, with an SC value of 0.78, a CHI value of 3,964, and a DBI value of 0.66. It divided the customers of Bank A into four clusters: value customers, core customers, general customers, and easy-to-lose customers, for which corresponding profiles were established. Moreover, targeted marketing strategies for these clusters were developed. The results verify the applicability of customer profiling and cluster analysis in the financial field, demonstrating that they can be applied in the actual bank retail market.

Keywords

Cluster Analysis, Customer Profile, Marketing Strategy, Retail Market

1. Introduction

In the face of fierce market competition, various industries must utilize existing technical means to pursue development towards informatization and digitalization [1]. Current marketing method tend to prioritize the customer [2]. By processing and analyzing customer data, businesses can create detailed customer profiles that inform targeted marketing strategies, ultimately reducing costs and effectively reaching the right audience [3]. This method has been applied successfully in many fields [4]. Lin and Guan [5] proposed an electric power marketing user profile, designed the power payment user value model, and then implemented user profiling using the k-means algorithm. This method provides a strong foundation for electric power companies to assess user value and offer differentiated services. Liu et al. [6] analyzed the marketing mode of clothing enterprises and constructed an improved user profile model.

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received July 16, 2026; first revision October 20, 2025; second revision December 4, 2025; accepted December 16, 2025.

*Corresponding Author: Shu Ren (renshurs@outlook.com)

Anyang University, Anyang, Henan, China (renshurs@outlook.com)

Their tests indicated that this method could boost sales by more than 30%, offering a reference for the digital transformation of clothing enterprises. Tanuwijaya et al. [7] used the k-means algorithm to divide Netflix's potential customers into three clusters for personalized marketing. Through this approach, operators can improve marketing accuracy and obtain better marketing effects through this method. Sabuncu et al. [8] used the recency, frequency, and monetary (RFM) model to analyze data from 1,015 customers at a gas station in Turkey and subdivided them into five categories through cluster analysis. They found that truck drivers were the most valuable customers. Kumar et al. [9] utilized historical sales and customer data to analyze demand forecasting and customer segmentation in the medical field through statistical techniques and machine learning algorithms. Their analysis revealed that these two methods can complement each other to achieve resilience and flexibly respond to market demand fluctuations. Zervoudakis and Tsafarakis [10] adopted the flying fox optimization algorithm for customer segmentation. When applied to two different datasets, this method demonstrated excellent performance in identifying different customer groups, thereby facilitating the formulation of targeted marketing strategies. Maghsoudi et al. [11] proposed a method that combines sentiment analysis with customer segmentation, enabling manufacturers to effectively align product design with the ever-changing consumer demands. Mahendra et al. [12] analyzed the application of graph neural networks (GNNs) in customer segmentation within immersive virtual reality environments. The results showed that successful customer segmentation can be achieved with a GNN. Applying targeted marketing strategies in the financial sector can reduce marketing costs while both attracting new customers and retaining existing ones [13]. With the rapid development of Internet finance, banks face significant challenges in their retail markets. In order to further expand their customer base, banks must implement targeted marketing strategies. However, the classic RFM model widely used today fails to accurately capture the complex and diverse business relationships between customers and banks. For example, a customer with a high transaction frequency who only uses a savings account has different value and potential than a customer who engages in multiple services such as wealth management and credit. It is rather difficult for traditional models to make these distinctions. To address this issue, this paper developed customer profiles through cluster analysis of customer data from Bank A using the k-means algorithm and then put forward several suggestions on targeted marketing strategies based on these profiles. The research in this paper introduces the total number of transactions (amount) on the basis of the classic RFM model to construct a new amount-combined RFM (ARFM) model. The new approach aligns more closely with the complexities of banking services and diverse product offerings. Through a systematic comparison, the advantages of the proposed method on bank retail market customer data were demonstrated. This paper provides a reference for expanding the retail market of banks and enhancing the level of traditional bank retail business. It also provides theoretical support for the application of clustering algorithms and customer profiling in banking and other financial fields.

2. Cluster Analysis

2.1 k-Means

The k-means algorithm is a partition-based method [14]. Its principle is to divide a set of samples into k non-overlapping clusters. It has good performance in image processing [15], data classification [16],

and various other scenarios. The algorithm follows these steps:

- 1) randomly select k initial centroids,
- 2) assign each sample to the nearest cluster,
- 3) recalculate the center of mass for each cluster,
- 4) repeat these steps until the centers of mass no longer change.

As a benchmark algorithm based on partitional clustering, k -means is highly efficient and performs well on spherical clusters and clusters with uniform scale, making it suitable for meeting the requirements of customer profiling in the banking retail market.

2.2 Agglomerative Nesting

Agglomerative nesting (AGNES) is a hierarchical method [17] that treats each sample as a single cluster and merges clusters based on their distances [18]. The process is as follows:

- 1) treat each object as a cluster;
- 2) merge the two closest clusters;
- 3) repeat the process until k clusters are generated.

AGNES is a representative of hierarchical clustering. When analyzing customer profiles in the banking retail market, this method is advantageous for exploring the hierarchical structure of customer value without relying on the selection of the initial center point.

2.3 DBSCAN

Density-based spatial clustering of applications with noise (DBSCAN) [19] exhibits strong performance in processing discrete data [20]. One advantage is that it does not need to input the number of clusters to be divided [21]. The method determines the connectivity between samples based on sample density. There are two critical parameters in DBSCAN. The first is the neighborhood radius ϵ : for data x , the circle with x as the center and ϵ as the radius is called the ϵ -neighborhood of x . The second parameter is the minimum number of sample points (MinPts). For data x , if its ϵ neighborhood contains at least MinPts, the points can be said to be the core points. The process is as follows:

- 1) select an initial point at random and determine the number of points in its ϵ -neighborhood;
- 2) if the number of points is greater than or equal to MinPts, these points are recorded as core points, and the samples in this neighborhood are grouped into the same cluster;
- 3) repeat the process until all points have been accessed.

As a density-based algorithm, the advantage of DBSCAN lies in its ability to automatically identify noise points (abnormal customers). Applying it to the research on customer profiling in the banking retail market can help verify whether there are special customer segments that are difficult to capture in the data, such as inactive customers.

2.4 Cluster Effect Evaluation

1) Contour coefficient (SC)

Silhouette coefficient (SC) measures how similar a sample is to other samples within the cluster [22]. The formula is:

$$SC = \frac{b - a}{\max(a, b)} \quad (1)$$

where a is the average distance of the sample from other samples within the cluster and b is the average distance of the sample from all points within the adjacent cluster.

2) Calinski-Harabasz Index (CHI)

CHI is the ratio of the degree of dispersion between clusters to the degree of dispersion within clusters [23]. The formula is:

$$CHI(k) = \frac{Tr(O_k)/(k-1)}{Tr(I_k)/(n-k)} \quad (2)$$

where n refers to the number of samples, k is the number of clusters, Tr is the trace of the matrix, O_k is the intercluster covariance matrix, and I_k is the intra-cluster covariance matrix.

3) Davies-Bouldin Index (DBI)

DBI is the ratio of the distance within a sample class to the distance between classes [24]. The formula is:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} \left(\frac{\bar{c}_j + \bar{c}_i}{d(z_j, z_i)} \right) \quad (3)$$

where z_i and z_j are the center of mass of clusters i and j , $\bar{c}_i$ and $\bar{c}_j$ are the average distance of all samples in clusters i and j from the center of mass, and d is the Euclidean distance.

The SC measures the classification rationality of a single sample point. The CHI focuses on the macroscopic variance ratio. The DBI measures clustering quality based on the worst-case pairwise similarity between clusters, with smaller values indicating better performance. Each of these three indicators has a distinct focus. Therefore, this paper integrates all three to comprehensively analyze the clustering effects of different algorithms.

3. Retail Market of Bank A and Customer Profile

3.1 Retail Market of Bank A

Bank A is a commercial bank branch in Henan Province and has made significant investments in the retail market. Its retail products mainly include:

- 1) wealth: deposits, financial management services, funds, gold, etc.;
- 2) credit: mortgages, operating loans, credit cards, etc.

In terms of customer management, Bank A currently carries out hierarchical management according to differences in the value of existing customers and allocates financial managers according to customer assets. However, this division method is somewhat simplistic, and the marketing strategy predominantly relies on traditional methods, which yield low success rates. In the face of increasingly fierce market competition, Bank A's retail sector is under significant pressure, leading to customer attrition. Consequently, there is an urgent need to develop targeted marketing strategies to enhance the competitiveness of its retail market.

3.2 Processing of Customer Data of Bank A

In terms of customer data processing, the RFM model has been extensively used [25], and it contains three indicators:

- R (Recency): the time since the customer’s most recent purchase,
- F (Frequency): how often the customer makes purchases,
- M (Monetary): the total amount the customer has spent recently.

The RFM model evaluates customer value [26] through these three indicators, enabling tasks such as customer segmentation [27] and customer profiling [28]. It has been applied to customer data processing at Bank A. Moreover, it can be combined with the total number of transactions (amount) based on the specific situation of the banking retail market to form the ARFM model.

The inclusion of “Amount” is justified by the unique characteristics of the banking retail business.

- 1) Comprehensive quantitative assessment: In the RFM model, the single “F” alone may not adequately distinguish between customers who make numerous small transactions and those who engage in fewer, higher-value transactions. The total number of customer transaction activities (amount) can more directly measure a customer’s contribution level.
- 2) Precise business orientation: The marketing strategies of bank retail businesses, such as those related to wealth management and credit, are closely related to a customer’s asset scale. Introducing the amount can make the customer segmentation results directly serve differentiated product recommendation and interest rate pricing strategies, offering clear business directivity.

For this study, customer data was drawn from the transaction records of Bank A. The indicators were defined as follows:

- A: total number of transactions a customer has made in the recent 6 months,
- R: time elapsed since the most recent transaction,
- F: transaction frequency in the recent 6 months,
- M: the transaction amount in the recent 6 months.

Transaction data from Bank A covering January 2024 to June 2024 was selected. Records lacking key fields such as transaction amount and frequency were deleted. Records falling below the 1st percentile or above the 99th percentile were defined as outliers and deleted. After processing, data from 24,587 customers was obtained. Part of the data is shown in Table 1.

Table 1. Raw customer data of Bank A

	Amount (n)	Recency (day)	Frequency (time)	Monetary (yuan)
125145 **	5	2.685	18.652	1,985,246
412586 **	6	1.854	19.365	2,012,548
412536 **	1	9.332	3.215	1,125,415
123522 **	6	1.554	17.445	1,865,235
522367 **	7	5.685	10.116	3,985,412
965236 **	1	8.549	1.225	1,025,412
411252 **	2	7.657	2.374	1,136,258
451285 **	1	9.465	2.263	1,202,541
854211 **	8	6.654	6.258	3,215,216
854236 **	2	8.665	3.245	2,122,147

Due to the different units of the four dimensions, the customer data in Table 1 was standardized:

$$x_{new} = \frac{x - \min}{\max - \min} \quad (4)$$

where *max* and *min* refer to the maximum and minimum values of original data *x*.

3.3 A Cluster Analysis of Customer Data of Bank A

The elbow method was used to determine the *k* value for the *k*-means algorithm. The cases where *k* ranged from 2 to 40 were traversed and calculated using a “for” loop, and the result showed that the optimal *k* value was 6. The number of clusters for AGNES also used the *k* value determined by the *k*-means algorithm, namely *k* = 6. In the DBSCAN algorithm, the radius of neighborhood ε ranged from 0.001 to 1, with a step size of 0.05, and the MinPts ranged from 2 to 15. All parameter combinations were traversed through a “for” loop. Finally, $\varepsilon = 0.95$, MinPts = 2. The comparison of the clustering performance between different algorithms for the customer data of Bank A is shown in Table 2.

Table 2. Clustering performance analysis

	SC	CHI	DBI
k-means	0.78	3964	0.66
AGNES	0.75	3845	1.68
DBSCAN	0.56	656	2.77

As shown in Table 2, the SC value for the *k*-means algorithm was closest to 1, and it also had the highest CHI value and the lowest DBI value, which showed the best clustering performance on the customer data of Bank A. Therefore, subsequent analysis was conducted based on the clustering results from the *k*-means algorithm. The *k*-means algorithm divided the customers of Bank A into four clusters.

Table 3. Customer clustering results of Bank A

	Number of customers	A	R	F	M
Cluster 1	11,145	6	2.779	15.625	2,015,697
Cluster 2	4,568	3	5.667	10.897	3,521,484
Cluster 3	5,307	2	8.942	3.451	1,252,364
Cluster 4	3,567	1	10.897	1.456	1,021,546

From Table 3, it can be found that Cluster 1 had the largest number of customers, accounting for 45.33%. This segment of customers had a high transaction volume, with an average of six transactions, indicating extensive use of various Bank A products, including deposits, wealth management, and loans. Moreover, they had a short average interval since their most recent transaction, at 2.779 days. They also had a high average transaction frequency over the recent 6 months, 15.625 times, suggesting frequent interactions with Bank A. The total transaction amount for this cluster was also high, reaching 2,015,697 yuan. Based on these results, this customer group is a significant source of stable income for Bank A. They not only make a relatively high current contribution but also have a relatively stable lifetime value.

The number of customers in Cluster 2 was relatively small, accounting for 18.58%. Customers in this cluster had a small number of transactions, with an average of 3, and a long interval since their most

recent transaction, with an average of 5.667 days. Therefore, they were not frequent transaction customers. Their transaction frequency within 6 months was at a moderate level, averaging 10.897 times, but their transaction amount was the highest among all clusters, reaching 3,521,484 yuan. These results suggest that their banking business may mainly focus on large-amount wealth management and asset custody, indicating that they are likely high-net-worth individuals.

Cluster 3 comprised 21.58% of the customers. This group had an average transaction volume of 2 and an average interval of 8.942 days since their most recent transaction. Their transaction frequency was low, at only 3.451 times, and their average transaction amount was lower than that of the first two clusters, at 1,252,364 yuan. These customers may only use basic savings accounts, show little interest in wealth management products, and do not maintain a close relationship with Bank A. As a result, these customers have a risk of churn.

Cluster 4 had a smaller proportion of customers, at 14.51%. This group had the lowest transaction activity, averaging only 1 transaction and the longest interval since their most recent transaction, exceeding ten days. Their transaction frequency was also the lowest, at just 1.46 times, and their transaction amount was the smallest, totaling 1,021,546 yuan. These customers may have opened accounts but rarely used them, or they may be long-standing customers who are on the verge of churning. The cost of maintaining these customers may exceed the benefits they generate.

In order to further validate the business credibility of the clustering results, they were cross-compared with the business labels that are known within Bank A but were not used in the modeling. The results showed that:

- 1) The average financial assets per capita of customers in Cluster 1 far exceeded the average level of the whole bank, and the proportion of customers holding high-net-worth or private banking customer labels exceeded 80%; in terms of product holdings, the cross-holding rate of these customers simultaneously holding financial products, funds, and insurance products was extremely high;
- 2) The monthly asset activity rate of customers in Cluster 2 remained stable at above 95%, and their response rate to newly launched bank financial products was also the highest, exceeding 30%. Thus, they were the most active customers and had the greatest growth potential in the bank's ecosystem;
- 3) The average account balance of customers in Cluster 3 was relatively low, and for more than 50% of these customers, their main accounts were either pure settlement accounts or low-balance savings accounts; their product holdings were also very single, concentrated in basic savings and payment services;
- 4) In the past 6 months, for customers in Cluster 4, more than 60% exhibited behaviors such as a significant decline in account balance or termination of core products, and the proportion of customers marked as inactive by the bank's system exceeded 50%, which was consistent with the clustering results.

Through the above comprehensive systematic cross-validation using key business indicators, it can be concluded that the designed method is not only statistically valid from a data perspective but also has practical significance and interpretability in real banking business scenarios.

3.4 Customer Profile and Targeted Marketing Strategy

According to the clustering results in Table 3, the customers of Bank A were divided into four

categories. The customer profiles and targeted marketing strategies are as follows.

1) Value customers

This group is characterized by a wide variety of transactions, short intervals between transactions, high activity frequency, and medium transaction amounts. These customers demonstrate high engagement with the bank, holding multiple products and showing strong loyalty. As a result, they have created significant value for Bank A's retail market.

The number of these customers is large and stable, making them an important segment of the bank's retail operations. Therefore, it is necessary to focus on maintaining these customers. They can be encouraged to choose Bank A for wealth management and payment settlement. At the product level, advanced wealth management tools such as intelligent fixed-amount investment and fund portfolios can be offered to them. The bank could also provide them with complimentary personal asset allocation reports and high-limit credit cards with numerous benefits. At the service level, more very important person (VIP) activities can be provided for them. A points redemption platform can be established where points can be redeemed for airport VIP lounges and high-end hotel discounts to enhance the value of points. At the same time, the bank can create a fast track for valuable customers within the mobile banking app to reduce their waiting time and improve satisfaction.

2) Core customers

The characteristics of these customers are a limited variety of transactions, moderate trading intervals and frequencies, and high trading amounts. Although this type of customer is not as active as the previous group and holds fewer products, their transaction amounts are large, and they can create ultra-high value for the bank's retail market. Therefore, they can be considered core customers of Bank A.

These customers hold great value for Bank A and is very important. To better serve them, the bank should improve closeness of contact with customers by providing one-to-one service through senior account managers and offering customers with financial services with good privacy, high integration, and convenience. The bank can establish a dedicated channel for high-value services such as large-scale fund transfers and loan approvals. In terms of product offerings, exclusive wealth management products with high yields and low minimum investment thresholds should be created specifically for these core customers. When a significant inflow of funds occurs, the bank can proactively offer preferential subscription opportunities for private equity and popular trust products. In addition, it is necessary to conduct robust risk monitoring for these customers. When a large outflow of funds is detected, the customer manager should follow up to understand the reasons and develop retention strategies.

3) General customers

The characteristics of these customers are few types of transactions, long trading intervals, low frequency, and large transaction amounts. As a result, their activity, loyalty, and value can be considered average, bringing little value to the retail market of Bank A; they contribute only modestly to Bank A's retail market.

The value of these customers is average, but their numbers are significant. If similar management strategies as those used for the first two types of customers are adopted, there is a risk of incurring high marketing costs without corresponding returns. Therefore, the bank should consider low-cost automated marketing methods for these customers. For instance, different services can be offered based on their

behaviors. A growth ladder could be integrated into the mobile banking app, rewarding customers with 1,000 bonus points for purchasing their first financial product. Completing specific tasks would enable customers to earn corresponding benefits.

4) Easy-to-lose customers

These customers show few types of transaction activity, long transaction intervals, infrequent purchases, and low transaction amounts. Therefore, their activity level and value are low. They may fall into the categories of occasional or inactive customers within Bank A's retail market, displaying low loyalty and a high risk of attrition.

Given their limited value, marketing to these customers poses a significant challenge. Therefore, Bank A can reduce maintenance and marketing efforts for these customers, transitioning all service channels to the self-service mode and limit active investment in labor costs, thereby concentrating resources on customers with greater potential return value.

4. Conclusion

Based on the ARFM model, this paper divided the customers of Bank A into four clusters, value, core, general, and easy-to-lose customers, analyzed their customer profiles, and proposed targeted marketing suggestions for each customer category. This paper provides theoretical support for the development of the retail market of Bank A. However, the research in this paper also has some limitations. For example, the analysis relies solely on existing clustering methods, lacking exploration of advanced and hybrid techniques such as the Gaussian mixture model, k-means++, and ordering points to identify the clustering structure. Moreover, data such as customer age, occupation, income, and product holding status were not included. The marketing strategies proposed in this paper also lack practical verification. Therefore, in future work, more customer data will be incorporated to establish a more comprehensive customer tagging system, thereby making the marketing strategies more targeted. Moreover, more comparisons and combinations of clustering algorithms will be studied to capture more nuanced customer models. Then, the strategies proposed in this paper will be transformed into specific marketing activity plans, and the empirical cost-benefit of these strategies will be analyzed to scientifically evaluate their return on investment.

Conflict of Interest

The author declares no competing interests.

Funding

None.

References

- [1] O. Gumus, M. B. Agirkaya, F. Ekici, and M. Karadeniz, "Law and economics in Fintech with SWOT analysis: a theoretical approach," *Turkish Studies - Economics, Finance, Politics*, vol. 18, no. 3, pp. 885-906, 2023. <https://doi.org/10.7827/TurkishStudies.69246>
- [2] Y. Zintso, I. Fedorishina, H. Zaiachkovska, O. Kovalchuk, and Z. Tyagunova, "Analysis of current trends in the use of digital marketing for the successful promotion of goods and services in Ukraine," *Financial & Credit Activity: Problems of Theory and Practice*, vol. 3, no. 50, pp. 174-184, 2023. <https://doi.org/10.55643/fcaptp.3.50.2023.4080>
- [3] B. Ganguly and A. Ambhaikar, "Online e-commerce customer point of view and sectional interest analysis for custom user experiences," in *Advances in Data and Information Sciences*. Singapore: Springer, 2022, pp. 393-403. https://doi.org/10.1007/978-981-16-5689-7_35
- [4] R. Benkhelifa and N. Bouhyaoui, "Deep approach based on user's profile analysis for capturing user's interests," in *Artificial Intelligence and Its Applications*. Cham, Switzerland: Springer, 2022, pp. 177-186. https://doi.org/10.1007/978-3-030-96311-8_17
- [5] L. Lin and S. Guan, "Power marketing big data precision user profile," in *Proceedings of 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*, Jishou, China, 2019, pp. 117-120. <https://doi.org/10.1109/ICVRIS.2019.00038>
- [6] J. Liu, Z. Liu, and Y. Xu, "Research on the Innovation of marketing management mode under the digital transformation of textile and garment enterprises," *Journal of Information & Knowledge Management*, vol. 23, no. 1, article no. 2450003, 2024. <https://doi.org/10.1142/S0219649224500035>
- [7] S. Tanuwijaya, A. Alamsyah, and M. Ariyanti, "Mobile customer behaviour predictive analysis for targeting Netflix potential customer," in *Proceedings of 2021 9th International Conference on Information and Communication Technology (ICoICT)*, Yogyakarta, Indonesia, 2021, pp. 348-352. <https://doi.org/10.1109/ICoICT52021.2021.9527487>
- [8] B. Sabuncu, E. Turkan, and H. Polat, "Customer segmentation and profiling with RFM analysis," *Turkish Journal of Marketing*, vol. 5, no. 1, pp. 22-36, 2020. <https://doi.org/10.30685/TUJOM.V5I1.84>
- [9] K. G. Kumar, S. Harini, and R. Perumalraja, "Optimizing supply chain management in the medical field through demand forecasting and customer segmentation," in *Proceedings of 2024 13th International Conference on System Modeling & Advancement in Research Trends (SMART)*, Moradabad, India, 2024, pp. 424-430. <https://doi.org/10.1109/SMART63812.2024.10882528>
- [10] K. Zervoudakis and S. Tsafarakis, "Customer segmentation using flying fox optimization algorithm," *Journal of Combinatorial Optimization*, vol. 49, article no. 5, 2025. <https://doi.org/10.1007/s10878-024-01243-6>
- [11] M. Maghsoudi, N. Mohammadi, and M. Bakhtiari, "A novel approach to customer segmentation for product development on social media data: integrating aspect-based sentiment analysis and text mining," *Knowledge-Based Systems*, vol. 328, article no. 114269, 2025. <https://doi.org/10.1016/j.knosys.2025.114269>
- [12] K. B. Mahendra, V. P. Sriram, V. Dadhich, T. Rajasri, E. Slepceva, and H. Pallathadka, "Enhanced integration of graph neural networks for customer segmentation in collaborative interfaces within immersive virtual reality settings," in *Proceedings of 2024 2nd International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)*, Lucknow, India, 2024, pp. 1-5. <https://doi.org/10.1109/IC3TES62412.2024.10877662>
- [13] S. Devaraju, S. Jawahar, M. Manimaran, A. Somasundaram, and M. Thenmozhi, "Performance analysis of marketing campaign with customer profile using machine learning," in *Recent Advances in Artificial Intelligence & Smart Applications*. Singapore: Springer, 2024, pp. 221-233. https://doi.org/10.1007/978-981-97-3485-6_17
- [14] V. Agalya, M. Kandasamy, E. Venugopal, and B. Maram, "CPRO: competitive poor and rich optimizer-

- enabled deep learning model and holoentropy weighted-power k-means clustering for brain tumor classification using MRI,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 36, no. 4, article no. 2252008, 2022. <https://doi.org/10.1142/S0218001422520085>
- [15] L. B. Lisha, and C. Helen Sulochana, “Highly accurate blood vessel segmentation using texture-based modified K-means clustering with deep learning model,” *Concurrency & Computation: Practice & Experience*, vol. 35, no. 7, article no. e7590, 2023. <https://doi.org/10.1002/cpe.7590>
- [16] E. Ondigo, and C. Wekesa, “Economic viability of distribution network upgrade deferral through BESS sizing from k-means clustered annual load profile data,” *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14517-14524, 2024. <https://doi.org/10.48084/etasr.7189>
- [17] J. Yan, J. Chen, J. Zhan, S. Song, Y. Zhang, M. Zhao, Y. Liu, and W. Xu, “Automatic identification of rock discontinuity sets using modified agglomerative nesting algorithm,” *Bulletin of Engineering Geology and the Environment*, vol. 81, article no. 229, 2022. <https://doi.org/10.1007/s10064-022-02724-w>
- [18] S. Sadhukhan, H. Root-Gutteridge, and B. Habib, “Identifying unknown Indian wolves by their distinctive howls: its potential as a non-invasive survey method,” *Scientific Reports*, vol. 11, article no. 7309, 2021. <https://doi.org/10.1038/s41598-021-86718-w>
- [19] A. S. Parihar, S. K. Chakraborty, A. Sharma, and M. Saini, “A comparative study and proposal of a novel distributed mutual exclusion in UAV assisted flying ad hoc network using density-based clustering scheme,” *Wireless Networks*, vol. 29, no. 6, pp. 2635-2648, 2023. <https://doi.org/10.1007/s11276-023-03327-3>
- [20] J. Li, “Research of density based spatial clustering of applications with noise on TCM patent law evaluation,” in *Proceedings of 2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE)*, Yunlin, Taiwan, 2021, pp. 463-466. <https://doi.org/10.1109/ECICE52819.2021.9645658>
- [21] M. Y. Wang, and R. Cheng, “Research on multi-label classification method of transformer based on DBSCAN (Density-Based Spatial Clustering of Applications with Noise) clustering algorithm,” *Journal of Physics: Conference Series*, vol. 2132, article no. 012008, 2021. <https://doi.org/10.1088/1742-6596/2132/1/012008>
- [22] F. Wang, J. Chen, and F. Liu, “Keyframe generation method via improved clustering and silhouette coefficient for video summarization,” *Journal of Web Engineering*, vol. 20, no. 1, pp. 147-170, 2021. <https://doi.org/10.13052/jwe1540-9589.2018>
- [23] X. Wang, and Y. Xu, “An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index,” *IOP Conference Series: Materials Science and Engineering*, vol. 569, article no. 052024, 2019. <https://doi.org/10.1088/1757-899X/569/5/052024>
- [24] F. Ros, R. Riad, and S. Guillaume, “PDBI: a partitioning Davies-Bouldin index for clustering evaluation,” *Neurocomputing*, vol. 528, pp. 178-199, 2023. <https://doi.org/10.1016/j.neucom.2023.01.043>
- [25] C. Rungruang, P. Riyapan, A. Intarasit, K. Chuarkham, and J. Muangprathub, “RFM model customer segmentation based on hierarchical approach using FCA,” *Expert Systems with Application*, vol. 237(Part B), article no. 121449, 2024. <https://doi.org/10.1016/j.eswa.2023.121449>
- [26] J. Wu, L. Shi, L. Yang, X. Niu, Y. Li, X. Cui, S. B. Tsai, and Y. Zhang, “User value identification based on improved RFM model and K-means++ algorithm for complex data analysis,” *Wireless Communications and Mobile Computing*, vol. 2021, article no. 9982484, 2021. <https://doi.org/10.1155/2021/9982484>
- [27] Y. Liu and C. Chen, “Improved RFM model for customer segmentation using hybrid meta-heuristic algorithm in medical IoT applications,” *International Journal of Artificial Intelligence Tools*, vol. 31, no. 1, article no. 2250009, 2022. <https://doi.org/10.1142/S0218213022500099>
- [28] A. H. L. Chen, Y. C. Liang, and M. Z. Fang, “RFM model and K-means clustering analysis of transit traveller profiles: a case study,” *Journal of Advanced Transportation*, vol. 2022, article no. 1108105, 2022. <https://doi.org/10.1155/2022/1108105>



Shu Ren <https://orcid.org/0009-0003-2988-599X>

She was born in February 1982, and graduated from Dongbei University of Finance and Economics with a master's degree in 2009. She is working at Anyang University as an associate professor. She is interested in marketing.