International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 7539-7552

https://doi.org/10.70517/ijhsa464650

Meta-analysis and structural equation modeling analysis of the synergistic mechanism of marketing strategies for new energy vehicle consumer behavior patterns

Xiaoyi Guo[1,*]
[1] Beijing Chengpeng Automobile Sales and Service Co., Ltd, Beijing, 100000, China
Corresponding authors: (e-mail: alan_guoxiaoyi_edu@163.com).

Abstract In the current wave of automotive manufacturing, understanding consumer preferences and needs and implementing precise marketing strategies can help reduce corporate costs and bring greater economic benefits to automakers. This article focuses on accurately understanding consumer purchasing behavior in a complex market environment, providing strong data support for new energy vehicle marketing strategies. This paper constructs a consumer purchasing behavior pattern recognition algorithm. By improving the SOM neural network algorithm to address the pendulum effect in samples, and combining it with the K-means algorithm, a new P2SOM-K-means clustering algorithm is proposed, which features a shorter learning cycle and faster convergence speed. Using the proposed algorithm to analyze the massive consumer data collected, the paper delves into key behavioral patterns of new energy vehicle consumers, including their purchasing reasons, preferences, purchasing methods, and usage channels. The paper categorizes the collected consumers into five groups, with the fifth group being the primary purchasers. This group is primarily characterized by individuals aged 26–35, company employees, and commuting needs. Finally, after deeply analyzing consumer behavior data, the paper proposes targeted strategies for optimizing marketing.

Index Terms SOM neural network, P2SOM-K-means clustering, purchasing behavior patterns, marketing

I. Introduction

The new energy vehicle market has been one of the fastest-growing sectors globally in recent years, driven primarily by its potential advantages in addressing environmental issues and promoting energy conservation and emissions reduction. From electric vehicles to plug-in hybrids and hydrogen fuel cell vehicles, the variety of new energy vehicles is expanding, and their technological capabilities are continuously improving, fostering sustainable development [1]-[4]. With ongoing technological advancements and policy support, new energy vehicles are not only gradually matching or even surpassing traditional internal combustion engine vehicles in terms of technical performance but also demonstrating significant advantages in environmental protection and energy conservation. This trend is not only reshaping the competitive landscape of the automotive industry but also exerting a profound impact on global energy consumption patterns and environmental protection [5]-[8].

The rapid development of the new energy vehicle market holds significant importance for driving global energy transition and achieving the goal of green and low-carbon development. However, the further development of the new energy vehicle market still faces numerous challenges, such as insufficient consumer awareness of new energy vehicles, inadequate pricing strategies, lack of innovation in marketing methods, and insufficient distribution channels. These issues highlight a key challenge in the market: how to effectively formulate marketing strategies to stimulate and enhance consumer purchasing decisions [9]-[12]. Additionally, issues such as cost concerns, limitations in driving range, and insufficient charging infrastructure directly impact consumers' purchasing decisions [13], [14]. Furthermore, policy support plays a crucial role, with government measures such as purchase subsidies, tax incentives, and the development of charging infrastructure aimed at reducing barriers to purchasing and using new energy vehicles, thereby increasing consumers' willingness to adopt them [15]-[17].

Multiple studies have shown that consumer behavior is closely related to new energy vehicle sales. Literature [18] utilized data mining and deep learning analysis to determine that the primary factors influencing new energy vehicle sales in China are national policies, vehicle infrastructure, demographic characteristics, and safety factors. Literature [19] employs a factor theory model to explore consumer purchasing behavior for new energy hybrid vehicles. Consumer perceived behavioral control, behavioral attitudes, and subjective norms influence actual behavior through behavioral intentions, and there are differences in consumer introversion and extroversion under behavioral attitudes and subjective norms. Literature [20] applies structural equation modeling to analyze female consumers, with significant purchasing decision factors including new energy vehicle performance, effort expectations, social influence, convenience, and enjoyment. Literature [21] explores consumer demand for vehicles through online reviews. New energy vehicle consumers compare brands or models with gasoline vehicles during purchase and place greater emphasis on vehicle service performance. Literature [22] uses machine learning to analyze consumer emotions, which have a positive promotional effect on new energy vehicles. This effect varies significantly across different consumer contexts, while adequate public charging stations can alleviate range anxiety and promote consumption. Literature [23] explores the attention displayed by consumers in their searches and finds that the three search variables “new energy vehicles,” “new energy vehicle batteries,” and “charging stations” positively influence sales, while “vehicle fires” have a negative impact.

Consumer behavior theory provides a framework for analyzing and explaining various behaviors of consumers during the purchase of new energy vehicles [24]. Consumer behavior includes the consumer's purchasing decision-making process, decision-making factors, and preferences. This information is valuable to marketers. By combining marketing strategies with consumer behavior theory, marketers can more effectively design products that meet market demand and adopt appropriate promotional strategies to promote the sales and market penetration of new energy vehicles [25], [26]. In such a market environment, deeply understanding and identifying consumer purchasing behavior, and exploring the underlying motivations, attitudes, and preferences, is of great significance for the formulation of marketing strategies.

To mitigate the pendulum effect caused by small differences between consumer samples, the principle of weight adjustment for forgetting the second-place option was proposed, and the weight adjustment equation was derived. Based on this equation, the SOM neural network was improved and applied to the selection of the number of categories K and initial cluster centers required for the K-means clustering algorithm, thereby developing the P2SOM-K-means clustering algorithm that addresses the parameter pre-selection issue. Subsequently, this paper uses the collected new energy vehicle consumer data as samples, applies the improved P2SOM-K-means algorithm to perform cluster analysis on consumer purchasing behavior, and categorizes new energy vehicle consumers based on their purchasing habits according to the obtained cluster results. Furthermore, it conducts a detailed analysis of consumer preferences regarding vehicle models, price ranges, purchasing reasons, and timing.

II. Data acquisition and preprocessing

II. A. Data Acquisition

II. A. 1) Obtaining word-of-mouth review data from users who have purchased cars

Throughout the project, word-of-mouth plays a crucial role. It not only includes evaluations of the vehicle's quality by existing owners but also information such as their purchase time, location, and vehicle model. The review texts are categorized into ten aspects for evaluation. These specifically include the most satisfying aspect, the least satisfying aspect, space, power and handling, fuel efficiency, comfort, exterior design, interior design, and cost-effectiveness. The latter eight aspects each have corresponding ratings.

Each of the ten word-of-mouth metrics mentioned above is accompanied by approximately 1,000 words of text content from each car owner, and the reviews are quite thorough. We have collected around 800,000 such posts, equivalent to 800,000 articles evaluating cars, which constitutes an excellent resource for text classification. Using the collected text content and ratings, we will train a text classification language model using the Naive Bayes algorithm, then predict the text classification of forum posts, and analyze the sentiment trends of potential car buyers. Another important function of reputation metrics is the ability to use the LDA topic model to predict the themes of posts within each vehicle series, enabling us to identify keywords. For example, if we identify keywords such as “insufficient power” or “too small space” for a particular vehicle series, this indicates that improvements are needed for that series.

II. A. 2) Obtaining personal information data of registered car users

The number of registered users on the AutoHome website has reached 430 million. Analyzing the demographic information of user groups helps automobile manufacturers make decisions about product marketing. For example, if the age of users who have purchased a certain car model is around 30, automobile manufacturers will invest most of their advertising budget in young people around 30 years old. Such value is still very significant for automobile manufacturers.

II. A. 3) Acquisition of forum data posted by automobile users

According to statistics, the AutoHome website features over a thousand vehicle models, each with its own forum section. Thousands of people post comments daily within these model-specific forums. After analyzing the data, our project team defined users actively participating in forums who have not undergone vehicle ownership verification as potential car buyers. For automakers, this data holds particular significance, as identifying potential buyers and targeting them with precise marketing efforts translates to greater profits. In this research project, we crawled the forums of seven vehicle models. Forum content is updated daily, and we typically update the forum data once a month. The data is stored in text files, but the specific storage format uses a key-value structure, which facilitates search and statistical operations.

II. B. Data processing

II. B. 1) Data acquisition by web crawlers

The crawler primarily collects data from three main sources: high-quality posts from users who have purchased vehicles on Koubei, user-posted threads on forums, and user personal information. The posts published by users who have purchased vehicles on Koubei amount to approximately 3.5 GB, while forum data amounts to approximately 6 GB, resulting in a large data volume of around 110 GB. The data scraped from the internet is stored in TXT files and placed on Linux machines in the laboratory. The specifics have already been explained in Section 2.1, so we will not repeat them here.

II. B. 2) Chinese word segmentation

This project uses the Stutter word segmentation method for Chinese word segmentation, which mainly employs the following technologies:

(1) The Trie tree structure is used to efficiently scan words and graphs, generating a directed acyclic graph (DAG) composed of all possible word combinations of Chinese characters in a sentence.

(2) Dynamic programming is used to find the path with the highest probability and identify the maximum segmentation combination based on word frequency.

(3) For unregistered words, an HMM model based on the ability of Chinese characters to form words is used, employing the Viterbi algorithm.

II. B. 3) Filtering invalid words and invalid posts

Invalid words and invalid posts refer to certain posts in the forum that do not contribute to categorization or mislead our judgment regarding the sentiment direction of potential car buyers.

The forum is organized with one forum per car model, and each post receives numerous replies. After research, comments with the following characteristics are classified as spam comments:

(1) Posts with too few words, such as “top,” “sofa,” or “bench.”

(2) Posts unrelated to cars, such as advertisements.

(3) Posts with abnormal titles, such as those named with numbers or special symbols.

Based on the content of the forum, the above three forms of spam comments are defined.

III. Consumer purchasing behavior pattern recognition algorithm

Pattern recognition is an important component of information science and artificial intelligence, often referred to as pattern classification. It involves processing and analyzing various forms of information (numerical, textual, and logical relationships) that represent objects or phenomena, in order to describe, identify, classify, and interpret those objects or phenomena. Specifically, it involves using certain features or criteria, along with specific methods and tools, to organize and study originally disorganized and scattered objects. This facilitates a macro-level understanding and grasp of the abstract characteristics and commonalities of known objects, and makes it easier to make inferences about unknown or future new objects.

Among these, behavioral pattern recognition is an abstract generalization of a group of people performing a certain task, used to measure and analyze what they are doing. Consumer purchasing behavior pattern recognition involves studying consumers and abstractly generalizing their participation in new energy vehicle consumption, transforming the originally chaotic phenomena into regular patterns, thereby facilitating targeted marketing to consumers.

III. A. Algorithm improvement ideas for consumers

Generally speaking, when data is collected from a randomly sampled population, the resulting consumer purchase data is approximately normally distributed. Considering the common characteristics of the sample population, sampling a population restricted to these conditions will result in smaller differences between samples compared to those from a random population. During the training process of a SOM neural network, the dot product between each sample and the corresponding weight vector of each output neuron must be calculated. Samples with smaller differences and similar weight vectors may lead to the pendulum effect in the SOM neural network during the later stages of iteration, where the winning node repeatedly jumps between several local extrema points, causing the network to fail to converge.

To mitigate this situation as much as possible, this paper proposes a weight adjustment principle that forgets the second-place neuron to improve the SOM neural network. After selecting the winning neuron and adjusting the weights of its neighboring neurons, the second-place neuron with the next highest matching score is adjusted according to a set forgetting rate, further widening the gap between its weight vector and that of the best-matching neuron. This differentiated weight vector effectively reduces the pendulum effect between local extrema.

Under this weight adjustment rule, each input of the car user sample adjusts the winning neuron and its surrounding weight vectors to a state more similar to the sample, while distinguishing similar weight vectors to enhance their matching degree with subsequent samples of other types, thereby better reflecting the distribution of the samples.

III. B. SOM Neural Network Related Functions

The influence of the winning neuron in the SOM network on its neighboring neurons propagates from near to far, gradually transitioning from excitation to inhibition, as shown in Figure 1(a). Therefore, not only does the winning neuron itself need to adjust its weight vector, but the neurons surrounding it also need to adjust their weight vectors to varying degrees under its influence [27], [28]. This adjustment can be represented by the following functions.

(1) Mexican hat function: The winning node has the largest weight adjustment, while neighboring nodes have slightly smaller adjustments. As the distance from the winning node increases, the weight adjustment decreases until it reaches zero at a certain distance. When the distance increases further, the weight adjustment becomes negative, and then returns to zero at an even greater distance. As shown in Figure 1(b).

(2) Bowler Hat Function: This is a simplified version of the Mexican Hat Function, as shown in Figure 1(c).

(3) Chef's Hat Function: This is a simplified version of the Bowler Hat Function, as shown in Figure 1(d).

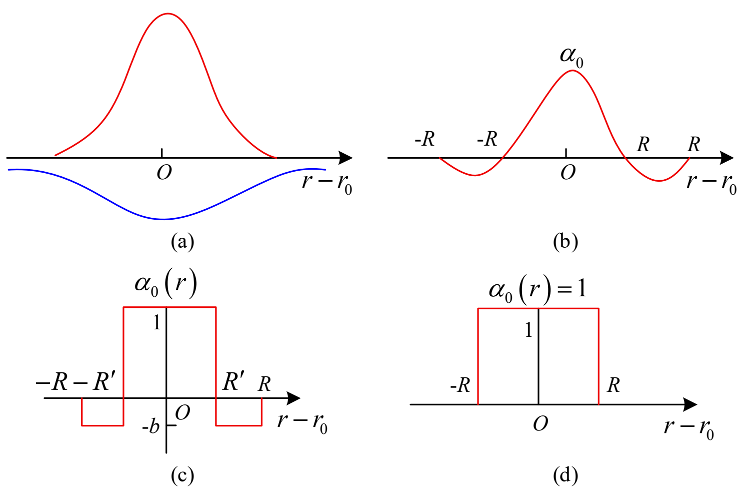


Figure 1: Function of SOM neural network

The function defined by the learning rate should decrease as the number of iterations t increases, as it affects the convergence speed of the model. At the beginning of training, the learning rate is initialized with a large value and then decreases rapidly, allowing the model to quickly understand the general sample structure of the input vectors. Subsequently, to make the competitive layer as closely aligned as possible with the sample distribution structure of the input space, the learning rate should decrease gradually from a smaller value to 0, enabling precise weight adjustment. The learning rate $\eta(t)$ can be defined in a form similar to the winning neighborhood radius, as their forms are quite similar. Common definitions include:

$$\eta(t) = C_2\left(1-\frac{t}{t_m}\right) \tag{1}$$

and a function where $\eta(t)$ decreases linearly to 0 over the training time:

$$\eta(t) = C_2 \exp\left(\frac{-B_2 t}{t_m}\right) \tag{2}$$

In the above equation, C_2 is a constant between 0 and 1. B_2 is a constant greater than 1.

Considering that the differences between car user samples are small, after multiple iterations, it is easy for the winning node to repeatedly jump between local extrema. This study will redesign the weight adjustment equation to weaken the pendulum effect of the SOM neural network in the later stages of iteration as much as possible.

III. C. P2SOM weight adjustment equation

This paper proposes a SOM neural network that uses the forget-second-place rule to adjust the weight vector, named the P2SOM neural network. The weight adjustment process is as follows:

Let the input vector be $X = \left(x_1, x_2, \ldots, x_p\right)^T$, and the competitive layer contain k output neurons. Then, the weight vector of output neuron j is:

$$w_j = \left(w_{1j}, w_{2j}, \ldots, w_{pj}\right)^T, \left(j = 1, 2, \ldots, k\right) \tag{3}$$

The best-matching node for the input vector X will become the cluster center of its affiliated category, defined as the competing layer node with the highest similarity to X, denoted by w^*, satisfying the condition:

$$S\left(X, w^*\right) \geq S\left(X, w_j\right), j = 1, 2, \ldots, k \tag{4}$$

Here, $S\left(A, B\right)$ represents the Euclidean distance, which is used to measure the similarity between vectors and nodes.

The corresponding definition formula is as follows:

$$S\left(X, w_j\right) = \sqrt{\left(x_1 - w_{1j}\right)^2 + \ldots + \left(x_p - w_{pj}\right)^2} \tag{5}$$

In the above equation, p represents the dimension of the input vector X, and the point with the maximum value of $S\left(X, w_j\right)$ corresponds to the best matching node.

To make the adjustment of the weight vector more suitable for the distribution characteristics of the car user samples, the “forget the second” weight adjustment rule is introduced here to increase the differences between similar weight vectors. The coefficient μ_j of the “forget the second” rule is defined as:

$$\mu_j = \begin{cases} 1 & j = c \\ -1 & j = r \\ 0 & \text{Otherwise} \end{cases} \tag{6}$$

Let m_i represent the number of times $\mu_i = 1$, and define γ_j as follows:

$$\gamma_j = m_j \Big/ \sum_{i=1}^{k} m_i \tag{7}$$

Let c represent the winning node, which is the best matching node, hereinafter referred to as the winner, satisfying the following conditions:

$$\gamma_c * S(X, c) \geq \gamma_j * S\left(X, w_j\right) \tag{8}$$

Let r represent the second matching node, referred to below as the second place, which satisfies the following conditions:

$$\gamma_r * S(X, r) \geq \gamma_j * S_{j \neq c}\left(X, w_j\right) \tag{9}$$

Introducing the concept of sample density, let $d\left(x_i\right)$ denote the density of sample x_i, whose specific definition formula is:

$$d\left(x_i\right) = \sum_{j=1}^{k} \frac{d\left(x_i, x_j\right)}{\sum_{i=1}^{n} d\left(x_i, x_j\right)} \tag{10}$$

Let α denote the learning rate and β denote the forgetting rate, where $\alpha \square \beta$. Combining the above formulas, we obtain the following weight adjustment equation. During the iteration process, the P2SOM neural network will simultaneously adjust the weights of the winner c and the runner-up r according to this equation:

$$\begin{cases} \frac{dw_{ij}}{dt} = \alpha * d(x_i) * [X_i(t) - w_{ij}(t)] & j = c \\ \frac{dw_{ij}}{dt} = -\beta * d(x_i) * [X_i(t) - w_{ij}(t)] & j = r \\ \frac{dw_{ij}}{dt} = 0 & \text{Otherwise} \end{cases} \tag{11}$$

Due to the small differences between car user samples, conventional SOM neural networks may exhibit a pendulum effect where winning neurons repeatedly jump between several local extrema nodes after multiple iterations, which can affect the stability of convergence. The weight adjustment equation using the forgetting rate mentioned above can increase the gap between the top two neural nodes. The improved P2SOM neural network can differentiate the weight vectors in each iteration, even when the differences between automotive user samples are small, enabling the distribution of weight vectors to better reflect the sample distribution characteristics of each category.

III. D. P2SOM-K Mean Clustering Algorithm

The algorithm can be divided into two parts: first, the P2SOM neural network performs an initial clustering of the data samples. Similar feature vectors are identified and classified by the neural network during training, resulting in K clusters of sample data. This determines the number of clusters in the K-means clustering algorithm and yields the initial cluster centers Z. Subsequently, the results obtained earlier are used as initial values for the K-means clustering algorithm, and after iterative calculations, the final clustering results are obtained [29], [30].

(1) Weight initialization: Let w_j be the weight vector connecting the input node to the j th output node, and assign it a random initial value.

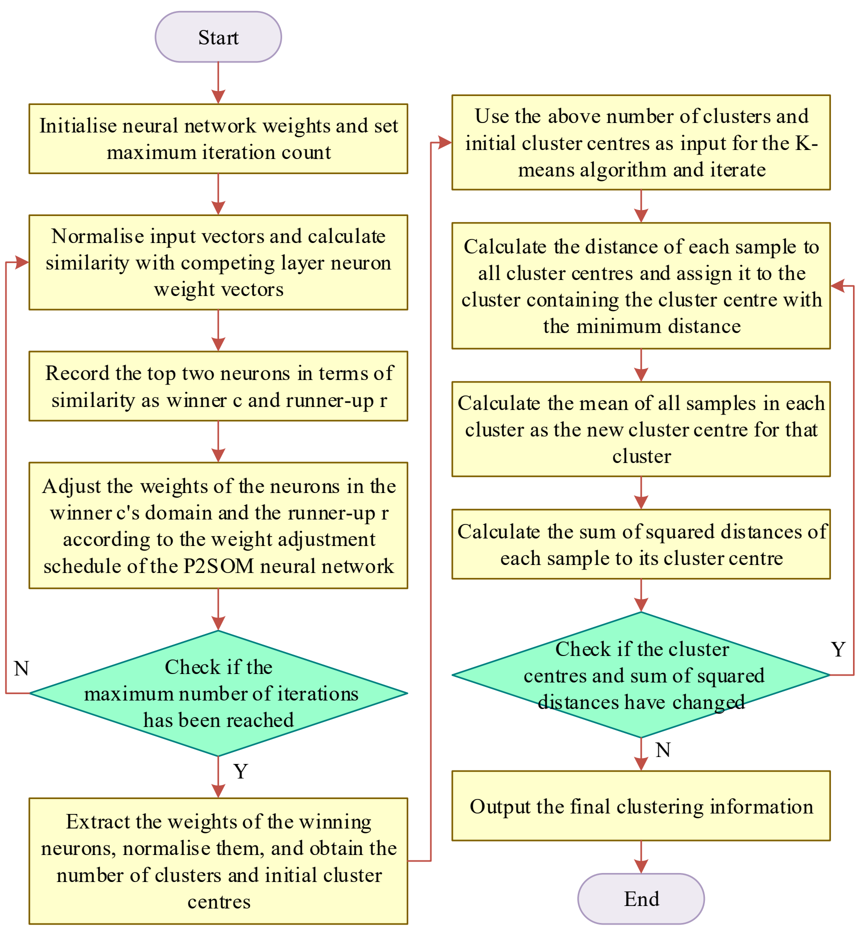


Figure 2: P2SOM-K-mean clustering algorithm flowchart

(2) Find the connection weight vector c with the smallest distance from each input sample $X_i\left(i=1,2,\cdots,m\right)$ in w_j, and the weight vector r with the second smallest distance from X_i. For the winner c and runner-up r, refer to the above P2SOM neural network for the weight vector adjustment equation for the neighborhoods of the two. Repeat the training process. When the weight coefficients of the competitive layer neural network stabilize, convergence is achieved, and the initial clustering of the samples is completed based on the feedback from the output nodes.

(3) Output the number of clusters K and the initial cluster centers $Z=(z_1,z_2,\cdots,z_K)$.

(4) The initial input parameters for the $K-$ mean clustering algorithm are iterated using the above output results (cluster number K and initial cluster centers Z).

(5) Iterate until convergence, using the output results from step (3) as the initial input values for the $K-$ mean clustering algorithm, until the convergence condition is met.

(6) Output the final clustering information.

The specific flowchart of the P2SOM-K-mean clustering algorithm is shown in Figure 2 below:

III. E. Selection of K value

The number of clusters in this data processing was determined using the elbow method. The “norm” function in MATLAB was used to calculate the distance between the cluster centers and the sample points. The number of clusters k was set to a range from 2 to 10 in a loop. The sum of squared errors (SSE) within clusters obtained for different values of k were compared and plotted to generate a line chart showing the SSE values for different k values, as shown in Figure 3. The optimal k value is selected at the point where the line becomes smooth, so k=5 was chosen for clustering.

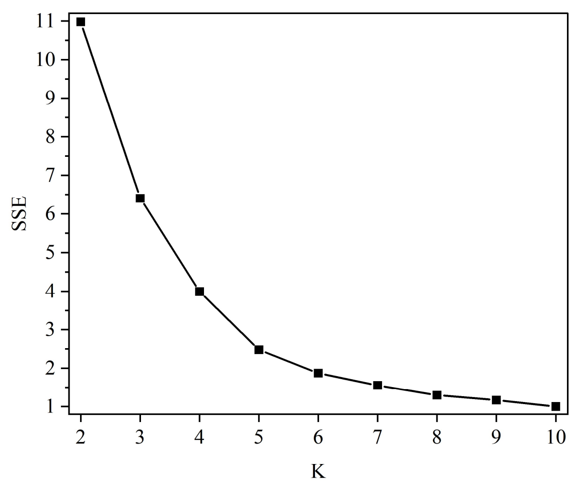


Figure 3: Elbow diagram

IV. Analysis of consumer purchasing behavior

IV. A. Clustering Results Analysis

Based on the information expressed in the elbow diagram above, an initial attempt was made to perform cluster analysis on the valid instances in the acquired data using Python programming when the number of clusters k = 5.

Using the P2SOM-K-mean clustering algorithm for preliminary calculation, the initial cluster data distribution is shown in Figure 4. Based on a rough judgment of the data, the five centers are clustered, which is basically the same as the optimal k value of 5 judged in Figure 3. This program is set with a maximum of 600 iterations. When the iteration reaches around 400 times, the cluster center positions change little, resulting in a relatively stable cluster data distribution diagram, as shown in Figure 5.

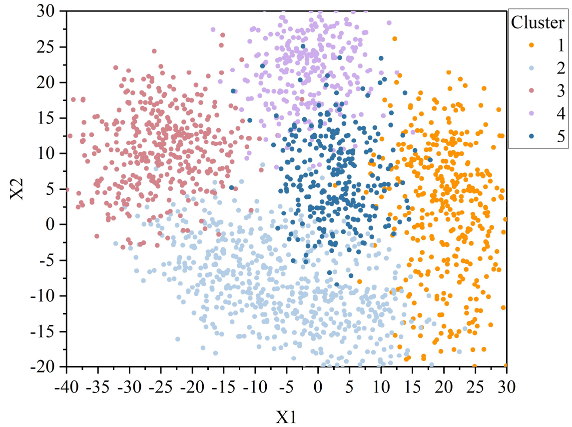
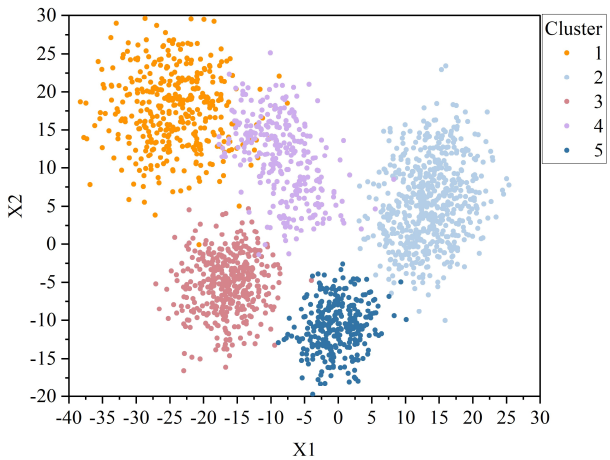
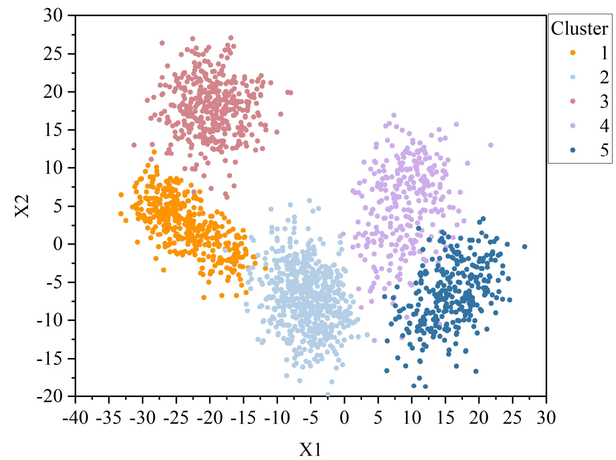


Figure 4: Initial clustering data profile



(a)Intermediate process



(b)Cluster final results

Figure 5: K= 5 hours are based on the iterative process profile

To provide a more detailed view of the importance of feature attributes for different numbers of individuals and different groups within each cluster, this paper presents intra-class importance as percentages. The overall cluster analysis results are shown in Table 1. The following explanations are provided for the numerical values in the table:

(1) Gender: Male 1, Female 2.

(2) Age: 1: under 25 years old, 2: 26–35 years old, 3: 36–45 years old, 4: 46 years old and above.

(3) Region: 1: Northern region, 2: Southern region, 3: Region with vehicle restrictions.

(4) Occupation: 1: Public institution staff, 2: Company employee, 3: Student, 4: Freelancer, 5: Other.

(5) Annual income: 1: Under 100,000, 2: 100,000–200,000, 3: 200,000–300,000, 4: 300,000–500,000, 5: Over 500,000.

(6) Car purchase needs: 1: Daily commuting for work, 2: Family needs, 3: Weekend outings, 4: Hobbies, 5: Other.

(7) Low living costs: 5: Strongly agree, 4: Agree, 3: Neutral, 2: Disagree, 1: Strongly disagree.

(8) Tax incentives: 5: Strongly agree, 4: Agree, 3: Neutral, 2: Disagree, 1: Strongly disagree.

(9) Supporting infrastructure development: 5: Strongly agree, 4: Agree, 3: Neutral, 2: Disagree, 1: Strongly disagree.

(10) Range: 5: Strongly agree, 4: Agree, 3: Neutral, 2: Disagree, 1: Strongly disagree.

(11) Battery life: 5: Strongly agree, 4: Agree, 3: Neutral, 2: Disagree, 1: Strongly disagree.

(12) Promotion: 5: Strongly agree, 4: Agree, 3: Neutral, 2: Disagree, 1: Strongly disagree.

(13) Purchase intention: 1: Yes, 2: No.

Category 5 accounts for 52.81%, Category 3 accounts for 18.66%, and so on, with Category 1 accounting for 5.58%.

Category 5 consists of individuals aged 26–35 who are company employees with commuting needs. This category represents the majority of consumers, specifically those with a strong intention to purchase new energy vehicles.

The fourth category requires new energy vehicles to have strong policy support, well-developed supporting infrastructure, and long-range capabilities, representing a group with high demands for new energy vehicles.

The third category consists of freelancers who purchase new energy vehicles out of personal interest under favorable tax incentives, accounting for the second-largest share. This group also requires strong policy support.

The second category comprises individuals around 45 years old with family-related needs, though their share is low, their demands are still evident.

The first category is similar to the fourth category, but they do not require very high external conditions to purchase new energy vehicles. In summary, the proportion of essential demand groups accounts for more than 50%. This clustering result is highly consistent with people's real-world perceptions, which not only validates the reliability of the survey data but also further confirms the scientific and accurate nature of the clustering analysis method.

Table 1: Cluster analysis results (k=5)

Cluster	1	2	3	4	5
Size	5.58%	9.54%	18.66%	13.41%	52.81%
Input	Fiscal and tax preferential policies 3(99.29%)	Car demand 2(56.77%)	Occupation 4(73.68%)	Battery life 5(91.88%)	Car demand 1(69%)
	Supporting facilities 3(98.95%)	Occupation 5(35.72%)	Car demand 4(54.48%)	Supporting facilities 5(95.48%)	Occupation 2(64.5%)
	Endurance range 3(88.86%)	Age 4(49.4%)	Fiscal and tax preferential policies 4(88.33%)	Fiscal and tax preferential policies 5(91.02%)	Age 2(38.7%)
	Battery life 3(96.43%)	Endurance range 3(57.73%)	Age 1(98.97%)	Endurance range 5(83.8%)	Publicity 5(48.34%)
	Low cost of living 3(90.2%)	Fiscal and tax preferential policies 4(48.49%)	Supporting facilities 4(79.56%)	Low cost of living 5(71.26%)	Annual income 2(39.4%)
	Publicity 3(96.29%)	Supporting facilities 3(47.54%)	Endurance range 4(81.43%)	Car demand 4(34%)	Region 2(49.78%)
	Occupation 4(94.86%)	Publicity 3(52.57%)	Low cost of living 4(67.08%)	Occupation 4(59.07%)	Fiscal and tax preferential policies 5(51.54%)
	Car demand 4(56.99%)	Battery life 4(49.01%)	Publicity 4(51.97%)	Gender 1(73.48%)	Supporting facilities 5(52.01%)
	Age 1(95%)	Annual income 1(78.58%)	Region 3(66.29%)	Age 1(82.99%)	Endurance range 5(49.42%)
	Gender 1(72.02%)	Region 2(43.02%)	Battery life 4(67.58%)	Publicity 5(62.79%)	Low cost of living 4(39.93%)
	Region 3(74.73%)	Whether or not 1(81.99%)	Annual income 2(41.53%)	Annual income 1(62.74%)	Gender 2(66.08%)
	Annual income 1(64.25%)	Gender 2(69.12%)	Whether or not 1(84.72%)	Region 3(55.64%)	Battery life 4(48.4%)
	Whether or not 1(91.15%)	Low cost of living 4(45.26%)	Gender 2(57.38%)	Whether or not 1(87.48%)	Whether or not 1(90.61%)

IV. B. Analysis of User Purchasing Behavior

In the previous section, this study utilized the P2SOM-K-mean clustering algorithm to analyze data collected on new energy vehicle consumers, successfully dividing consumers into five distinct groups and conducting an in-depth analysis of their purchasing needs and preferences. These analyses not only revealed potential trends in the new energy vehicle market but also provided valuable insights into consumer purchasing behavior. However, understanding consumers goes beyond this. In this chapter, we will further delve into the information behind these data to provide a more detailed and thorough description of the purchasing behavior of new energy vehicle consumers.

IV. B. 1) Purchase Target

User requirements and expectations for new energy vehicles, from the perspective of energy type, show that pure electric vehicles are currently the most favored by users, followed by plug-in hybrid electric vehicles, while a portion of users still prefer gasoline-powered vehicles. The preferred vehicle models are shown in Table 2. From the perspective of vehicle models, microvans are the most popular in the market, with 39.52% of surveyed users indicating a preference for purchasing microvans. Only 4.36% of surveyed users indicated a preference for purchasing heavy-duty trucks. Due to the immaturity of battery technology and insufficient range, users remain skeptical about high-power models such as heavy-duty trucks. The price range under consideration is shown in Table 3. From the perspective of price, 56.65% of users consider a price range below 100,000 yuan, exceeding half of the total, indicating that users are highly sensitive to purchasing costs.

Table 2: Tends to buy models

Model	Proportion /%
Microsurface	39.52
Microcalorie	14.92
Light card	16.35
Medium card	8.07
Heavy card	4.36
Pickup truck	5.59
Ling	4.93
Sea lion	6.26

Table 3: Considers the price of the purchase

Price /Yuan	0~50000	50000~100000	100000~150000	150000~200000	More than 200000
Proportion /%	32.18	24.47	18.93	6.27	18.15

IV. B. 2) Reasons for purchase

The motivations for purchasing new energy vehicles are shown in Table 4, with the primary factors being energy conservation and environmental protection, vehicle performance, low operating costs, price subsidies, priority registration, road access, and tax incentives, accounting for 56.03%, 17.96%, 13.62%, 5.85%, 3.73%, 2.31%, and 0.5%, respectively. It is evident that the inherent advantages of new energy vehicles in terms of energy conservation, environmental protection, and low operational costs significantly stimulate consumer purchasing intent, while government favorable policies play a certain promotional role.

Table 4: Reasons for buying

Motive	Proportion/%
Energy conservation and environmental	56.03
Protection	17.96
Performance	13.62
Low cost	5.85
Price subsidy	3.73
Preferred	2.31
The right of road is open	0.5

IV. B. 3) Information channels/purchase locations/regions of use

Before purchasing a vehicle, users typically gather information through multiple channels and make a purchasing decision only after thoroughly understanding the vehicle's specifications. The channels through which users obtain vehicle information and the regional distribution of vehicle usage are shown in Table 5. Among the information channels for new energy vehicles, the internet accounts for 67.43%, while recommendations from acquaintances account for 13.16%. This indicates that the internet is the most important channel for users to obtain vehicle information, while the role of recommendations from acquaintances should not be overlooked.

Group users, such as internet logistics platforms and car rental companies, primarily customize personalized vehicles directly through manufacturers. Among individual users, 80.35% of surveyed users indicated a preference for purchasing at brand specialty stores, while 4.69% of surveyed users indicated a preference for purchasing through online marketplaces. This indicates that offline physical brand specialty stores remain the primary purchasing channel, while online marketplaces also attract a small number of early adopters.

In terms of usage regions, most new energy vehicles are currently used within cities, with a significant proportion also used in townships. The rural market segment also holds a notable share.

Table 5: Information channel

	Proportion/%
Information channel	
Network	67.43
Used car market	5.11
Store publicity	5.07
Acquaintance introduction	13.16
Television advertising	4.63
Newspapers	4.60
Area of use	
In the city	52.50
Urban and rural	14.56
Intercity	6.69
Countryside	7.84
Township	18.41

IV. B. 4) Purchase Methods

With the widespread adoption of online payment platforms, their convenience and reliability have gained widespread popularity. WeChat Pay and Alipay have become the primary payment methods for car buyers. The promotional methods preferred by consumers are shown in Table 6, with price discounts being the most popular at 46.13%, followed by after-sales maintenance and repair discounts.

Table 6: Promotion Methods

Promotion mode	Price discount	After-sales maintenance/ Maintenance discount	Giveaway	Policy subsidy	Trade-in/ Deduction of price for old vehicles
Proportion /%	46.13	31.36	6.97	11.84	3.7

IV. B. 5) Comparison and analysis of various brands of new energy vehicles

Table 7 shows the sales performance of various brands in the 2024 new energy vehicle market, reflecting the preference trends of ordinary new energy vehicle consumers. Given the numerous new energy vehicle brands available in the market, this paper specifically selects the top ten brands by sales volume in the 2024 new energy vehicle market for in-depth analysis, providing consumers with more targeted reference information.

As shown in the table, many consumers still prefer to choose products from traditional automakers such as BYD when purchasing new energy vehicles. As a leader in the new energy vehicle market, BYD has achieved comprehensive development through deep investment in research and development and business model innovation, exerting a profound influence on fields such as batteries, new energy vehicles, and rail transit. Its products perform excellently in terms of driving range, safety performance, and comfort, and have become the top seller in 2024 sales data.

Table 7: 2024 different new energy car brand market performance

Brand	Sales (vehicles)	Year-on-year growth /%
Byd	4104354	12.5
Geely	888154	91.04
Tesla	1790665	-1.3
Saic	12340228	9.56
Gac	430492	-

Great Wall	322853	23.15
Chery	583335	235.68
K wei	221765	39.37
Xiao peng	190522	83.18
Beiqi group	81158	172.4

IV. B. 6) Analysis of New Energy Vehicle Consumer Purchase Timing

Consumers should thoroughly understand market trends and combine them with their personal needs to choose the right time to purchase a vehicle in order to obtain the best purchasing experience and value for money. In response to this issue, this study compiled statistics on the monthly sales volume of new energy vehicles from 2022 to 2024, as shown in Figure 6. New energy vehicle sales reach their peak at the end of the year, with multiple intertwined factors contributing to this trend. On one hand, the end of the year is traditionally a peak season for vehicle purchases, as consumers often plan for the new year and anticipate increased travel during the Spring Festival, leading them to choose this time to purchase vehicles. As the Spring Festival and other holidays approach, travel demand increases, and new energy vehicles are highly favored due to their environmental and energy-saving characteristics. On the other hand, automakers typically launch a variety of promotional activities and discounts to meet their annual sales targets, further attracting consumer attention and stimulating purchasing desire. Additionally, as consumer awareness and acceptance of new energy vehicles increase, coupled with ongoing technological advancements, the new energy vehicle market has become increasingly mature, providing strong support for the surge in year-end sales. Furthermore, this study observed a noticeable sales trend. From January to April, vehicle sales generally remained at a low level. This is primarily because, after the Spring Festival, most people are busy with work and daily life, leaving them little time to consider purchasing a vehicle. As May arrives, vehicle sales begin to gradually recover. This is mainly due to important holidays such as Labor Day and National Day, which provide consumers with favorable opportunities to purchase vehicles.

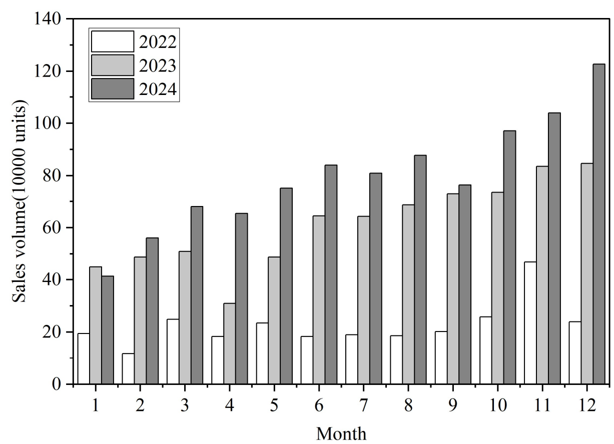


Figure 6: New energy auto month sales change trend

V. Marketing strategies for new energy vehicles

After analyzing consumer car purchasing behavior using the P2SOM-K-mean clustering algorithm, combined with marketing theory, recommendations are made for the selection of marketing strategies for the new energy vehicle market.

V. A. Focus on marketing and promotion strategies

When formulating marketing strategies, companies should focus on long-term and sustainable approaches, shifting away from traditional marketing and promotional concepts. They should continuously build consumers' environmental awareness and understanding of new energy vehicles for household use, thereby enhancing consumers' positive attitudes toward such vehicles. Companies can adopt innovative promotional strategies to educate the market about the technological advancements and environmental benefits of new energy vehicles. By leveraging influential media channels, they can accurately report on urban traffic pressures and environmental issues, using statistical data to demonstrate the tangible effects of new energy vehicles in improving environmental quality. Regarding new energy vehicle products themselves, companies should focus on enhancing consumer understanding through marketing strategies, transforming negative perceptions, and stimulating interest in new energy vehicles.

V. B. Focus on word-of-mouth marketing strategies

The new energy vehicle industry is currently at a critical stage of development. Government support has attracted a significant portion of the existing consumer base. In their marketing efforts, companies should focus on strengthening consumer word-of-mouth marketing strategies to influence purchasing behavior through the subjective norms of consumers.

Satisfaction follow-ups are a crucial component of word-of-mouth marketing strategies. By conducting surveys to understand existing consumers' evaluations of home-use new energy vehicle products, companies can promptly address issues encountered during use, enhance after-sales service, and improve consumer satisfaction, thereby fostering a positive reputation within the user community. Additionally, companies can regularly conduct product-related social surveys, facilitate topic discussions among consumer groups, and establish consumer communication platforms such as “car owner groups” or “membership programs.” Regularly monitor public opinion and integrate it with the product manufacturing and R&D process to determine the product's core value proposition, aligning it with consumer word-of-mouth and increasing the user base for existing products. Strengthen corporate brand image construction to elevate the company's status in consumers' minds, thereby exerting a positive normative influence within the group. Through the group's subjective norms, drive the occurrence of purchasing behavior for new energy vehicles.

V. C. Focus on price marketing strategies

The government provides consumers with certain subsidies and tax breaks for purchasing vehicles, and there are corresponding convenience policies for subsequent use, including: no restrictions on vehicle use, no restrictions on driving, and reduced parking fees, This significantly reduces the purchase and usage costs for consumers. For new energy vehicle manufacturers, it is important to seize this policy opportunity, combine it with the company's own production costs, and adopt corresponding pricing strategies, including government subsidy programs, installment payments, and tiered pricing. Additionally, companies should focus on actively reducing production and organizational management costs to lower the overall cost of new energy vehicles and achieve a competitive advantage in pricing. On the other hand, the government should also focus on improving relevant subsidy policies, further clarifying subsidy standards and scope, collaborating with companies to stabilize new energy vehicle market prices, and increasing infrastructure construction efforts to create a favorable consumer market environment.

In terms of new energy vehicle technology development, companies should focus on researching and developing high-quality, high-performance products, enhancing research and development capabilities, and reasonably allocating internal resources to maximize efficiency and benefits.

VI. Conclusion

Any company seeking sustained growth and a solid foothold in a highly competitive market economy must not overlook the importance of in-depth analysis and insights into consumer purchasing behavior characteristics. This paper employs data collection, data processing, and the development of a consumer purchasing behavior pattern recognition algorithm to uncover the characteristics of new energy vehicle consumers' purchasing behavior. According to the results, energy efficiency and environmental protection, vehicle performance, and low usage costs are the primary reasons consumers purchase new energy vehicles, accounting for 56.03%, 17.96%, and 13.62% respectively. 67.43% of consumers obtain vehicle information online but choose to purchase vehicles at offline specialty stores. New energy vehicle users primarily utilize their vehicles in urban and rural areas. Therefore, companies should intensify online promotional efforts for their vehicles while enhancing the service quality of their offline stores. Based on consumer insights, this paper proposes a practical marketing strategy framework encompassing promotion, reputation, and pricing.

References

[1] Liu, Z., Hao, H., Cheng, X., & Zhao, F. (2018). Critical issues of energy efficient and new energy vehicles development in China. Energy Policy, 115, 92-97.

[2] Lee, E., & Mah, J. S. (2021). Environmental protection and development of technology-intensive industries: The case of new energy vehicle industry in Korea. Science, Technology and Society, 26(3), 413-432.

[3] Zhang, L., Wang, L., & Chai, J. (2020). Influence of new energy vehicle subsidy policy on emission reduction of atmospheric pollutants: a case study of Beijing, China. Journal of Cleaner Production, 275, 124069.

[4] Rao, Y. (2020). New energy vehicles and sustainability of energy development: Construction and application of the Multi-Level Perspective framework in China. Sustainable Computing: Informatics and Systems, 27, 100396.

[5] Peng, L., & Li, Y. (2022). Policy evolution and intensity evaluation of the Chinese new energy vehicle industry policy: The angle of the dual-credit policy. World Electric Vehicle Journal, 13(5), 90.
[6] Hu, R., Cai, T., & Xu, W. (2024). Exploring the technology changes of new energy vehicles in China: Evolution and trends. Computers & Industrial Engineering, 191, 110178.
[7] Liao, H., Peng, S., Li, L., & Zhu, Y. (2022). The role of governmental policy in game between traditional fuel and new energy vehicles. Computers & Industrial Engineering, 169, 108292.
[8] Yang, T., Xing, C., & Li, X. (2021). Evaluation and analysis of new-energy vehicle industry policies in the context of technical innovation in China. Journal of Cleaner Production, 281, 125126.
[9] Zhang, J., & Wang, R. (2019). Research on the marketing strategy of new energy vehicles in SL company. American Journal of Industrial and Business Management, 9(2), 306-314.
[10] Dezhi, P., Navavongsathian, A., Chen, W. K., Rungruang, P., Assawasitisilp, D., & Sukkawan, J. (2022). The New Energy Vehicles Marketing Strategies: A Case Study of Li Auto Inc. APHEIT International Journal of Interdisciplinary Social Sciences and Technology, 11(2), 104-115.
[11] Hasan, F., & Islam, M. R. (2022). New energy vehicles from the perspective of market and environment. Journal of Business Strategy Finance and Management, 4(1), 38.
[12] Sofwan, F., & Sukaris, S. (2023). Marketing Challenges in The Era of The Transition of Fuel Oil Vehicles to Electric Vehiclese. Innovation Research Journal, 4(2), 101-107.
[13] Pang, J., Ye, J., & Zhang, X. (2023). Factors influencing users' willingness to use new energy vehicles. Plos one, 18(5), e0285815.
[14] Liua, Q. (2024). An Empirical Study on the Development of New Energy Vehicle Technology and Consumer Car Purchase Behavior. Journal of Advanced Academic Research and Studies (JAARS), 1(3), 27-36.
[15] Tang, X., Feng, J., Feng, B., Mao, X., & Wei, X. Z. (2024). Policy analysis on the promotion of new energy vehicles in China considering consumers' car purchasing choices in the "post-subsidy era": Based on the study of a three-party evolutionary game. Environment, Development and Sustainability, 1-32.
[16] Lou, Y., Wang, W., & Yang, X. (2017). Customers' attitude on new energy vehicles' policies and policy impact on customers' purchase intention. Energy Procedia, 105, 2187-2193.
[17] Tian, X., Zhang, Q., Chi, Y., & Cheng, Y. (2021). Purchase willingness of new energy vehicles: A case study in Jinan City of China. Regional Sustainability, 2(1), 12-22.
[18] Wang, L., Fu, Z. L., Guo, W., Liang, R. Y., & Shao, H. Y. (2020). What influences sales market of new energy vehicles in China? Empirical study based on survey of consumers' purchase reasons. Energy policy, 142, 111484.
[19] Guo, Q., & You, W. (2023). Research on psychological attributions and intervention strategies of new energy hybrid vehicle purchase behavior. Scientific Reports, 13(1), 9853.
[20] Zhao, J., Su, Y., Fang, M., & Su, M. (2024). Embracing new energy vehicles: An empirical examination of female consumer perspectives. Journal of Retailing and Consumer Services, 80, 103925.
[21] Wang, X., Cheng, Y., Lv, T., & Cai, R. (2023). Fuel vehicles or new energy vehicles? A study on the differentiation of vehicle consumer demand based on online reviews. Marketing Intelligence & Planning, 41(8), 1236-1251.
[22] Liu, Y., Zhang, M., Chen, X., Li, K., & Tang, L. (2024). The Impact of Consumer Sentiment on Sales of New Energy Vehicles: Evidence from Textual Analysis. World Electric Vehicle Journal, 15(7), 318.
[23] Jiang, Z., Long, Y., & Zhang, L. (2021). How Does the Consumers' Attention Affect the Sale Volumes of New Energy Vehicles: Evidence From China's Market. Frontiers In Energy Research, 9, 782992.
[24] Hu, X., Yusof, R. N. R., & Mansor, Z. D. (2025). Consumers' Purchase Intentions Towards New Energy Vehicles Based on the Theory of Planned Behaviour on Perceived Value: An Empirical Survey of China. World Electric Vehicle Journal, 16(3), 120.
[25] Hu, Z. (2022). Research on the consumer behavior characteristics and marketing strategy of new energy vehicles—Taking BYD and tesla as examples. BCP Business & Management, 31, 168-175.
[26] Sun, B., & Ju, Z. (2023). Research on the promotion of new energy vehicles based on multi-source heterogeneous data: consumer and manufacturer perspectives. Environmental Science and Pollution Research, 30(11), 28863-28873.
[27] Liu Xu,Wang Xuewen,Du Xiaowei & Gu Peng. (2022). Analysis of Efficiency of Human Resource Management Evaluation Model Based on SOM Neural Network. Security and Communication Networks,2022,
[28] Sun X.,Collins R. & Kim J. (2001). A COMPARISON OF SOM NEURAL NETWORKS AND K-MEANS CLUSTERING USING REAL WORLD DATA: CHINESE CONSUMER ATTITUDES TOWARDS IMPORTED FRUIT. Acta Horticulturae,(566),185-191.
[29] Okereke George Emeka,Bali Mohamed Chaker,Okwueze Chisom Nneoma,Ukekwe Emmanuel Chukwudi,Echezona Stephenson Chukwukanedu & Ugwu Celestine Ikechukwu. (2023). K-means clustering of electricity consumers using time-domain features from smart meter data. Journal of Electrical Systems and Information Technology,10(1),
[30] Taghipour Anari Esmat,Zegordi Seyed Hessameddin & Albadvi Amir. (2025). The type of supplier involvement in new product development in the automotive industry: metaheuristic-based K-means clustering and analytic hierarchical process methods. Journal of Advances in Management Research,22(1),90-110.