Risk Assessment Method for New Energy Vehicle Supply Chain Based on Hierarchical Holographic Model and Matter Element Extension Model

Qiankun Jiang[1,2*], Haiyan Wang[1]
[1]School of Transportation and Logistics Engineering, Wuhan University of Technology
Wuhan 430063, China
[2]Department of Management, Hubei University of Technology Engineering and Technology College
Wuhan 430068, China
E-mail: 290317@whut.edu.cn, jiangqiankun163@sina.com
*Corresponding author

Keywords: new energy vehicles, hierarchical holographic model, matter element extension, risk identification, scientific evaluation

Received: August 21, 2024

New energy vehicles provide new solutions for low-carbon emissions. With the continuous expansion of the new energy vehicle industry, a more scientific supply chain management system is needed to effectively identify and evaluate risks. This study proposes a more scientific and comprehensive risk assessment method for the supply chain system of new energy vehicles based on a hierarchical holographic model and matter element extension model. The results showed that the proposed algorithm improved the classification performance of risk factors by 1.2%, 1.3%, and 1.5% compared to the clustering performance of the nearest frequency amount clustering, particle swarm optimization, and self-organizing mapping algorithms. In terms of noise processing effectiveness, the Rand index has improved by an average of 55% and 41% compared to the kernel density threshold algorithm and spectral clustering algorithm, with smaller fluctuations and significant differences ($P<0.05$). The accuracy, recall, and F-measure were 9.4%, 8.5%, and 9.6% higher than traditional spectral clustering algorithms, with smaller fluctuations and significant differences ($P<0.05$). While reducing the risk handling time by 21%, the effect has improved by 6%, and the fluctuations during the risk handling process were smaller than those of the self-exploration model. Therefore, the proposed algorithm can cope with many uncertain factors in the complex supply chain management system, ensuring the sustainability and stability of the development of the new energy vehicle industry supply chain.

Povzetek: Predlagana je metoda ocenjevanja tveganj v dobavni verigi električnih vozil, ki združuje hierarhični holografski model in model razširitve materialnih elementov.

1 Introduction

The issue of global warming caused by the increase in carbon emissions has become increasingly prominent and has received widespread attention from the international community. The importance of developing clean energy and energy transformation has deeply penetrated people's hearts [1]. New Energy Vehicles (NEVs) are an important way to achieve low-carbon emissions. The growth in the social demand for NEVs is being driven by national policies, while the Supply Chain System (SCS) for automobiles has also expanded significantly [2]. The SCS of NEVs involves business processes in all aspects of production, manufacturing, and sales. The companies involved in the process from raw materials to delivery to consumers have a significant amount of management and operational risks that need to be identified and evaluated [3]. Traditional supply chain risk identification and assessment rely on manual labor, which cannot avoid evaluation errors caused by many uncertain factors. Therefore, there is an urgent need for a new risk assessment method for the supply chain risks of NEVs. Many scholars have researched supply chain risk systems in various fields and proposed some identification and evaluation methods [4].

The first step in risk assessment is to identify the risks in SCS, and many experts have made some research progress in this direction [5]. Dhruv et al. proposed a method for exploring business process problems based on simulation data to avoid risk transfer during business operations. This method could effectively identify key risks in business processes and take targeted measures [6]. Sun et al. proposed an algorithm based on financial and supplier operational risks, which can analyze uncontrollable factors in the market, effectively capture risk incentives, and create conditions for in-depth analysis of risk causes. This method could effectively control the hidden risks of SCS in agricultural products [7]. Ge et al. proposed a supply chain risk identification model based on supply demand. This model conducted research from the end of the supply chain through reverse deduction to obtain the optimal solution for risk identification. The

identification model established by this method could effectively analyze the impact of decision preferences on risk [8]. Patibandla et al. proposed a global liquidity index evaluation model. This model could identify risk propagation for important nodes in the overall process and effectively improve the accuracy and reliability of risk identification [9]. Wu et al. proposed a risk identification model based on supplier standard selection. It combined the consideration of both price and quality factors to obtain a more reliable method for risk assessment. This method provided a basic framework for risk identification and guided supply chain risk identification [10].

After effectively identifying risks, the application of Risk Assessment Models (RAMs) to rate risks has received widespread attention. Yu W et al. proposed an NEVs RAM based on the functional-coefficient method, which can achieve supply chain risk warning and provide a quantitative method for risk assessment. This method has been applied to actual enterprise financial risk warnings, achieving high accuracy and feasibility [11]. Ghazal et al. proposed a risk matrix-based analysis and evaluation model and provided diverse and specific risk assessment methods for multi-criteria decision fuzzy problems. This method has the potential to effectively assess the degree of risk in complex environments, thereby providing a reliable decision-making basis for decision-makers [12]. Sari et al. proposed a supply chain risk assessment method based on a coordination mechanism. This method provided a reliable measurement basis for the impact of uncertain factors in the market on risk assessment. Its effectiveness in assessing the supply chain risks of NEVs was highly scientific and accurate [13]. Sardar et al. proposed a logistics RAM based on Matter Element Extension (MEE) to optimize supply chain logistics business processes and reduce transportation costs and personnel work efficiency. This model could address nonlinear influencing factors in logistics transportation processes to obtain reliable risk nodes, providing a reference for risk control in physical transportation processes [14]. Li et al. proposed a Bayesian-based RAM that can effectively address the impact of subjective and objective environmental factors to ensure the accuracy of assessment data. This model could be applied to the SCS of Otis Elevator Company to accurately classify risk levels [15].

In conclusion, a variety of risk identification and assessment techniques have been developed for use in Supply Chain Management (SCM) systems. These have contributed to the advancement of SCM, ensuring its continued healthy development [16-17]. However, the accuracy and feasibility of existing NEV SCS risk assessment methods in meeting personalized supply chain needs and considering various complex factors still need to be improved, as shown in Table 1.

This study proposes a risk assessment method for NEVs supply chain based on Hierarchical Holographic Model (HHM) and MEE model. This method helps identify important risk indicators in SCS, and enables weight calculation and priority ranking of indicators to reduce the scope of risk events. Its innovation lies in the combination of HHM and MEE models, which can identify risks in a framework and conduct systematic risk assessment, providing new ideas for NEVs SCM.

Table 1: Summary of related work.

Authors	Method/Model	Key Contributions	Strengths	Limitations	Improvement of the proposed method
Dhruv et al.	Simulation-based business process problem exploration	Avoid risk transfer phenomena	Accurate identification of key risks	May overlook non-linear factors	Consider more complex factors
Sun et al.	Algorithm based on financial and supplier operational risks	Capture risk triggers	Incorporate uncontrollable market factors	Limited adaptability to personalized needs	Provide more personalized solutions
Ge et al.	Supply-demand-based supply chain risk identification model	Analyze the impact of decision preferences on risk	Reverse derivation for optimal solution	Algorithm may be complex	Simplify the assessment process
Patibandla et al.	Overall liquidity index-based evaluation model	Improve risk identification accuracy	Analysis of critical nodes	Logistics characteristics may not be fully considered	Strengthen logistics risk management
Wu et al.	Supplier standard selection-based risk identification model	Provide a foundational framework	Integrated price and quality considerations	Generality needs improvement	Enhance model flexibility

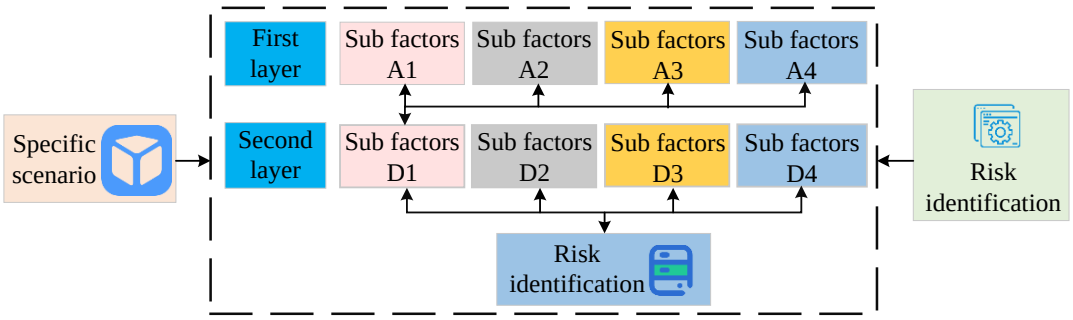


Figure 1: Schematic diagram of HHM risk identification framework.

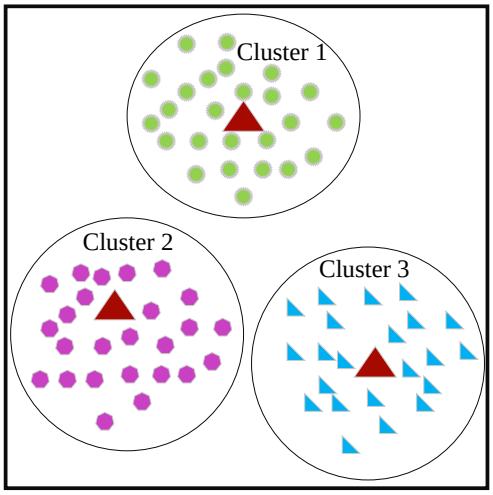


Figure 2: K-means effect diagram.

2 Methods and materials

This section constructs an HHM new energy supply chain risk identification model based on K-means Clustering (K-means), obtains important risk indicators, and further classifies them. Then, an MEE RAM based on a Long Short-Term Memory (LSTM) network is constructed, and the classified indicators are weighted and ranked according to their importance.

2.1 Risk identification model for HHM new energy supply chain based on K-Means

In the process of increasing consumer demand for NEVs, effective management of the NEVs supply chain becomes particularly important [18]. The risk identification of NEVs is the first step in supply chain risk management, which aims to detect risks and classify them for management. The process of NEVs, from raw material procurement to production and sales, is fraught with significant risks. It is therefore essential to implement a systematic and comprehensive method for the effective identification and classification of risks [19]. This study introduces a risk management model based on HHM, aiming to identify global risk factors from different perspectives. Fig.1 is a framework diagram for identifying supply chain risks.

Through the HHM framework, all risks in the process of NEVs from production to sales have been identified, and noteworthy risk factors have been identified. Due to the large number of risk factors involved, these factors need to be further summarized and categorized [20]. This study introduces the K-means algorithm to further mine the supply chain risk data identified by the HHM framework, cluster similar data, and provide important basis for the analysis of risk causes. The principle of K-means for big data is shown in Fig.2.

The first step of HHM is to classify risks from different perspectives, which can be achieved through the K-means algorithm. This algorithm aims to divide the dataset into clusters centered around K by selecting the cluster center K. Before clustering, the data are standardized to ensure that the dimensions of each risk factor data are consistent, as shown in equation (1).

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

In equation (1), z represents the Z-score standardization method. x representing risk factor data. μ is the mean. σ represents standard deviation. After data preprocessing, the process of selection and clustering is continuously repeated to ultimately obtain the best data classification results. The data classification method is shown in equation (2).

$$C_i = \arg\min_j x_i - \mu_j^2 \quad (2)$$

In equation (2), x_i is the training sample of the risk dataset. u_j is a randomly selected cluster center. j is the current number of cluster centers. C_j is the evaluation attribute of the current cluster. To further obtain more representative data classification and obtain the minimum value of the $fitness\left(A[1], A[2], ..., A[n]\right)$ objective function, there is a calculation formula as shown in equation (3).

$$fitness(A[1], A[2], ..., A[n]) = \sum_{1}^{K}\sum_{i}^{N} Dist(x_i^2, C_k^2) \quad (3)$$

In equation (3), K and N are risk datasets selected from two different perspectives. $Dist\left(x_i^2, C_k^2\right)$ is the distance between x_i^2 first level risk indicators and the central point C_k^2. According to the primary indicator attribute C_k^2, the number of risk points x_i^2 under this risk attribute is calculated, and C_k^2 is optimized through the calculation results to converge the objective function to the optimal clustering center [21-22]. By using the K-means algorithm for clustering weights, the dataset is subjected to distance calculation to obtain clustering clusters, as shown in equation (4).

$$\Delta\varepsilon_k = \frac{1}{2m}(w_k \cdot dist(c_k, x))^2 \quad (4)$$

In equation (4), W_k is the weight factor calculated by the algorithm clustering. $\Delta\varepsilon_k$ is the standard deviation for calculating the distance between all data and the center point. Assuming the cluster center attribute target is C_k^3, after k^3 rounds of clustering on the dataset, the distance increment is represented as $\Delta\varepsilon_k$. Following the preliminary risk classification, the second step of the HHM algorithm is initiated. This involves further optimization of the dataset partitioning perspective and the clustering process, with the objective of combining the cross effects between various risk clustering clusters. By combining the K-means algorithm with neural networks, the algorithm learns and clusters data unsupervised, resulting in better clustering centers [24].

The neural network structure can determine the selection of K-means centers through autonomous learning, while avoiding the influence of subjectivity. Therefore, neural networks are used to improve K-means. It takes the shortest distance between the original data and the neurons as the winning distance, and iteratively optimizes and adjusts the weights between all clustering centers and various risk data. The adjustment method is shown in equation (5).

$$y_{ij}(t+1) = y_{ij}(t) + \eta(t, N^0)(x_{kj} - y_{ij}(t)) \quad (5)$$

In equation (5), x_{kj} is the initially selected cluster center point. $y_{ij}(t)$ is the weight of the risk data node at time t. t is the training time of the model. $\eta(t, N)$ is the distance function of the optimal clustering center when the training time reaches t. The distance between each risk data is calculated using the selected best clustering center, as shown in equation (6).

$$d\left(x^0, y^0\right) = \sqrt{\sum_{i^0=1}^{m^0}\left(x_{i^0}{}^0 - y_{i^0}{}^0\right)^2} \quad (6)$$

In equation (6), $x^0 = \left(x_1, x_2, x_3, ..., x_{i^0}{}^0\right)$ and $y^0 = \left(y_1, y_2, y_3, ..., y_{i^0}{}^0\right)$ are the set of cluster centers and the set of data points within the relevant clusters. Among them, i^0 is the dimension in which the current risk data point observation perspective is located. $x_{i^0}{}^0$ and $y_{i^0}{}^0$ are variables in the current dataset dimension. $d\left(a^0, b^0\right)$ is the distance between the data points within the cluster and the cluster center. Different perspectives will lead to different risk factors, and the risk factors from different perspectives are interrelated, intersecting, and overlapping. Therefore, ineffective factors are eliminated through the risk factor identification framework of the HHM.

When the initial dataset has multi-dimensional variables due to different perspectives, its calculation is shown in equation (7).

$$\sum_{X^1} = \frac{1}{k^1}(X^1 - \mu_X)^{T^1}(X^1 - \mu_X) \quad (7)$$

In equation (7), $\sum_{X^1}$ is the covariance matrix of X^1. μ_X is the mean risk point of the X^1 risk dataset. T^1 is the transpose. k^1 is the total number of risk indicators in the dataset. After the eigenvectors of the covariance matrix are rotated, the resulting variance matrix is shown in equation (8).

$$\sum_F = \frac{1}{k^{2T^2}} U(X^2 - \mu_X{}^2)^{T^2} U(X^2 - \mu_X{}^2) = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_m{}^2 \end{pmatrix} \quad (8)$$

In equation (8), U is the rotation matrix of the risk matrix. The coordinate of rotation matrix U is F. $\sum_F$ is the covariance matrix of the rotation matrix F. $\mu_X{}^2$ is the sample mean of the risk dataset. $\sqrt{\sigma}$ is the variance of the risk indicator. k^2 is the total number of datasets. T^2 is transposed. The improved algorithm distance is shown in equation (9).

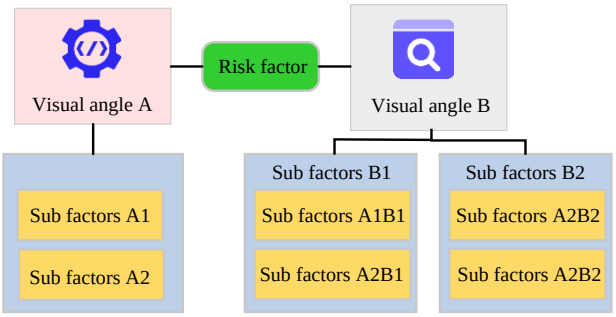


Figure 3: Risk assessment framework.

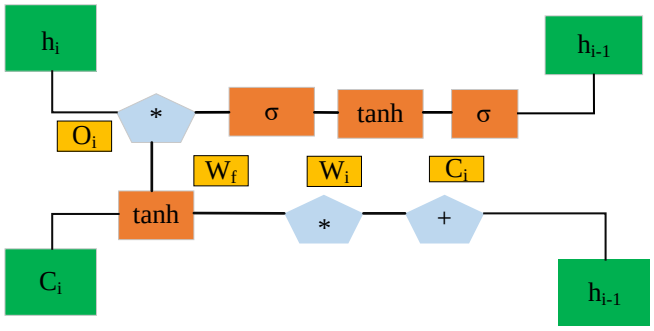


Figure 4: LSTM architecture diagram.

$$d\left(x^5, y^5\right) = \sqrt{\left(x^5 - y^5\right)^{T^5} \sum_{x^5, y^5}^{-1} \left(x^5 - y^5\right)} \quad (9)$$

In equation (9), $x^5 = \left(x_1, x_2, x_3, ..., x_h{}^5\right)$ and $y^5 = \left(y_1, y_2, y_3, ..., y_l{}^5\right)$ are datasets from different perspectives. T^3 is transposed. $\sum_{x^5, y^5}^{-1}$ is the inverse matrix of X^5.

2.2 Construction of MEE RAM combining hybrid weight method

The resolution of important risk identification issues has laid the foundation for the construction of RAMs. The RAM for new energy can effectively judge the degree of risk after identifying it, providing important basis for the formulation of risk solutions. This study proposes the LSTM algorithm, which can avoid the influence of nonlinear factors in the SCM process and has a certain degree of inclusiveness in evaluating missing risk data. Fig.3 shows the risk assessment framework.

In Fig.3, the behavior of the LSTM model is controlled by gates. Its structure mainly consists of three gate gates: forget gate, input gate, and output gate, which can effectively detect the feature sequence of risk data indicators, assign weights to risk indicators, and sort them according to their importance. The LSTM model used in the study consists of two hidden layers, each with 128 hidden units, and is trained using the Adam optimizer. In the data preprocessing stage, the raw data are standardized and the first 80% of the dataset is divided into a training

set and the last 20% as a testing set to address uncertainty in the risk assessment process. Fig.4 shows the structure of LSTM.

Based on the risk identification results of K-means, the classified risk indicators are sorted by their importance to obtain the weights of each indicator. By using gradient descent, the weights and thresholds of the loss function are repeatedly corrected to improve the accuracy of the algorithm, as shown in equation (10).

$$g_t = \frac{w_{t-1} - w_t}{y} \tag{10}$$

In equation (10), weight calculation begins at point $t=1$. w_t and $w_{t\text{-}1}$ are the weight values at times t and t-1. g_t is the gradient. The weight attenuation coefficient of risk assessment indicators is shown in equation (11).

$$\begin{cases} m_t = \varepsilon m_{t-1} + (1-\varepsilon) g_t \\ v_t = \varepsilon v_{t-1} + (1-\varepsilon) g_t^2 \end{cases} \tag{11}$$

In equation (11), m_t and v_t are the first and second momenta. ε is the attenuation coefficient. The calculation steps for weights are as follows: the first is to adaptively adjust the learning rate, and then calculate the difference between the predicted and true values. The loss function is shown in equation (12).

$$\begin{cases} MAE = |G - f(X)| \\ MSE = \sum_n G - f^2(X)) \\ LOG = -\log f(\frac{G}{X}) \end{cases} \tag{12}$$

In equation (12), MAE and MSE are the absolute and mean square values of the loss function. LOG is the logarithmic loss. G is the risk level prediction data output by the function. X is the true data of risk level. The final step in weight calculation is accuracy judgment. If the accuracy does not meet the requirements, it will be recalculated until it converges to the optimal value. To avoid getting stuck in local optima during algorithm training and reduce deep learning time, a topology structure of adding three layers of neural networks to the neural network is used. The establishment of the classical domain is shown in equation (13).

$$R_j = (N_j, C_i^*, V_{ji}) = \begin{bmatrix} N_j, C_1, (a_{j1}, b_{j1}) & \\ C_2, (a_{j2}, b_{j2}) & \\ \cdots & \cdots \\ C_n, (a_{jn}, b_{jn}) & \end{bmatrix} \tag{13}$$

In equation (13), N_j is the risk level. j is the risk indicator to be sorted. C_i^* is the characteristic value of the ranking of various supply chain risk indicators. V_{ji} is the ranking range of risk level characteristic values between indicators j and i. The asymmetric closeness of risk level indicators is shown in equation (14).

$$K_j(N) = 1 - \frac{1}{n(n+1)} \sum_{i=1}^{n} D_{ij} w_i \tag{14}$$

In equation (14), $K_j(N)$ is the closeness between the risk element to be evaluated and the existing risk level. D_{ij} is the distance of the risk indicator. w_i is the comprehensive weight value of each risk indicator. The maximum closeness obtained from the calculation is used as the final output of the MEE RAM, which is the risk assessment indicator of the model. The final risk assessment process is shown in Fig.5.

In Fig.5, the supply chain risk assessment process for NEVs includes risk identification, risk assessment, risk ranking, and final risk indicator output.

3 Results

This section introduces the dataset used for model training and presents the results of risk identification performance testing on the HHM model based on the K-means algorithm.

The results are then compared with the performance of other risk identification algorithms. Subsequently, the LSTM-based MEE RAM is trained on different datasets and its effectiveness is analyzed.

3.1 Performance testing of HHM risk identification model based on K-Means

The experimental platform adopts AMD Ryzen 55600 H with Radeon Graphics, with a main frequency of 3.30 GHz and 16 GB of memory. The dataset is collected from manufacturing, supply, and sales enterprises of NEVs. Data indicators are about the management risks and operational risk factors of the automotive supply chain. To verify the risk identification effectiveness of the K-means-based HHM and its classification performance on risk indicator datasets, three additional clustering algorithms

are used for performance analysis. They include Recent-Frequency-Monetary Clustering (RFM), Particle Swarm Optimization (PSO), and Self-Organizing Map (SOM). The clustering results of each algorithm on the risk indicator datasets of manufacturing, supply, and sales enterprises are shown in Fig.6. In Fig.6 (d), the research algorithm selects cluster centers from three datasets, with data clusters concentrated in [-0.5, 1.5], [2, 4.1], and [0.2,

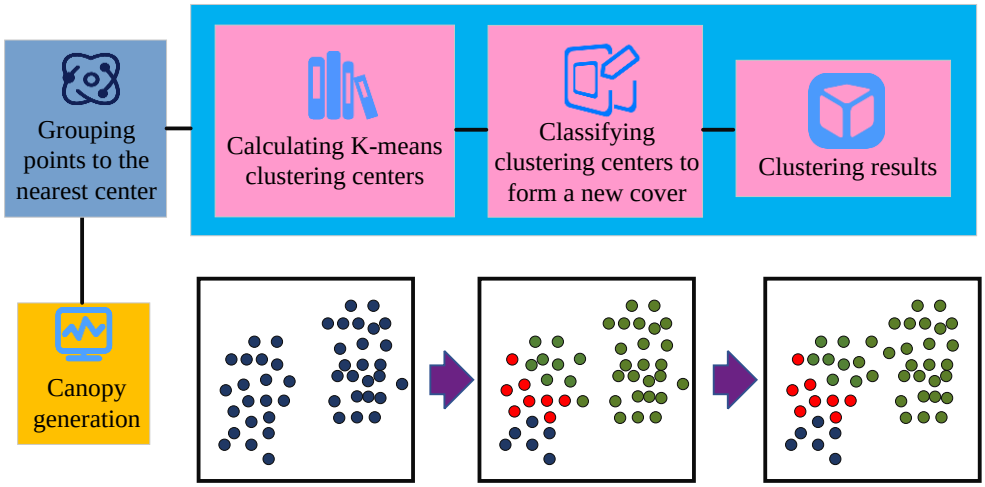


Figure 5: Schematic diagram of RAM structure.

1.8]. The clustering clusters of the comparative algorithm are significantly more dispersed, with an average distance reduction of 1.2%, 1.3%, and 1.5% from each data point to the center point. There are individual data adhesion and excessive dispersion phenomena, with smaller fluctuations and significant differences ($P<0.05$). The research algorithm can effectively cluster different risk assessment indicator datasets and has high reliability and accuracy.

To verify the noise resistance of the proposed HHM risk identification model, the noise processing capability of the K-means algorithm is tested. Dynamic Localized Clustering with Kernel Density Thresholding (DLCKDT), spectral clustering, and the proposed algorithm are tested and compared using 30% and 70% random noise. The Rand index is used as an evaluation indicator, and the results are shown in Fig.7. In Fig.7 (a), when the research algorithm is subjected to a 30% random noise distribution, the Rand index curve consistently remains above the other two clustering algorithms, exhibiting minimal fluctuations, thereby indicating high stability. The Rand index of the DLCKDT and spectral clustering algorithms is, on average, 55% and 41% lower than that of the proposed algorithm. Furthermore, the curve demonstrates a decline following an initial increase, indicating reduced stability with diminished fluctuations and notable discrepancies ($P<0.05$). In Fig.7 (b), at a higher 30% random noise distribution, the Rand indices of all three types show a decreasing trend. However, the Rand curve of the proposed algorithm still performs the best, within the range of [0.65, 0.9], with a fluctuation range 45% smaller than that of the DLCKDT algorithm. The curve of spectral clustering algorithm first increases and then decreases, gradually becoming ineffective. This indicates that the research algorithm has better noise processing capabilities and can maintain high risk identification stability in practical applications.

In NEVs SCM, risk datasets for different business processes are collected, including planning risk, production risk, sales risk, and transportation risk. To verify the accuracy of the research algorithm in risk identification, the algorithm is applied to performance testing and compared with traditional spectral clustering algorithms. The results are shown in Table 2. The research algorithm performs the best in accuracy, recall, and F-measure on different datasets, with each indicator being 9.4%, 8.5%, and 9.6% higher than traditional spectral clustering algorithms, with smaller fluctuations and significant differences ($P<0.05$). The performance test results of the two algorithms remain consistent in each dataset, with a difference of less than 3%. This indicates that the proposed algorithm has high stability, accuracy, and feasibility in practical risk assessment.

3.2 Analysis of the effectiveness of MEE RAM based on LSTM

To verify the learning and practical application effectiveness of the LSTM-based MEE RAM, the model training period is set to 200. The time and effectiveness of risk management are used as

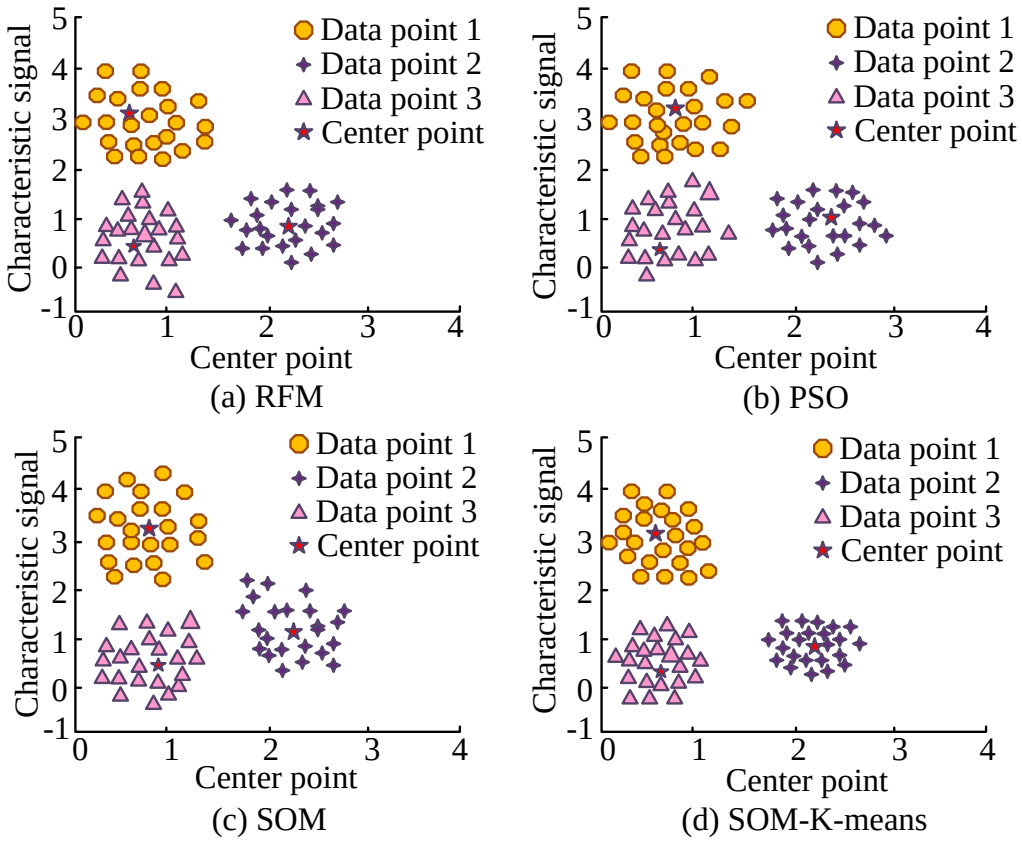
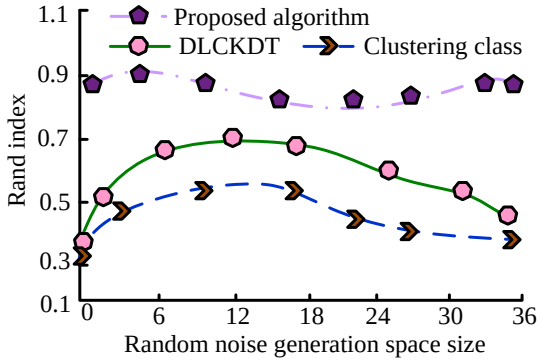
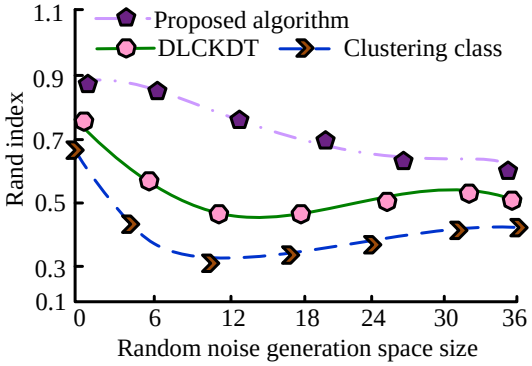


Figure 6: Comparison of the effectiveness of data categorization by algorithms.



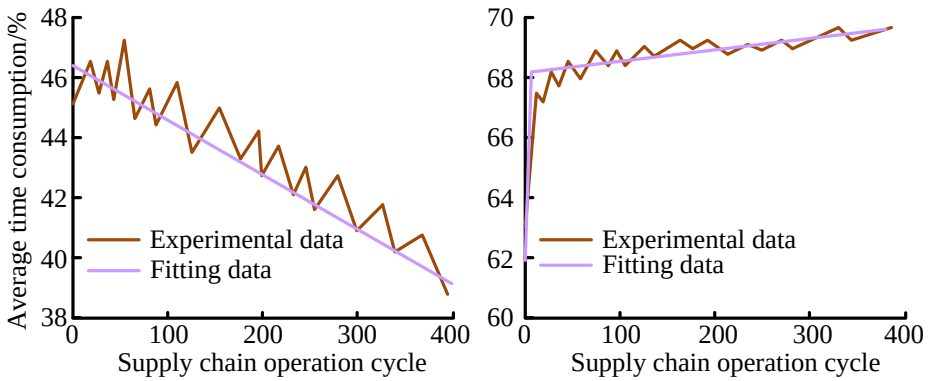
(a) 30% random noise



(b) 70% random noise

Figure 7: Comparison of noise resistance among various algorithms. Table 2: Comparison of testing performance of various algorithms.

Data set	K-means			Spectral clustering algorithm		
	Accuracy	Recall	F-measure	Accuracy	Recall	F-measure
Planning risk	0.74	0.64	0.61	0.64	0.59	0.57
Production risk	0.75	0.68	0.60	0.63	0.60	0.63
Sales risk	0.76	0.67	0.63	0.70	0.61	0.60
Transportation risk	0.81	0.63	0.67	0.73	0.58	0.59



(a) Risk management time consuming (b) Risk management effectiveness

Figure 8: Time and effect curves of the risk of the MEE model.

evaluation indicators for the model, and the results are shown in Fig.8. In Fig.8 (a), as the SCM cycle increases, the RAM shows a gradually decreasing trend in the processing time curve for risks. The management cycle decreases by an average of 23% from 0 to 200, resulting in a decrease in the average processing time of the model, with smaller fluctuations and significant differences ($P<0.05$). In Fig.8 (b), as the SCM cycle increases, the risk management effectiveness of the RAM continues to improve. When the operational management cycle reaches 20, the risk management capability improves the most, increasing by 6%. The proposed model can effectively perform deep learning on risk data, resulting in continuously improving risk management capabilities.

To further validate the research model's ability to handle supply chain risks, the data were collected from January 2022 to December 2023, covering the entire operational cycle of the NEV industry chain. This ensured the timeliness and representativeness of the data. The data come from four primary risk assessment indicators: planning, manufacturing, production, and technology. For each primary indicator, the data contain specific characteristics of multiple risk factors, and the planning indicators include production plan changes and supply plan delay features. Manufacturing indicators include production equipment failure rate and characteristics of raw material supply interruption. Production indicators include production efficiency and product quality qualification rate characteristics. Technical indicators include the cycle of technological innovation and the cost characteristics of technology introduction. Each primary indicator dataset contains 100 data points. Comparing the performance of the autonomous exploration models, as the risk data continue to increase, the processing time curves of each model are displayed in Fig.9. In Fig.9 (a), the risk treatment time curve of the research model shows a continuous downward trend, with a total reduction of 40s in treatment time, and has high stability, with a fluctuation range between [50, 75]. The autonomous exploration model spends an average of 21% more time on risk management than the research model, and the processing time fluctuates greatly, with smaller fluctuations and significant differences ($P<0.05$). Overall, there is no significant improvement. In Figures

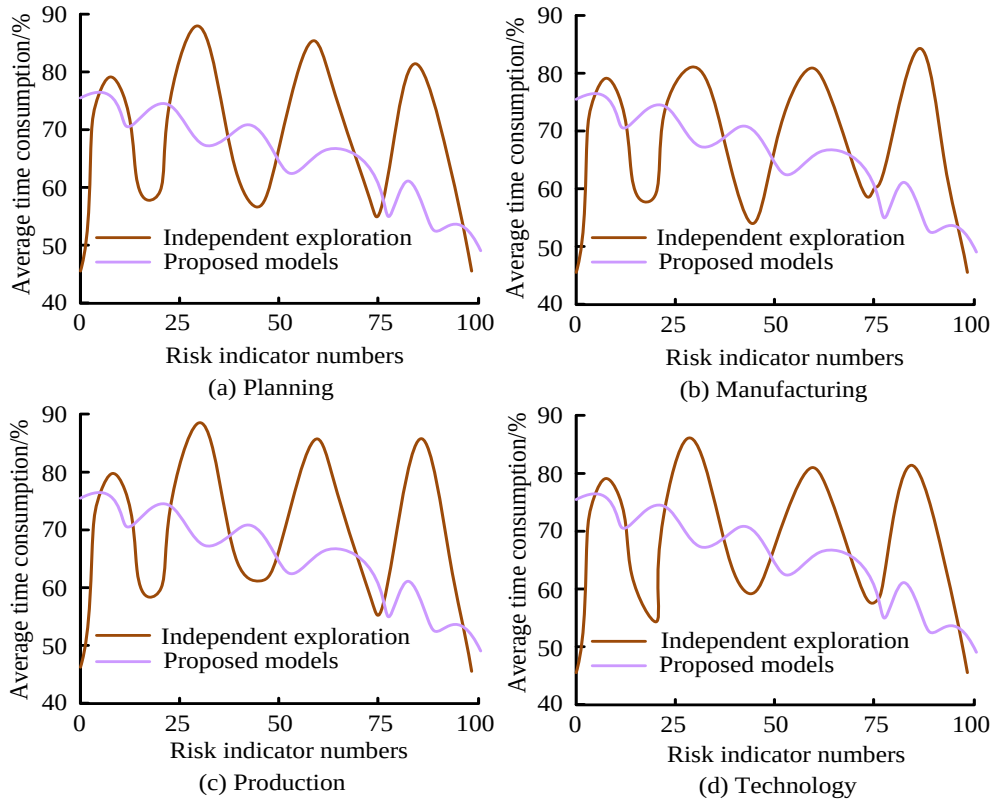


Figure 9: Comparison of risk management effects of different models.

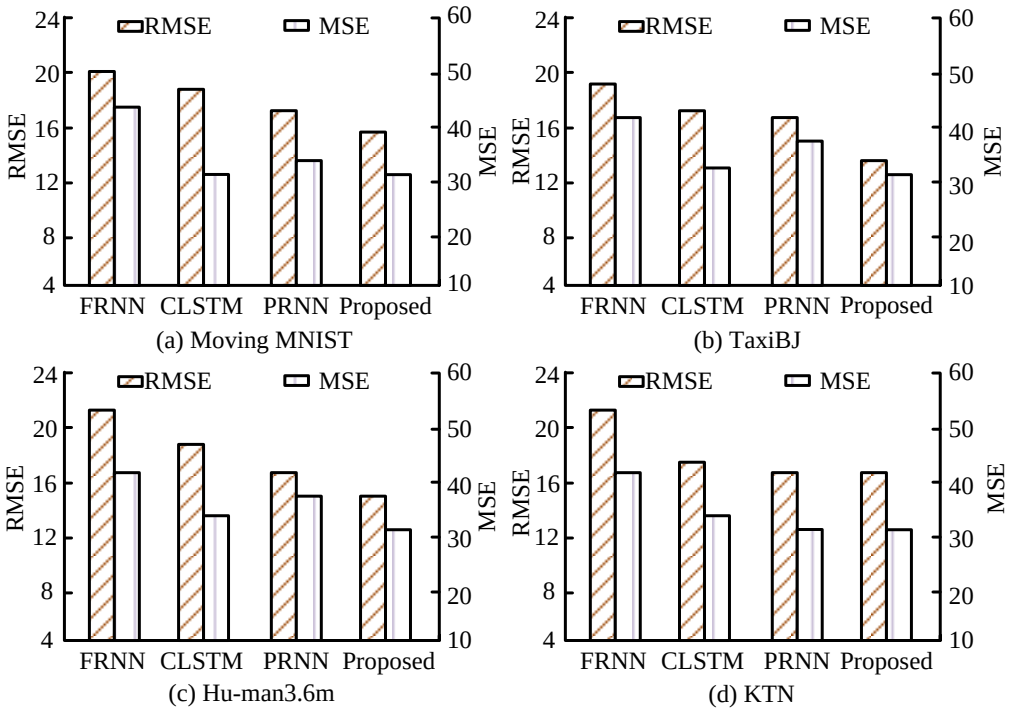
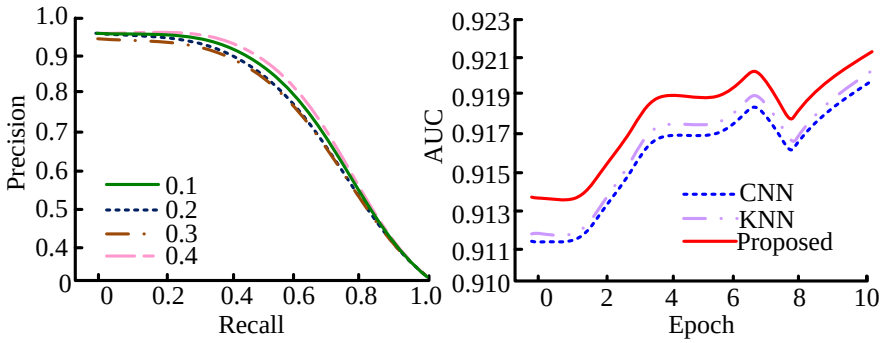


Figure 10: Comparison of detection errors among various algorithms.



(a) PR curve (b) AUC curve

Figure 11: Comparison of prediction performance of various models.

9 (b), (c), and (d), there is no significant change in the processing performance of the two models, and the average data difference is within 5% compared to Fig.9 (a). The research method can effectively enhance the processing capability of supply chain risk assessment indicators and has high stability and accuracy.

To verify the actual evaluation effect of the research model on risk indicators, Root Mean Square Error (RMSE) and Mean Square Error (MSE) are used as evaluation indicators. The research model is applied to the Moving MNIST, TaxiBJ, Hu-man 3.6m, and KTN datasets for testing, and compared and analyzed with models based on other prediction algorithms. The comparison models are Fuzzy Recurrent Neural Network (FRNN), ConvLSTM, and PredRNN algorithms, and the results are shown in Fig.10. Research algorithms have minimum error values in different datasets. Among them, in Fig.10 (a), the RMSE of the research algorithm is 16, which is 3.2%, 2.1%, and 1.5% lower than the FRNN, CLSTM, PRNN, and MIM algorithms. Its MSE is 32, which is 5.8%, 0.7%, and 1.7% lower than the FRNN, ConvLSTM, and PredRNN algorithms, with smaller fluctuations and significant differences ($P<0.05$). In Figures 10 (b), (c), and (d), there is no significant difference in the test data results of each algorithm among the three different datasets, with a difference within 5%. The proposed model has higher accuracy and stability in risk assessment.

To verify the effectiveness and generalization ability of the proposed MEE model, different sampling rates are used to validate its risk assessment performance. Precision-Recall (PR) curve is used as an evaluation indicator. The Area Under the Curve (AUC) during its training process is compared with other CNN and K-Nearest Neighbors (KNN) algorithms, and the results are shown in Fig.11. In Fig.11 (a), the research method has high PR values at different sampling rates, with the best performance observed at sampling rates of 0.1 and 0.4. In Fig.11 (b), as the number of iterations increases, the AUC values of each method show a continuous increasing trend. The average AUC value of the research algorithm is the highest, reaching 0.915, which is 9% and 3% higher than traditional CNN and KNN algorithms, with smaller fluctuations and significant differences ($P<0.05$). Therefore, when studying algorithms for weight calculation and importance ranking of risk indicators, it has high accuracy and reliability.

To verify the universality of the proposed method, two similar algorithms, Random Forest Classifier (RFC) and Naive Bayes Classifier (NBC), are compared, and accuracy, Receiver Operating Characteristic Area under the Curve (AUC-ROC), and processing time are used as evaluation metrics. The results are shown in Table 3. Table 3 shows that the proposed method has the highest accuracy, reaching 97.2%, which is significantly better than the 87.6% and 89.8% of RFC and AUC-ROC, with smaller fluctuations and significant differences ($P<0.05$). The AUC-ROC of the proposed method is 0.921, which is close to 1 and higher than the 0.882 and 0.871 of RFC and AUC-ROC, respectively, with smaller fluctuations and significant differences ($P<0.05$). The proposed method has the shortest processing time, with a reduction of 3.2 and 2.2 on the basis of RFC and AUC-ROC, respectively. The results show that the proposed method exhibits significant superiority in all indicators, is more suitable for handling complex multidimensional data, and has strong universality.

Table 3: Performance comparison of various algorithms.

Algorithm	**Accuracy**	**AUC-ROC**	**Processing Time (s)**
RFC	87.6%	0.882	18.4
AUC-ROC	89.8%	0.871	17.4
Research method	97.2%	0.921	15.2

4 Discussion and conclusion

The SCM of NEVs is crucial for the development of the NEV industry and provides important directions for the transformation and upgrading of the automotive industry. At present, various supply chain RAMs have been proposed. However, in complex and ever-changing environments, many existing methods are difficult to balance the accuracy and real-time performance of the assessment. Therefore, a NEV supply chain risk assessment method based on HHM and MEE models has been proposed, which can effectively globally identify and classify risk factors, and classify and evaluate key risk indicators.

The results indicated that the proposed risk identification model based on HHM had good performance in clustering risk data, with an improvement of 1.2%, 1.3%, and 1.5% compared to the other three clustering algorithms, with significant differences ($P<0.05$). Christie C R et al. proposed a supply chain RAM based on fuzzy comprehensive evaluation. Although the interaction of risk factors was considered, the ability to capture local features was limited when processing complex supply chain data, and the clustering accuracy was poor [23]. The proposed method improved the Rand index by an average of 55% and 41% compared to DLCKDT and spectral clustering algorithms, with smaller fluctuations and significant differences ($P<0.05$). Thesenvitz et al. proposed a supply chain risk identification model based on Analytic Hierarchy Process (AHP), which can analyze the relationship between risk factors at different levels of the supply chain. However, it was sensitive to data noise and had average noise resistance performance [24]. The accuracy of the proposed method was 9.4%, 8.5%, and 9.6% higher than traditional spectral clustering algorithms, with significant differences ($P<0.05$). The average processing time of the proposed MEE model decreased by an average of 23%, and the risk management ability increased by 6%, with significant differences ($P<0.05$). Compared with the independent exploration model, there has been a 21% reduction in time and less volatility in risk management. RMSE and MSE were significantly lower than the other three algorithms by 3.2%, 2.1%, and 1.5% on average ($P<0.05$). Tamsah et al.'s supply chain RAM based on hesitant fuzzy evaluation considers the complexity of the supply chain, but it takes a long time to process large-scale data, making it difficult to meet real-time decision-making requirements [25].

In summary, the proposed NEV supply chain risk assessment method based on HHM and MEE models outperforms existing methods in clustering accuracy, noise resistance, processing efficiency, and error rate. The proposed algorithm has high accuracy and feasibility in risk identification and assessment, and can be effectively used in automotive SCM to obtain a more comprehensive and scientific evaluation method. However, the training process of the proposed algorithm mainly utilizes existing risk factor data. In the development process of NEVs, the scientificity of SCM may also be affected by the emergence of new battery technologies or new market demands. Therefore, predicting new influencing factors and taking preventive measures in advance can be a future development direction.

5 References

[1] Kamar Zekhnini, Anass Cherrafi, Imane Bouhaddou, and Youssef Benghabrit. Jose Arturo Garza-Reyes. Supply chain management 4.0: a literature review and research framework. Benchmarking: An International Journal, 28(2):465-501, 2021. https://doi.org/10.1108/BIJ-04-2020-0156

[2] Matjaž Gams, and Tine Kolenik. Relations between electronics, artificial intelligence and information society through information society rules. Electronics, 10(4):514, 2021. https://doi.org/10.3390/electronics10040514

[3] Nicholas P. Simpson, Katharine J. Mach, Andrew Constable, Jeremy Hess, Ryan Hogarth, and Mark Howden. A framework for complex climate change risk assessment. One Earth, 4(4):489-501, 2021. https://doi.org/10.1016/j.oneear.2021.03.005

[4] Liua B, and Sunb FH. Research on the risk assessment method of PPP project based on the improved matter element model. Scientia Iranica, 27(2):614-624, 2020. https://doi.org/10.24200/SCI.2018.5295.1187

[5] Laura Kaikkonen, Tuuli Parviainen, Mika Rahikainen, Laura Uusitalo, and Annukka Lehikoinen. Bayesian networks in environmental risk assessment: A review. Integrated Environmental Assessment and Management, 17(1):62-78, 2021. https://doi.org/10.1002/ieam.4332

[6] Albert A. Koelmans, Paula E. Redondo-Hasselerharm, Nur Hazimah Mohamed Nor, Vera N. de Ruijter, Svenja M. Mintenig, and Merel Kooi. Risk assessment of microplastic particles. Nature Reviews Materials, 7(2):138-152, 2022. https://doi.org/10.1038/s41578-021-00411-y

[7] Bei Sun, Xudong Liu, Jiayuan Wang, Xuezhe Wei, Hao Yuan, and Haifeng Dai. Short-term performance degradation prediction of a commercial vehicle fuel cell system based on CNN and LSTM hybrid neural network. International Journal of Hydrogen Energy, 48(23):8613-8628, 2023. https://doi.org/10.1016/j.ijhydene.2022.12.005

[8] Qian Ge, Yu Liu, Yinghao Zhao, Yuetian Sun, Lei Zou, Yuxing Chen, and Anqun Pan. Efficient and accurate simrank-based similarity joins: experiments, analysis, and improvement. Proceedings of the VLDB Endowment, 17(4):617-629, 2023. https://doi.org/10.14778/3636218.3636219

[9] R S M Lakshmi Patibandla, and N Veeranjaneyulu. A SimRank based ensemble method for resolving

challenges of partition clustering methods. Journal of Scientific & Industrial Research, 79(4):323-327, 2022. https://doi.org/10.56042/jsir.v79i4.68681

[10] Tianhao Wu, Ji Cheng, Chaorui Zhang, Jianfeng Hou, Gengjian Chen, Zhongyi Huang, Weixi Zhang, Wei Han, and Bo Bai. ClipSim: A GPU-friendly parallel framework for single-source simrank with accuracy guarantee. Proceedings of the ACM on

[11] Weiren Yu, Julie McCann, Chengyuan Zhang, and Hakan Ferhatosmanoglu. Scaling high-quality pairwise link-based similarity retrieval on billion-edge graphs. ACM Transactions on Information Systems (TOIS), 40(4):1-45, 2022. https://doi.org/10.1145/3495209

[12] Taher M. Ghazal. Performances of k-means clustering algorithm with different distance metrics. Management of Data, 1(1):1-26, 2023. https://doi.org/10.1145/3588707
Intelligent Automation & Soft Computing, 30(2):735-742, 2021. https://doi.org/10.32604/iasc.2021.019067

[13] Indah Purnama Sari, Al-Khowarizmi Al-Khowarizmi, and Ismail Hanif Batubara. Cluster analysis using K-Means algorithm and fuzzy C-Means clustering for grou students' abilities in online learning process. Journal of Computer Science. Information Technology and Telecommunication Engineering, 2(1): 139-144. https://doi.org/10.30596/jcositte.v2i1.6504

[14] Tanvir Habib Sardar, and Zahid Ansari. An analysis of distributed document clustering using MapReduce based K-means algorithm. Journal of the Institution of Engineers (India): Series B, 101(6):641-650, 2020. https://doi.org/10.1007/s40031-020-00485-2

[15] Yongyi Li, Zhongqiang Yang, and Kaixu Han. Research on the clustering algorithm of ocean big data based on self-organizing neural network. Computational Intelligence, 36(4):1609-1620, 2020. https://doi.org/10.1111/coin.12299

[16] Mohamed Sakkari, and Mourad Zaied. A convolutional deep self-organizing map feature extraction for machine learning. Multimedia Tools and Applications, 79(27):19451-19470, 2020. https://doi.org/10.1007/s11042-020-08822-9

[17] Xinwang Liu. Simplemkkm: Simple multiple kernel k-means. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4):5174-5186, 2022. https://doi.org/10.1109/TPAMI.2022.3198638

[18] Zengyi Huang, Haotian Zheng, Chen Li, and Chang Che. Application of machine learning-based k-means clustering for financial fraud detection. Academic Journal of Science and Technology, 10(1):33-39, 2024. https://doi.org/10.54097/74414c90

[19] Sani Saminu, Guizhi Xu, Shuai Zhang, Isselmou Ab El Kader, Hajara Abdulkarim Aliyu, Adamu Halilu Jabire, and Yusuf Kola Ahmed. Applications of artificial intelligence in automatic detection of epileptic seizures using EEG signals: A review. Artificial Intelligence and Applications, 1(1):11-25, 2023. https://doi.org/10.47852/bonviewAIA2202297

[20] Srikanta Pal, Ayush Roy, Shivakumara Palaiahnakote, and Umapada Pal. Adapting a swin transformer for license plate number and text detection in drone images. Artificial Intelligence and Applications, 1(3):145-154, 2023. https://doi.org/10.47852/bonviewAIA3202549

[21] Shen Deng, Siyu Mou, and, Huifang Liu. A comprehensive evaluation framework for China's oil import sustainability with a projection pursuit and matter-element extension model. Energy Sources, Part B: Economics, Planning, and Policy, 16(7):650-668, 2021. https://doi.org/10.1080/15567249.2021.1937402

[22] Jingwei Too, and Abdul Rahim Abdullah. A new and fast rival genetic algorithm for feature selection. The Journal of Supercomputing, 77(3):2844-2874, 2021. https://doi.org/10.1007/s11227-020-03378-9

[23] Christopher R. ChristieLuke E. K. Achenie, and Oluwafemi B. Ayeni. A model-based approach to diagnosing hypercalcemia. Industrial & Engineering Chemistry Research, 62(5):2263-2274, 2023. https://doi.org/10.1021/acs.iecr.3c00051

[24] Jodi Thesenvitz, Shelby Corley, Lana Solberg, and Chris Carvalho. Home health monitoring during the COVID pandemic: Results from a feasibility study in Alberta primary care. Healthcare Management Forum, 35(1):29-34, 2022. https://doi.org/10.1177/08404704211041969

[25] Hasmin Tamsah, and Yusriadi Yusriadi. Quality of agricultural extension on productivity of farmers: Human capital perspective. Uncertain Supply Chain Management, 10(2):625-636, 2022. https://doi.org/10.5267/j.uscm.2021.11.003