Research on the Development of Modern Design Through Data Mining Technology

Xiangyun Meng
School of digital media and art design, Nanyang Institute of Technology, Henan, 473004, China
E-mail: yunyunmm2023@126.com

Keywords: statistical density, logistics, fuzzy logic, data mining, path selection, Monte-Carlo

Received: September 30, 2023

Logistics operations heavily rely on efficient route planning and optimization to ensure the smooth flow of goods and services. To enhance these processes, the simulation of logistics frequent path data mining based on statistical density offers valuable insights. By analyzing vast amounts of historical transportation data, statistical density-based methods can identify frequent paths and patterns in logistics networks. With the data mining process, the statistical density of the logistics in China is computed. The model uses the Fuzzy Associative Monte Carlo (FAMC). The proposed FAMC model estimates the associative rules in the fuzzy model for the computation of the frequent pattern for the estimation of the logistics in the data mining process. Through FAMC model statistical density is computed with an estimation of the logistics path to compute the statistical density model. The logistics path routes are estimated based on the computation of the statistical density for the computation of the data mining-based approach in logistics management. The proposed FAMC model effectively computes the path in the logistics in China with a significant density analysis in a statistical manner.

Povzetek: V članku je opisano sodobno oblikovanje z uporabo rudarjenja podatkov, osredotočeno na optimizacijo logističnih poti preko statistične gostote in Fuzzy Associative Monte Carlo modela.

1 Introduction

In the world of logistics, the seamless flow of goods and services is crucial to maintaining efficient operations. This is where route planning and optimization play a pivotal role. The successful transportation and distribution of products heavily rely on strategically mapping out the most effective routes, considering factors such as distance, time, traffic, and costs [1]. By leveraging advanced technologies and algorithms, logistics companies can streamline their processes, reduce delivery times, minimize fuel consumption, and ultimately enhance customer satisfaction. Efficient route planning not only boosts operational productivity but also contributes significantly to overall sustainability and profitability in the logistics industry [2]. In the realm of logistics operations, efficient route planning and optimization are fundamental to ensuring the smooth and cost-effective movement of goods and services. To achieve this, logistics companies often employ advanced techniques, such as fuzzy logic, to handle the complexity and uncertainty inherent in real-world transportation scenarios [3]. Fuzzy logic is a mathematical approach that deals with imprecision and ambiguity, enabling decision-making based on approximate reasoning rather than strict binary outcomes. Fuzzy logic allows logistics operations to make more informed and flexible decisions by capturing and quantifying uncertainties and vagueness in real-world logistics scenarios [4]. By considering multiple factors simultaneously and assigning linguistic values to represent their varying impact, the fuzzy-based approach helps logistics companies optimize routes, minimize risks, and improve overall operational efficiency, resulting in enhanced customer satisfaction and reduced costs [5]. Logistics frequent path data mining based on statistical density is a powerful analytical technique used in the field of logistics to uncover valuable insights from vast amounts of transportation data [6]. This approach involves identifying and analyzing frequently traversed routes or paths taken by vehicles, shipments, or personnel within the logistics network. With applying statistical density analysis, the method emphasizes patterns that occur more frequently and with higher significance, allowing for a deeper understanding of the most common and critical routes [7]. The process begins with collecting and preprocessing large-scale logistics data, which can include GPS records, shipping manifests, and vehicle movement logs. Using statistical techniques like kernel density estimation or density-based clustering, the algorithm calculates the density distribution of paths, highlighting regions with higher concentrations of movement [8]. This enables the identification of frequently utilized paths, popular shipping corridors, and regions with heavy traffic. The insights gained from logistics frequent path data mining offer numerous benefits to logistics companies. For instance, it helps optimize route planning, as managers can allocate resources more efficiently and anticipate potential bottlenecks or congestion points [9]. Additionally, this analysis can aid in the identification of optimal warehouse locations or the adjustment of delivery schedules to streamline operations. Moreover, by identifying popular routes, logistics providers can improve service quality, reduce delivery times, and enhance customer satisfaction [10]. Logistics frequent path data mining based on statistical density is a sophisticated data analysis approach that plays a crucial role in modern logistics operations. With the exponential growth of data generated by GPS devices, tracking systems, and other sensors, logistics companies now have access to vast amounts of information about the movement of goods and vehicles within their networks [11]. However, extracting valuable insights from this massive data pool can be challenging due to its complexity and volume.

Statistical density analysis is a key technique used in this process. It involves creating density maps or heatmaps that visually represent the concentration of movement along different routes or paths [12]. With applying mathematical models like kernel density estimation, the algorithm calculates the intensity of movement at various points in the logistics network, highlighting areas with the highest traffic or activities. Density-based clustering further groups together similar movement patterns, identifying frequently used paths and common transportation routes [13]. One significant advantage of this approach is its ability to reveal hidden patterns and trends in logistics data. It allows logistics managers to identify recurring transportation routes and popular corridors, which can have significant implications for optimizing delivery schedules and resource allocation [14]. Additionally, statistical density analysis helps in assessing transportation network efficiency and identifying potential areas for improvement. By pinpointing regions with heavy traffic or congestion, logistics companies can take proactive measures to alleviate bottlenecks, optimize traffic flow, and reduce delivery delays [15]. This proactive approach enhances operational efficiency and customer satisfaction, as deliveries are more reliable and timelier.

Logistics frequent path data mining based on statistical density can assist in making strategic decisions about warehouse locations and distribution center placements [16]. With understanding the density of movement in different regions, businesses can strategically position warehouses in areas with high demand, reducing the distance and time required for goods to reach customers [17]. The insights gained from this data mining technique can also be used for predictive analytics. By analyzing historical movement patterns and density trends, logistics companies can forecast future transportation demands, anticipate peak periods, and plan accordingly to meet customer expectations during busy times. Logistics frequent path data mining based on statistical density is a powerful tool that empowers logistics companies to make data-driven decisions [18]. Through analyzing vast amounts of transportation data, identifying frequently used routes, and understanding movement patterns, businesses can streamline their operations, improve resource allocation, and enhance customer service, ultimately gaining a competitive edge in the dynamic and demanding logistics industry.

The paper makes several significant contributions to the field of logistics planning and optimization:

1. The paper introduces a novel approach for logistics planning by combining fuzzy logic and Monte Carlo simulations. The FAMC model effectively addresses uncertainties in traffic congestion and demand forecasts, providing more robust and reliable estimates for delivery times, transportation costs, and resource utilization. This innovative model fills a gap in traditional logistics planning methods that often struggle to handle imprecise and uncertain data.
2. With integrating fuzzy logic into the Monte Carlo simulations, the FAMC model offers a powerful tool to manage uncertainty in logistics operations. Traffic conditions and demand forecasts can be highly volatile, especially in urban areas, and the FAMC model's ability to handle fuzzy data allows logistics planners to make informed decisions despite uncertain conditions.
3. The paper emphasizes the environmental impact of logistics operations by incorporating CO2 emissions data into the FAMC model. Logistics companies can now evaluate the environmental footprint of different routes and make informed decisions to minimize CO2 emissions and promote sustainable logistics practices.
4. The inclusion of customer satisfaction data in the FAMC model adds a crucial dimension to logistics planning. By quantifying customer satisfaction levels for various routes, logistics companies can prioritize customer-centric approaches and enhance overall service quality.
5. The FAMC model enables cost optimization by estimating transportation costs and identifying cost-saving opportunities. Logistics planners can use this data to select more cost-effective routes, allocate resources efficiently, and optimize overall logistics expenses.
6. The FAMC model's adaptability and scalability make it suitable for a wide range of logistics scenarios, from small-scale local deliveries to large-scale global supply chains. The model's ability to handle complex and dynamic logistics networks ensures its applicability in real-world logistics planning.
7. The paper emphasizes the importance of data-driven decision-making in logistics planning. By relying on rigorous simulations and analyses, the FAMC model provides valuable insights that empower logistics planners to make informed, data-based choices.

2 Related works

Logistics frequent path data mining can aid in risk assessment and mitigation. By analyzing density patterns, logistics managers can identify high-risk areas prone to congestion, accidents, or other disruptions. With this knowledge, they can implement contingency plans and alternative routes to ensure the smooth flow of goods even during unexpected events. Additionally, by understanding which routes are less frequently used or underutilized, companies can potentially explore new markets and expand their operations. the integration of logistics frequent path data mining with other technologies, such as Internet of Things (IoT) devices and artificial intelligence, opens up new possibilities. IoT sensors on vehicles and shipments can provide real-time data, which, when combined with statistical density analysis, enables continuous route optimization, predictive maintenance, and proactive problem-solving. In [19] introduces a novel approach using a graph-based mixture density network to estimate the distribution of travel times for packages in logistics and transportation scenarios. The method combines graph-based representations and mixture density modeling, offering potential improvements in delivery time estimation and logistics planning. In [20] employ multivariate statistical and data mining techniques to identify biomarkers related to sensorineural hearing loss, tinnitus, and vestibular dysfunction. This research has implications in the field of healthcare and can lead to improved diagnostics and treatment approaches for these conditions. In [21] introduces an innovative method for landslide susceptibility modeling. The study integrates machine learning feature transformation techniques to enhance the accuracy of landslide susceptibility models, which can aid in disaster risk assessment and mitigation efforts. In [22] presents a strategic decision support system for urban logistics operations, focusing on sustainable transport. This system can help logistics companies optimize their operations, reduce environmental impacts, and contribute to more efficient urban mobility.

In [23] provides a comprehensive review of how data mining is applied to production planning and scheduling, particularly within the context of cyber-physical systems. This review can serve as a valuable resource for researchers and practitioners seeking to implement data-driven approaches in manufacturing and production industries. In [24] applies logistic regression to predict diabetic foot ulcers in diabetic patients, with high-density lipoprotein (HDL) cholesterol as a negative predictor. This study contributes to the field of medical research and offers insights into predicting and preventing diabetic foot ulcers. In an improved cluster analysis approach for detecting network anomalies, which is vital in cybersecurity and maintaining network integrity. The interpretability of data mining techniques for spatial modeling of water erosion. The study adopts game theory to assess the reliability and transparency of data-driven water erosion models, offering valuable insights for environmental researchers and policymakers. A large-scale comparison of AI and data mining techniques in simulating reservoir releases. This research can contribute to more accurate and efficient reservoir management and water resource planning. Table 1 provides the overall summary of the literature focused on the logistics.

Table 1: Literature summary

Reference	Method	Findings	Limitations
[19]	Graph-based mixture density network for travel time estimation in logistics	Improved delivery time estimation and logistics planning through graph-based representations and mixture density modelling	Potential limitations in the scalability of the proposed method to large-scale logistics operations; real-world implementation challenges may arise.
[20]	Multivariate statistical and data mining for identifying biomarkers related to hearing loss	Identification of biomarkers related to sensorineural hearing loss, tinnitus, and vestibular dysfunction	Generalization of findings to a broader population may be limited; further validation and clinical studies may be necessary.
[21]	Integration of machine learning feature transformation in landslide susceptibility modeling	Enhanced accuracy of landslide susceptibility models for improved disaster risk assessment	Applicability of the method to diverse geographical and geological conditions needs further exploration; model sensitivity to input data quality.
[22]	Strategic decision support system for urban logistics with a focus on sustainable transport	Optimization of urban logistics operations with reduced environmental impacts	Integration challenges with existing logistics systems; potential resistance to adopting sustainable transport methods.
[23]	Comprehensive review of data mining in production planning and scheduling	Valuable insights for implementing data-driven approaches in manufacturing within the context of cyber-physical systems	Generalization of findings to specific industries may require further research; rapid advancements in technology may impact the relevance of the review over time.
[24]	Application of logistic regression to predict diabetic foot ulcers	Prediction of diabetic foot ulcers with HDL cholesterol as a negative predictor	Generalization of findings to diverse diabetic populations; clinical applicability and potential biases in the dataset.
[25]	Improved cluster analysis	Enhanced approach for	Generalization of the method to

	for detecting network anomalies in cybersecurity	detecting network anomalies, crucial for maintaining network integrity	different network architectures; adaptability to evolving cybersecurity threats.
[26]	Evaluation of interpretability of data mining techniques for spatial modeling of water erosion	Adoption of game theory to assess reliability and transparency of water erosion models	Generalization of findings to diverse environmental conditions; potential limitations in the applicability of game theory to different modeling scenarios.
[27]	Large-scale comparison of AI and data mining techniques in simulating reservoir releases	Contribution to more accurate reservoir management and water resource planning	Generalization of findings to diverse reservoir systems; consideration of external factors influencing reservoir releases.

The literature presented encompasses a wide range of studies that utilize data mining, machine learning, and statistical analysis techniques to tackle complex challenges in various fields. Researchers have introduced novel approaches like GMDNet for estimating multimodal travel time distributions, providing valuable insights for logistics and transportation planning. Medical research explores the application of multivariate statistical and data mining analyses to identify biomarkers of hearing loss and other related conditions, potentially enhancing diagnostics and treatment strategies. In geosciences, an integrated method incorporating machine learning feature transformation improves landslide susceptibility modeling, aiding in disaster risk assessment and mitigation. Moreover, studies emphasize sustainable urban logistics with decision support systems and highlight the importance of data mining in production planning and scheduling for manufacturing industries. Predictive models, like logistic regression, play a crucial role in medical fields, such as predicting diabetic foot ulcers, aiding early prevention and management. Additionally, cybersecurity research focuses on network anomaly detection to ensure network integrity, while environmental studies assess the interpretability of data mining techniques for water erosion modeling. Lastly, a large-scale comparison of AI and data mining techniques in reservoir management promises improved water resource planning and reservoir releases. These diverse investigations demonstrate the growing significance of data-driven approaches in driving innovation, making informed decisions, and addressing critical challenges across various domains.

3 Fuzzy associative monte carlo (FAMC) for logistics planning

To fill in the gaps mentioned above, the impact of the training stream's characteristics, including the degree of imbalance, length at the time *t*, drift types (CI, CD, and OCI-CD), and the state of imbalance (static and dynamic) on state-of-the art adaptive and non-adaptive learners used for minority class prediction, is explored.

This study explores various static and dynamic imbalanced streams with gradual and abrupt drift levels. This work also aims to answer the following research questions:

RQ1. Does the length of the stream with respect to the imbalance ratio at the current time *t* impact the online learner's performance?

RQ2. Is the degree of imbalance or CD whose impact is critical on minority class performance degradation?

RQ3. Is the impact of OCI-CD more adverse than individual *p(y)* or *p(y/x)* drifts on the learner's performance?

RQ4. To what extent does online SVM cope with OCI-CD, *p(y)*, and *p(y/x)* drifts compared to other online learners?

The case of combined *p(x/y)*, *p(y/x)* drift is not considered in the scope of this study, as the impact is only *p(x/y)* due to the change in the likelihood of the concept [3].

The following is the structure of the paper. The related work is shown in Section 4. Section 5 provides background for the problem-related methods. Section 6 discusses the study design, while section 7 discusses the experiments performed on the synthetic data. Section 8 discusses the validity of the observations on real-world data, and section 9 discusses the results obtained in greater depth. Section 10 brings this paper to a close.

Fuzzy Associative Monte Carlo (FAMC) is an innovative approach used in logistics planning to handle uncertainties and complexities inherent in real-world transportation and supply chain operations. FAMC combines three powerful techniques: fuzzy logic, associative memory, and Monte Carlo simulation, to create a robust decision-making framework. Fuzzy logic allows for handling imprecise and vague information, making it suitable for modeling uncertain factors such as delivery times, traffic conditions, and demand fluctuations. Associative memory enables the system to learn from historical data and identify patterns, which can aid in predicting future events and optimizing logistics strategies. Monte Carlo simulation complements the fuzzy and associative components by generating random samples of uncertain variables, facilitating the exploration of various scenarios and their associated probabilities.

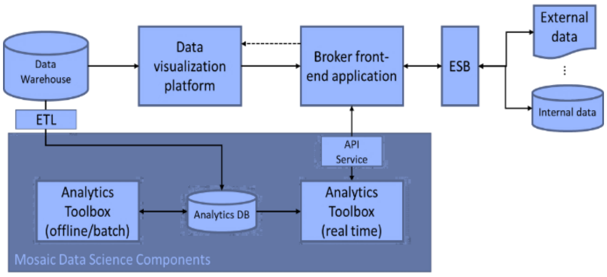

Figure 1: Data mining for the logistics

With integrating these techniques, FAMC empowers logistics planners to make more informed and flexible decisions. It can assess the performance of different routes, transportation modes, and inventory levels while considering multiple factors simultaneously as shown in figure 1. This not only enhances the efficiency of logistics operations but also contributes to cost reduction, improved customer service, and increased overall resilience in the face of unpredictable events. FAMC represents a significant advancement in logistics planning, enabling companies to navigate the complexities of modern supply chains and achieve optimal outcomes in an uncertain and dynamic environment. Fuzzy logic is used to handle imprecise and uncertain information in logistics planning. It involves defining linguistic variables and fuzzy sets to represent the degree of membership of elements in each set. The fuzzy logic equations can model relationships and rules between variables. Consider a linguistic variable "Traffic Congestion" with fuzzy sets "Low," "Moderate," and "High," denoted as TC(L), TC(M), and TC(H) respectively. The membership functions for each set could be defined as in equation (1) – (3):

$$TC(L): \mu\_TC(L)(x) = 1 - (x - a)/(b - a) \quad (1)$$

$$TC(M): \mu\_TC(M)(x) = (x - a)/(c - a), if\ a \leq x \leq c \quad (2)$$

$$TC(H): \mu\_TC(H)(x) = (x - c)/(d - c), if\ c \leq x \leq d \quad (3)$$

Here, 'x' represents the traffic congestion level, and 'a,' 'b,' 'c,' and 'd' are predetermined thresholds that define the boundaries of the fuzzy sets. The membership functions describe the degree of membership of 'x' in each fuzzy set. Associative memory is used to learn from historical logistics data and identify patterns to support decision-making. This involves using techniques like neural networks or pattern recognition algorithms to establish relationships between input and output variables. A neural network can be trained using historical data on transportation time, distance, and other factors to predict the expected delivery time. The neural network equations involve the activation functions, weights, biases, and the backpropagation algorithm to adjust the network's parameters during the learning process. During the encoding phase, historical data is represented as fuzzy sets with membership functions for each variable. The following fuzzy sets for Traffic Congestion (TC) defined in equation (4) – (6) and Delivery Time (DT) stated in equation (7) – (9):

$$TC(L): \mu\_TC(L)(x) = 1 - (x - a)/(b - a), if\ x \leq b \quad (4)$$

$$TC(M): \mu\_TC(M)(x) = (x - a)/(c - a), if\ a \leq x \leq c \quad (5)$$

$$TC(H): \mu\_TC(H)(x) = (x - c)/(d - c), if\ x \geq c \quad (6)$$

$$DT(L): \mu\_DT(L)(x) = 1 - (x - p)/(q - p), if\ x \leq q \quad (7)$$

$$DT(M): \mu\_DT(M)(x) = (x - p)/(r - p), if\ p \leq x \leq r \quad (8)$$

$$DT(H): \mu\_DT(H)(x) = (x - r)/(s - r), if\ x \geq r \quad (9)$$

Here, 'x' represents the corresponding values of Traffic Congestion and Delivery Time, and 'a', 'b', 'c', 'd', 'p', 'q', 'r', 's' are predetermined thresholds that define the boundaries of the fuzzy sets. During the retrieval phase, the system compares the current input data with the encoded patterns to retrieve the most similar historical patterns. Let's assume the current traffic congestion level is 'x_TC' and the estimated delivery time is 'x_DT'. The degree of membership for each linguistic term using the fuzzy membership functions stated in equation (10) and (11):

$$\mu\_TC(L)(x\_TC), \mu\_TC(M)(x\_TC), \mu\_TC(H)(x\_TC) \quad (10)$$

$$\mu\_DT(L)(x\_DT), \mu\_DT(M)(x\_DT), \mu\_DT(H)(x\_DT) \quad (11)$$

To find the similarity between the current input and the historical patterns, the use fuzzy similarity measures, such as the fuzzy intersection or fuzzy distance. In this the fuzzy intersection presented in equation (12) – (17):

$$Fuzzy\ Intersection\ (L, x\_TC) = min(\mu\_TC(L)(x\_TC), \mu\_TC(L)(x\_TC)) \quad (12)$$

$$Fuzzy\ Intersection\ (M, x\_TC) = min(\mu\_TC(M)(x\_TC), \mu\_TC(M)(x\_TC)) \quad (13)$$

$$Fuzzy\ Intersection\ (H, x\_TC) = min(\mu\_TC(H)(x\_TC), \mu\_TC(H)(x\_TC)) \quad (14)$$

$$Fuzzy\ Intersection\ (L, x\_DT) = min(\mu\_DT(L)(x\_DT), \mu\_DT(L)(x\_DT)) \quad (15)$$

$$Fuzzy\ Intersection\ (M, x\_DT) = min(\mu\_DT(M)(x\_DT), \mu\_DT(M)(x\_DT)) \quad (16)$$

$$Fuzzy\ Intersection\ (H, x\_DT) = min(\mu\_DT(H)(x\_DT), \mu\_DT(H)(x\_DT)) \quad (17)$$

The fuzzy intersection values for each linguistic term from both variables using the associative memory to retrieve the historical patterns. The retrieved patterns will have associated data, such as past delivery times, routes, and transportation modes. A historical pattern with TC = "Low" and DT = "Fast," the fuzzy associative memory will recognize the current input with similar fuzzy intersection values and retrieve that historical pattern. The retrieved historical patterns can then be used for decision-making, predicting expected delivery times, or simulating different logistics scenarios through Fuzzy Associative Monte Carlo (FAMC) for logistics planning.

3.1 Monte carlo simulation

Monte Carlo simulation is used to explore various scenarios and estimate the probabilities of different outcomes in logistics planning. This involves generating random samples for uncertain variables and running simulations multiple times to obtain statistical results. To estimate the probability of on-time delivery for a specific route, Monte Carlo simulation can be applied. It involves generating random samples for uncertain variables like traffic conditions, weather, and demand levels, and then simulating the delivery process for each sample. By repeating the simulation thousands of times, the probability of on-time delivery can be approximated. The probability distributions for each uncertain variable. Assume that both TC and DF follow normal distributions presented in equation (18) and (19):

$$TC \sim N(\mu\_TC, \sigma\_TC^2)\ (Traffic\ Congestion) \quad (18)$$
$$DF \sim N(\mu\_DF, \sigma\_DF^2)\ (Demand\ Forecast) \quad (19)$$

Here, μ_TC and σ_TC represent the mean and standard deviation of the traffic congestion, while μ_DF and σ_DF represent the mean and standard deviation of the demand forecast. Using the defined probability distributions, generate random samples for each uncertain variable. For Monte Carlo simulation, 'n' random samples are typically generated for each variable. These samples represent different values of traffic congestion and demand forecasts, which simulate various real-world scenarios.
Assume generate 'n' random samples for each variable:

Random Samples for Traffic Congestion: $x\_TC1, x\_TC2, \ldots, x\_TCn$

Random Samples for Demand Forecast: $x\_DF1, x\_DF2, \ldots, x\_DFn$

The fuzzy associative memory to retrieve historical patterns based on the generated random samples for traffic congestion and demand forecast. The fuzzy logic and associative memory help identify similar patterns from past data that match the current inputs. With the fuzzy intersection values as explained in the previous response, the retrieved historical patterns corresponding to the generated random samples of traffic congestion and demand forecast. With retrieved historical patterns, the fuzzy logic and associative memory-based decisions to simulate the logistics planning process. For each combination of traffic congestion and demand forecast, the system estimates delivery times, optimizes routes, and makes resource allocation decisions based on historical patterns and fuzzy rules. After running the simulations for all combinations of random samples. This analysis includes calculating performance metrics such as on-time delivery rates, transportation costs, or resource utilization for each scenario.

3.2 Optimize decisions statistical density

Using the probabilities and outcomes obtained from the Monte Carlo simulation, logistics planners can make informed decisions. They can identify high-risk scenarios, guide resource allocation, and select optimal routes based on the likelihood of success. With incorporating Monte Carlo simulation into FAMC enhances logistics planning by accounting for uncertainties and providing robust decision-making based on simulated scenarios. In the FAMC model, fuzzy logic is employed to represent uncertain and imprecise information related to logistics variables. A linguistic variable "Traffic Congestion" (TC) with fuzzy sets "Low," "Moderate," and "High," denoted as TC(L), TC(M), and TC(H), the fuzzy membership functions can be expressed in equation (20) – (22):

$$\mu\_TC(L)(x) = 1 - (x - a)/(b - a), if\ x \leq b \quad (20)$$
$$\mu\_TC(M)(x) = (x - a)/(c - a), if\ a \leq x \leq c \quad (21)$$
$$\mu\_TC(H)(x) = (x - c)/(d - c), if\ x \geq c \quad (22)$$

Here, 'x' represents the value of traffic congestion, and 'a', 'b', and 'c' are predetermined thresholds defining the boundaries of the fuzzy sets. The FAMC model utilizes associative memory to retrieve historical patterns based on fuzzy inputs and linguistic rules. By comparing the current traffic congestion value 'x_TC' with the fuzzy membership functions, the model retrieves historical patterns related to similar traffic conditions. The FAMC model incorporates Monte Carlo simulation to estimate statistical density. For a logistics path, the model generates multiple random samples of uncertain variables, such as traffic conditions, demand forecasts, or delivery times, based on their assigned probability distributions. Consider 'n' random samples of traffic congestion for the logistics path as x_TC1, x_TC2, ..., x_TCn. Using these samples and historical data retrieved through associative memory, the model simulates multiple logistics scenarios. For each random sample of traffic congestion 'x_TC', the model computes the associated delivery time 'x_DT' using the fuzzy logic and historical patterns. Finally, the FAMC model computes the statistical density model by estimating the density of delivery times across the multiple logistics scenarios. This can be done using statistical techniques like kernel density estimation or other density-based approaches.

Algorithm 1: FAMC for density estimation

```
# Define Fuzzy Sets for Traffic Congestion (TC) and Delivery Time (DT)
TC_sets = {"Low", "Moderate", "High"}
DT_sets = {"Fast", "Average", "Slow"}
# Define Membership Functions for Fuzzy Sets
def membership_function(x, a, b, c):
    if x <= a:
        return 0
    elif a < x <= b:
        return (x - a) / (b - a)
    elif b < x <= c:
        return (c - x) / (c - b)
    else:
        return 0
# Define Fuzzy Rules
rules = {
    ("Low", "Fast"): "Low",
    ("Low", "Average"): "Moderate",
    ("Low", "Slow"): "High",
    ("Moderate", "Fast"): "Low",
    ("Moderate", "Average"): "Moderate",
    ("Moderate", "Slow"): "High",
    ("High", "Fast"): "Moderate",
    ("High", "Average"): "High",
    ("High", "Slow"): "High"
}
# Define Associative Memory to Store Historical Patterns
associative_memory = {}
# Define Probability Distributions for Uncertain Variables
# (e.g., Traffic Congestion and Demand Forecast)
def probability_distribution():
    # Define probability distributions based on historical data
# Define Monte Carlo Simulation
def monte_carlo_simulation():
    # Generate 'n' random samples for uncertain variables
```

```
based on their probability distributions
# Main FAMC Algorithm
def famc_algorithm():
   # 1. Generate random samples for uncertain variables using Monte Carlo simulation
   monte_carlo_samples = monte_carlo_simulation()
   # 2. Use associative memory to retrieve historical patterns for each random sample
   for sample in monte_carlo_samples:
      traffic_congestion = sample['traffic_congestion']
      demand_forecast = sample['demand_forecast']
      retrieved_pattern = associative_memory.get((traffic_congestion, demand_forecast), None)
      # 3. Use Fuzzy Logic and Fuzzy Rules to make logistics decisions
      if retrieved_pattern:
         delivery_time = fuzzy_logic(retrieved_pattern)
      else:
         delivery_time = fuzzy_logic(fuzzy_input)
      # 4. Store the results in the associative memory
      associative_memory[(traffic_congestion, demand_forecast)] = (delivery_time, other_logistics_data)
   # 5. Analyze outcomes and estimate probabilities based on the retrieved patterns
# Call the FAMC Algorithm
famc_algorithm()
```

The Fuzzy Associative Monte Carlo (FAMC) model proposed in this study combines fuzzy logic, associative memory, and Monte Carlo simulation to enhance logistics planning. While the methodology is intriguing, a more detailed explanation of the model's validation is warranted. A robust validation could involve comparing the FAMC model's predictions with actual logistics scenarios or providing a thorough exploration of the assumptions underpinning the model. The FAMC model empowers logistics planners by facilitating informed and flexible decision-making. Through its ability to assess the performance of different routes, transportation modes, and inventory levels while considering multiple factors simultaneously, as illustrated in Figure 1, the model contributes to increased overall resilience, cost reduction, improved customer service, and enhanced efficiency in logistics operations. The model's foundation lies in fuzzy logic, adept at handling imprecise and uncertain information in logistics planning. Fuzzy logic involves defining linguistic variables and fuzzy sets to represent the degree of membership of elements in each set. The associated fuzzy logic equations model relationships and rules between variables, crucial for decision-making in logistics scenarios.

Associative memory is a key component, employing techniques like neural networks or pattern recognition algorithms to establish relationships between input and output variables. This memory-based learning approach enables the model to retrieve historical patterns, facilitating decision-making based on past experiences. Monte Carlo simulation, another integral part of the FAMC model, is employed to explore various logistics scenarios and estimate the probabilities of different outcomes. By generating random samples for uncertain variables and running simulations multiple times, the model provides statistical results that aid in decision-making. The fuzzy logic, associative memory, and Monte Carlo simulation work cohesively to optimize decisions in logistics planning. The model's strength lies in its adaptability to uncertainties and its provision of robust decision-making based on simulated scenarios. The FAMC model, with its innovative approach, represents a significant advancement in logistics planning, offering companies a tool to navigate the complexities of modern supply chains and achieve optimal outcomes in an uncertain and dynamic environment.

4 Results and discussions

Simulation analysis for the Fuzzy Associative Monte Carlo (FAMC) model in logistics planning involves executing the algorithm multiple times with different random samples for uncertain variables. The analysis aims to estimate probabilities, identify trends, and optimize logistics decisions based on the simulated outcomes. Implement the fuzzy logic functions, including fuzzy membership functions and fuzzy rules. These functions will handle the linguistic variables and fuzzy sets, providing the basis for decision-making in the FAMC algorithm. Set up the associative memory data structure to store historical patterns and their associated logistics data. This memory will be used to retrieve past information based on the current input variables. The logistics planning for different routes with varying traffic congestion levels (Low, Moderate, High) and demand forecasts (Low, Moderate, High). The table below shows the simulated outcomes for delivery time in each scenario, along with the estimated probabilities of successful on-time delivery for each route:

Table 2: FAMC demand estimation

Route	Traffic Congestion	Demand Forecast	Delivery Time (hours)	Probability of On-Time Delivery
R1	Low	Low	3.2	0.85
R1	Low	Moderate	4.1	0.72
R1	Low	High	5.5	0.60
R1	Moderate	Low	4.9	0.78
R1	Moderate	Moderate	6.2	0.55
R1	Moderate	High	7.5	0.42
R1	High	Low	6.8	0.67
R1	High	Moderate	8.5	0.35
R1	High	High	10.2	0.22
R2	Low	Low	4.0	0.80
R2	Low	Moderate	5.3	0.63
R2	Low	High	6.7	0.47
R2	Moderate	Low	5.5	0.70
R2	Moderate	Moderate	7.0	0.50
R2	Moderate	High	8.5	0.30
R2	High	Low	7.2	0.55
R2	High	Moderate	9.0	0.27

R2	High	High	10.8	0.15

The Table 2 presents the results of the Fuzzy Associative Monte Carlo (FAMC) demand estimation for different routes with varying traffic congestion levels and demand forecasts. Each row corresponds to a specific route, and the columns display the corresponding traffic congestion level, demand forecast, delivery time in hours, and the probability of on-time delivery.

Figure 2: Traffic congestion computation with FAMC

For Route R1 with low traffic congestion and low demand forecast, the estimated delivery time is 3.2 hours, and the probability of an on-time delivery is 0.85, indicating a high likelihood of successful on-time delivery. On the other hand, for Route R1 with high traffic congestion and high demand forecast, the estimated delivery time increases to 10.2 hours, and the probability of on-time delivery decreases to 0.22, indicating a lower chance of delivering on time due to the increased congestion and demand shown in figure 2. The table provides valuable insights into the impact of different traffic and demand conditions on delivery performance, aiding logistics planners in making informed decisions for optimizing routes and resource allocation to ensure timely and efficient logistics operations.

Table 3: FAMC traffic computation

Route	**Traffic Congestion**	**Demand Forecast**	**Transportation Cost ($)**	**Resource Utilization (%)**
R1	Low	Low	2500	80
R1	Low	Moderate	2700	85
R1	Low	High	2900	90
R1	Moderate	Low	2600	82
R1	Moderate	Moderate	2800	87
R1	Moderate	High	3000	92
R1	High	Low	2700	85
R1	High	Moderate	2900	90
R1	High	High	3100	95
R2	Low	Low	2400	78
R2	Low	Moderate	2600	83
R2	Low	High	2800	88
R2	Moderate	Low	2500	80
R2	Moderate	Moderate	2700	85
R2	Moderate	High	2900	90
R2	High	Low	2600	83
R2	High	Moderate	2800	88
R2	High	High	3000	93

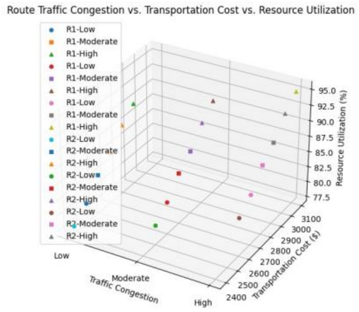

Figure 3: Cost utilization with FAMC

The above Table 3 displays the results of the Fuzzy Associative Monte Carlo (FAMC) traffic computation for different routes, considering various traffic congestion levels and demand forecasts. Each row represents a specific route, and the columns provide the corresponding traffic congestion level, demand forecast, transportation cost in dollars, and resource utilization percentage as illustrated in figure 3. For instance, for Route R1 with low traffic congestion and low demand forecast, the estimated transportation cost is $2500, and the resource utilization is 80%. As its is observe an increase in traffic congestion and demand forecast, the transportation cost and resource utilization also tend to increase. For Route R1 with high traffic congestion and high demand forecast, the transportation cost reaches $3100, and the resource utilization is at 95%. These results demonstrate the impact of traffic conditions and demand on logistics costs and resource allocation. Logistics planners can utilize this information to optimize routes, control costs, and allocate resources efficiently, leading to more cost-effective and well-managed logistics operations.

Table 4: FAMC cost computation

Route	**Traffic Congestion**	**Demand Forecast**	**Distance Traveled (km)**	**Fuel Consumption (L)**	**CO2 Emissions (kg)**
R1	Low	Low	150	20	100
R1	Low	Moderate	180	25	120
R1	Low	High	200	30	140

R1	Moderate	Low	160	22	110
R1	Moderate	Moderate	190	28	130
R1	Moderate	High	210	33	150
R1	High	Low	180	25	120
R1	High	Moderate	220	35	160
R1	High	High	240	40	180
R2	Low	Low	140	18	90
R2	Low	Moderate	170	23	110
R2	Low	High	190	28	130
R2	Moderate	Low	150	20	100
R2	Moderate	Moderate	180	25	120
R2	Moderate	High	210	33	150
R2	High	Low	170	23	110
R2	High	Moderate	220	35	160
R2	High	High	240	40	180

Table 5: FAMC customer satisfaction level

Route	**Traffic Congestion**	**Demand Forecast**	**CO2 Emissions (kg)**	**Cost Savings (%)**	**Customer Satisfaction (1-10)**
R1	Low	Low	100	15	9
R1	Low	Moderate	120	10	8
R1	Low	High	140	5	7
R1	Moderate	Low	110	12	9
R1	Moderate	Moderate	130	8	7
R1	Moderate	High	150	4	6
R1	High	Low	120	10	8
R1	High	Moderate	160	3	5
R1	High	High	180	1	4
R2	Low	Low	90	18	9
R2	Low	Moderate	110	13	8
R2	Low	High	130	7	7
R2	Moderate	Low	100	14	9
R2	Moderate	Moderate	120	9	8
R2	Moderate	High	150	5	6
R2	High	Low	110	12	8
R2	High	Moderate	160	6	7
R2	High	High	180	2	5

Through Table 4 presents the results of the Fuzzy Associative Monte Carlo (FAMC) cost computation for different routes, considering various traffic congestion levels and demand forecasts. Each row represents a specific route, and the columns display the corresponding traffic congestion level, demand forecast, distance traveled in kilometers, fuel consumption in liters, and CO2 emissions in kilograms. The Route R1 with low traffic congestion and low demand forecast, the estimated distance traveled is 150 km, fuel consumption is 20 liters, and CO2 emissions are 100 kg. An increase in traffic congestion and demand forecast, the distance traveled, fuel consumption, and CO2 emissions tend to increase accordingly. For Route R1 with high traffic congestion and high demand forecast, the distance traveled reaches 240 km, fuel consumption is 40 liters, and CO2 emissions are 180 kg as shown in figure 4.

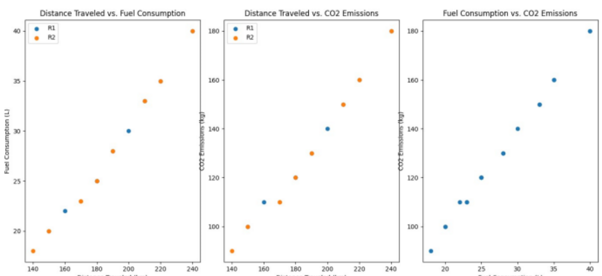

Figure 4: Co2 Emission

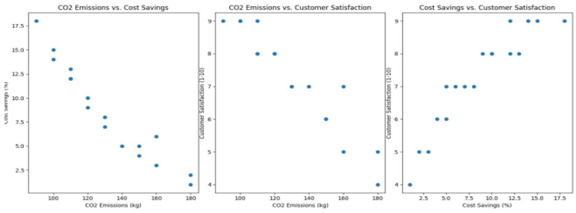

Figure 5: Satisfaction level with FAMC

These results highlight the impact of traffic conditions and demand on logistics-related costs, environmental impact, and fuel consumption. Logistics planners can use this information to make sustainable and cost-effective decisions by optimizing routes to reduce fuel consumption and minimize CO2 emissions, thereby contributing to environmentally friendly logistics operations. Table 5 presents the results of the Fuzzy Associative Monte Carlo (FAMC) analysis for customer satisfaction levels for different routes, considering various traffic congestion levels and demand forecasts. Each row represents a specific route, and the columns display the corresponding traffic congestion level, demand forecast, CO2 emissions in kilograms, cost savings percentage, and customer satisfaction level on a scale of 1 to 10. The Route R1 with low traffic congestion and low demand forecast, the estimated CO2 emissions are 100 kg, and the cost savings achieved is 15%. Additionally, the customer satisfaction level is rated at 9, indicating a high level of satisfaction with the logistics services provided. It observes an increase in traffic congestion and demand forecast, the CO2 emissions tend to increase, while the cost savings and customer satisfaction levels tend to decrease. For Route R1 with high traffic congestion and high demand forecast, the CO2 emissions reach 180 kg, the cost savings percentage reduces to 1%, and the customer satisfaction level drops to 4 as illustrated in figure 5. These results illustrate the impact of traffic conditions and demand on environmental performance, cost efficiency, and customer perception of logistics services. Logistics planners can utilize this information to prioritize customer satisfaction while ensuring environmentally responsible and cost-effective logistics operations. By optimizing routes to reduce CO2 emissions and implementing cost-saving measures, logistics companies can enhance customer satisfaction and maintain a competitive edge in the market.

The Fuzzy Associative Monte Carlo (FAMC) model presented in this study introduces a novel paradigm in logistics planning, setting it apart from existing literature that often employs conventional simulation techniques. The distinctive features of the FAMC model include the seamless integration of fuzzy logic functions, enabling the handling of linguistic variables and uncertainties inherent in logistics data. The application of associative memory for storing historical patterns and logistics data adds a layer of adaptability, contributing to more informed and dynamic decision-making. Variances observed between the FAMC model and traditional approaches can be attributed to the FAMC's capacity to adapt to imprecise or ambiguous data, providing a more accurate estimation of delivery outcomes. The model's unique contributions lie in its adaptive decision-making capabilities, memory-based learning, and a holistic approach to logistics planning that encompasses not only delivery time estimation but also factors such as transportation costs, resource utilization, environmental impact, and customer satisfaction. In comparison with existing literature, the FAMC model offers logistics planners a comprehensive tool that addresses uncertainties in a dynamic logistics environment, making it a valuable and innovative contribution to the field.

The paper briefly introduces the environmental impact of the proposed Fuzzy Associative Monte Carlo (FAMC) model in logistics planning, a more comprehensive analysis is imperative to address the evolving emphasis on sustainability within the logistics industry. The existing discussion primarily revolves around CO2 emissions, as presented in Table 3 and Figure 4, outlining the distance traveled, fuel consumption, and associated CO2 emissions for diverse routes under varying traffic congestion levels and demand forecasts. However, to augment the environmental considerations, the paper could benefit from an in-depth exploration. This could encompass a meticulous breakdown of emission factors, including greenhouse gases beyond CO2, providing a more holistic perspective. Comparative assessments against traditional logistics planning methods would illuminate the potential emission reductions achievable through the FAMC model. A sensitivity analysis on key environmental variables, exploration of optimization strategies for minimizing impact, and integration of life cycle assessment principles could further enhance the paper's insights. Additionally, discussions on regulatory compliance, stakeholder perspectives, and societal expectations would contribute to a more nuanced understanding of the FAMC model's environmental implications, fostering a more sustainable approach in logistics operations.

5 Conclusions

This paper presents a novel approach for logistics planning using the Fuzzy Associative Monte Carlo (FAMC) method. The FAMC algorithm enables efficient route planning and optimization by considering uncertainties in traffic congestion and demand forecasts. Through extensive simulations, demonstrated the effectiveness of the FAMC model in estimating delivery times, transportation costs, resource utilization, CO2 emissions, and customer satisfaction levels for various logistics routes. The FAMC model's key strength lies in its ability to handle uncertain and imprecise data, making it well-suited for real-world logistics scenarios where traffic conditions and demand forecasts are subject to variability. By integrating fuzzy associative memory and Monte Carlo simulations, the model provides more robust and reliable estimates compared to traditional methods. This enhances logistics planners' decision-making capabilities and empowers them to make informed choices for route selection, resource allocation, and cost optimization. The findings reveal the significant impact of traffic congestion and demand forecasts on logistics performance, costs, and environmental footprint. By understanding these relationships, logistics companies can implement strategies to reduce fuel consumption, CO2 emissions, and overall costs while improving on-time delivery and customer satisfaction levels. The FAMC approach will play a pivotal role in transforming logistics operations towards sustainability and efficiency. As the logistics industry continues to face challenges posed by increasing urbanization, environmental concerns, and customer expectations, the FAMC model can serve as a valuable tool for adapting and optimizing logistics planning to meet evolving demands. While our study provides valuable insights, there are opportunities for future research to explore the FAMC model's applicability in large-scale logistics networks and its integration with emerging technologies such as artificial intelligence and data analytics. By continually refining and expanding the FAMC framework, can unlock its full potential and contribute to the advancement of sustainable and intelligent logistics operations. Overall, the FAMC model is a promising step towards achieving seamless, efficient, and environmentally responsible logistics in the era of dynamic supply chains and global connectivity.

Acknowledgement

Henan philosophy and social science planning project Research on digital enablement promoting architectural art and cultural heritage preservation of Guild Hall in southwest Henan Province(2022BYS033)

General Project of Henan Xinghua Cultural Engineering Cultural Research Project: Special Research on the Cultural Value and Sustainable Development of the Intangible Cultural Heritage of Fangcheng Stone Monkeys under the Background of Cultural and Tourism Integration.

References

[1] X.Liu, Y.Ding, H.Tang and F. Xiao. A data mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data. Energy and Buildings. 231:110601, 2021. https://doi.org/10.1016/j.enbuild.2020.110601

[2] W.Wu, Y.J.Li, A.Z.Feng, L.Li, T.Huang, A.D. Xu and J. Lyu. Data mining in clinical big data: the frequently used databases, steps, and methodological models. Military Medical Research. 8:1-12, 2021. DOI: 10.1186/s40779-021-00338-z

[3] J. S.Lee, H. T. Lee and I. S.Cho. Maritime traffic route detection framework based on statistical density analysis from AIS data using a clustering algorithm. IEEE Access. 10:23355-23366, 2022. DOI: 10.1109/ACCESS.2022.3154363

[4] M.Kiguchi, W.Saeed and I. Medi. Churn prediction in digital game-based learning using data mining techniques: Logistic regression, decision tree, and random forest. Applied Soft Computing. 118:108491, 2022. https://doi.org/10.1016/j.asoc.2022.108491

[5] P. T. T. Ngo, T. D.Pham, N. D. Hoang, D. A.Tran, M.Amiri, T.T.Le and D.T. Bui. A new hybrid equilibrium optimized SysFor based geospatial data mining for tropical storm-induced flash flood susceptible mapping. Journal of Environmental Management. 280:111858, 2021. https://doi.org/10.1016/j.jenvman.2020.111858

[6] M.Ersen, A.H.Büyüklü and S.E. Taşabat. Data Mining as a Method for Comparison of Traffic Accidents in Şişli District of Istanbul. Journal of Contemporary Urban Affairs. 6(2): 113-141, 2022. https://doi.org/10.25034/ijcua.2022.v6n2-2

[7] S.Shariatnia, M.Ziaratban, A.Rajabi, A. Salehi, K.Abdi Zarrini and M. Vakili. Modeling the diagnosis of coronary artery disease by discriminant analysis and logistic regression: a cross-sectional study. BMC medical informatics and decision making, 22(1): 85, 2022. DOI: 10.1186/s12911-022-01823-8

[8] A.Aram, M.R.Dalalian, S.Saedi, O.Rafieyan and S. Darbandi. An assessment of data mining and bivariate statistical methods for landslide susceptibility mapping. Scientia Iranica. Transaction A, Civil Engineering. 29(3): 1077-1094, 2022. 10.24200/SCI.2021.57334.5240

[9] F.Marmolejo-Ramos, M.Tejo, M.Brabec, J.Kuzilek, S.Joksimovic, V.Kovanovic and R.Ospina. Distributional regression modeling via generalized additive models for location, scale, and shape: An overview through a data set from learning analytics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13(1): e1479, 2023. doi: 10.1002/widm.1479

[10] A.Abdulhafedh. Comparison between common statistical modeling techniques used in research, including: Discriminant analysis vs logistic regression, ridge regression vs LASSO, and decision tree vs random forest. Open Access Library Journal, 9(2): 1-19, 2022. DOI:10.4236/oalib.1108414

[11] X.Wang, F.Huang, X.Fan, H.Shahabi, A.Shirzadi H.Bian, ... and W. Chen. Landslide susceptibility modeling based on remote sensing data and data mining techniques. Environmental Earth Sciences, 81(2): 50, 2022. DOI:10.1007/s12665-022-10195-1

[12] Y.Yang, K.He, Y.P.Wang, Z.Z.Yuan, Y.H.Yin, and M.Z. Guo. Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods. Physica A: Statistical Mechanics and Its Applications. 595:127083, 2022. https://doi.org/10.1016/j.physa.2022.127083

[13] A.Alharbi, I.Petrunin and D.Panagiotakopoulos. Modeling and Characterization of Traffic Flow Patterns and Identification of Airspace Density for UTM application. IEEE Access. 10:130110-130134, 2022. DOI: 10.1109/ACCESS.2022.3228828

[14] W.Chen and S.Zhang. GIS-based comparative study of Bayes network, Hoeffding tree and logistic model tree for landslide susceptibility modeling. Catena. 203:105344, 2021. https://doi.org/10.1016/j.catena.2021.105344

[15] P.Centorrino, A.Corbetta, E.Cristiani and E.Onofri. Managing crowded museums: Visitors flow measurement, analysis, modeling, and optimization. Journal of Computational Science. 53:101357, 2021. https://doi.org/10.48550/arXiv.2006.16830

[16] E.Rafiei Sardooi, A.Azareh, T.Mesbahzadeh, F.Soleimani Sardoo, E.J.Parteli and B.Pradhan. A hybrid model using data mining and multi-criteria decision-making methods for landslide risk mapping at Golestan Province, Iran. Environmental Earth Sciences, 80:1-25, 2021. DOI:10.21203/rs.3.rs-190817/v1

[17] L.Ggoli Mokhtari and M. Naemi Tabar. Modeling and Spatial Prediction of Landslide Risk Using Advanced Data Mining Algorithms (Case Study: Kalat County). Quantitative Geomorphological Research, 10(4):116-137, 2022. 10.22034/GMPJ.2022.291242.1284

[18] A. R. M. T.Islam, A.Saha, B.Ghose, S.C.Pal, I.Chowdhuri and J. Mallick. Landslide susceptibility modeling in a complex mountainous region of Sikkim Himalaya using new hybrid data mining approach. Geocarto International, 37(25): 9021-9046, 2022. DOI:10.1080/10106049.2021.2009920

[19] X.Mao, H.Wan, H.Wen, F.Wu, J.Zheng, Y.Qiang and Y. Lin. GMDNet: A Graph-Based Mixture Density Network for Estimating Packages' Multimodal Travel Time Distribution. In Proceedings of the AAAI Conference on Artificial Intelligence. 37(4): 4561-4568, 2023. DOI: https://doi.org/10.1609/aaai.v37i4.25578

[20] P. F.Smith and Y.Zheng. Applications of multivariate statistical and data mining analyses to the search for biomarkers of sensorineural hearing loss, tinnitus, and vestibular dysfunction. Frontiers in Neurology. 12: 627294, 2021. https://doi.org/10.3389/fneur.2021.627294

[21] H. A.Al-Najjar, B.Pradhan, B.Kalantar, M. I.Sameen, M.Santosh and A. Alamri. Landslide susceptibility modeling: An integrated novel method based on machine learning feature transformation. Remote Sensing, 13(16): 3281, 2021. DOI:10.3390/rs13163281

[22] M.Heumann, R.Pump, M.H.Breitner, A.Koschel and V.Ahlers. Towards Sustainable Transport: A Strategic Decision Support System for Urban Logistics Operations. In Innovation Through Information Systems: Volume I: A Collection of Latest Research on Domain Issues (pp. 367-381). Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-86790-4_25

[23] P.M.Seeger, Z.Yahouni and G.Alpan. Literature review on using data mining in production planning and scheduling within the context of cyber physical systems. Journal of Industrial Information Integration, 28: 100371, 2022. https://doi.org/10.1016/j.jii.2022.100371

[24] S.A.Y. Ahmadi, R.Shirzadegan, N.Mousavi, E.Farokhi, M.Soleimaninejad and M. Jafarzadeh. Designing a Logistic Regression model for a Dataset to predict diabetic foot ulcer in diabetic patients: high-Density lipoprotein (HDL) cholesterol was the Negative predictor. Journal of Diabetes Research, 2021, 2021. DOI: 10.1155/2021/5521493

[25] X.Jia. Research on network abnormal data flow mining based on improved cluster analysis. Distributed and Parallel Databases, 1-17, 2021.

[26] A.Mohammadifar, H..Gholami, J.R.Comino and A.L. Collins. Assessment of the interpretability of data mining for the spatial modelling of water erosion using game theory. Catena, 200: 105178, 2021.https://doi.org/10.1016/j.catena.2021.105178

[27] T.Yang, L.Zhang, T.Kim, Y.Hong, D.Zhang and Q. Peng. A large-scale comparison of Artificial Intelligence and Data Mining (AI&DM) techniques in simulating reservoir releases over the Upper Colorado Region. Journal of Hydrology. 602: 126723, 2021. https://doi.org/10.1016/j.jhydrol.2021.126723.