International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Data-Driven Intelligent Clustering-Based Optimization for Enhancing Urban Logistics Delivery Systems: A Case Study in Casablanca, Morocco

Soufiane Reguemali[1], Abdellatif Moussaid[2*] and Abdelmajid Elouadi[1]

Submitted:11/03/2024 **Revised**: 26/04/2024 **Accepted**: 03/05/2024

Abstract: These In the field of supply chain management, making sure products get to people efficiently is not easy. There are issues like customers being unhappy, figuring out the optimizing itinerary of trucks, how much they can carry, and optimizing the delivery time. In this paper, we are introducing a smart system that makes the whole delivery process smoother, starting from managing inventory to reaching the customers. Our system leverages clustering techniques to automate and simplify this complex process. We collected data and tested our approach on a mass retail company in Casablanca, Morocco. This data includes information about customer locations, order details, and the available delivery trucks with their capacities. At the core of our solution lies a unique clustering algorithm, custom-made to handle our specific challenges. The approach starts by defining how far apart cus-tomers' locations can be, ensuring we don't group locations that are too distant from each other. Then, we use a straightforward method to create these groups based on proximity between locations and order details. This ensures the efficient allocation of customer orders to clusters, maximizing truck fill rates. In short, our innovative approach streamlines delivery operations, reduces cus-tomer complaints, optimizes fleet management and guarantees on-time, cost-effective deliveries with a highly satisfactory service rate.

Keywords: *DDS, Supply chain, Stock delivery, Unsupervised learning, Clustering*

1. Introduction

Distribution, often synonymous with delivery systems, serves as a pivotal component of modern supply chain management. It encompasses the intricate processes by which goods journey from manufacturers to the ultimate consumers. Paramount among the considerations of any company is ensuring the seamless accessibility of their products to consumers. Central to achieving this goal is the selection of an optimal delivery route, which not only influences the distance covered but also the accompanying costs [1–3].

The supply chain process, from stock to delivery, encompasses several vital steps. It begins with stock management, where inventory is acquired, stored, and managed effectively to meet customer demands.

Once inventory is in place, the supply chain process progresses to order processing, where customer orders are received, validated, and processed for fulfillment. Next, the process moves into the production or procurement phase, where products are manufactured, assembled, or sourced from suppliers to meet the demand generated by customer orders. Quality control and assurance measures are often integrated into this phase to ensure that products meet predetermined standards.

[1]*Ibn Tofail University, Ecole National des Sciences Appliquées, Kénitra-Morocco*
[2]*ENSIAS, Mohammed V University in Rabat, Rabat 10000,Morocco*
* *Corresponding author: abdelatif.moussaid@gmail.com*

After production or procurement, the supply chain process transitions to warehousing and distribution. In this stage, products are stored in strategically located warehouses or distribution centers to facilitate efficient order fulfillment. Inventory management plays a critical role here, as it helps maintain optimal stock levels to meet customer demand while minimizing carrying costs.

The final steps in the supply chain process revolve around order picking, packing, and shipping. Products are selected from the warehouse, packed according to specific customer requirements, and then shipped using chosen transportation methods.

Deliveries can be executed through drones [4,5] or motorcycles [6], eliminating the need for a local delivery system as they typically transport a single product per customer. However, for larger-scale deliveries to markets within neighborhoods and streets, vehicles like trucks are essential to transport maximum units in minimal time. Real-time tracking and visibility mechanisms become pivotal in monitoring shipment status, ensuring timely deliveries, and offering customers precise delivery estimates [7–10].

Currently, many companies grapple with manual and sporadic delivery processes, wherein goods travel from depots to consumers incurring significant time and cost implications [11,12]. In essence, these delivery systems often entail a complex web of logistical challenges. So, we can see that the delivery systems take a big part in the process of the supply chain.

To build a streamlined and effective delivery system, it is imperative to carefully consider various critical factors. Among these, the accurate localization data of customers, whether they are individual clients or patrons of supermarkets, plays a pivotal role [14,15]. Leveraging this geographical information, advanced algorithms can be applied to create clusters, facilitating the segmentation of customers into groups with common characteristics, such as proximity [16,17]. This approach enhances the precision and efficiency of the delivery system by tailoring services to specific customer clusters. In this context, unsupervised learning algorithms, a subset of Machine Learning, become useful in this scenario [18]. A prominent example is K-means [19], which employs Euclidean distance to organize data into clusters.

In the case of K-means, it is instrumental to delve deeper into its functionality. K-means is an iterative algorithm that partitions a dataset into 'k' clusters, where each data point belongs to the cluster with the nearest mean. The algorithm starts by initializing 'k' centroids, typically randomly. Then, it assigns each data point to the nearest centroid and recalculates the centroids' positions based on the data points within their respective clusters. This process iterates until the centroids no longer significantly change [20].

Hierarchical clustering is another clustering technique worth mentioning, as it constructs a hierarchy of clusters by iteratively merging or splitting existing clusters based on certain criteria [21]. Besides K-means and hierarchical clustering, there exist various other clustering techniques, each with its own strengths and applications [22].

Despite all these approaches, there are not several works that exploit unsupervised learning in order to create an intelligent delivery system to control and optimize this operation of the supply chain. Actually, there are some basic applications that have been used for several years. The idea is to use an electronic system and connect it to GPS to get data and recommend roads to the conductor of the delivery truck. But here the problem is: how the conductor can choose the best starting point, and how can control the point of delivery? Also, sometimes, there are some products that need to be delivered in the minimum of time, especially perishable goods like milk, yogurt, and more [23] . So here, we need a local system that divides the customers (supermarkets) into several groups to recommend the itinerary to the conductor. In most cases, there is a team inside the company that does this manually based on their experience. To automate this process, the use of clustering algorithms is necessary [5].

In this way, Taeho Kim et al. Introduced a clustering-based multiple ant colony algorithm to optimize delivery operations. The research aims to address the complex logistics challenges associated with delivering home appliances in Korea. The proposed algorithm combines clustering techniques with ant colony optimization to improve the efficiency and effectiveness of delivery routes [24].

Also, the paper of Bosona et al. presents a comprehensive approach to enhance the local food supply chain through the strategic utilization of clustering techniques within the context of food delivery logistics. By initially identifying and mapping local food producers and distribution centers, the study lays the foundation for more efficient logistics. It introduces clustering methodologies to categorize producers into clusters, promoting better coordination. Furthermore, the research addresses the optimization of Collection Center (CC) locations and the mapping of efficient routes for product collection and delivery [25].

On the other hand, Dhirendra et al. introduced a clustering-based routing heuristic (CRH) for fresh food delivering. The proposed CRH optimizes vehicle routing by clustering demand nodes, ensuring efficiency with single-vehicle serviceability. The algorithm's computational efficiency stands out, providing optimal solutions swiftly across various scenarios [26].

Despite these various applications, the intricate nature of delivery logistics for small urban markets, especially for daily items like milk, bread, and croissants, presents a significant challenge. The distinct product quantity requirements of each market pose difficulties in incorporating them into clustering algorithms based on customer locations, impeding the efficient determination of maximum truck delivery quantities.

However, within these challenges lies an opportunity for improvement through thoughtful analysis. The automation of manual operations, including customer clustering and truck recommendation based on optimal delivery capacities, becomes crucial in addressing diverse daily delivery needs. This streamlining enhances overall efficiency in last-mile logistics.

This paper introduces an innovative approach to tackle these distribution challenges, focusing on the Capacity Vehicle Routing Problem (CVRP) and a tailored clustering algorithm.

In contrast to previous attempts using generic clustering methods struggling with varying client orders, our personalized algorithm effectively overcomes this obstacle. By leveraging client data, GPS coordinates, and truck specifications, we not only define clusters based on geographic proximity but also introduce an additional layer of complexity by aligning these clusters closely with the capacity constraints of delivery trucks.

Our method, which is a meticulous step-by-step process of data sorting and allocation, results in the creation of distinct groups of stores. Each of these groups is carefully constructed to align with the quantities of products that will be delivered, closely matching the capacity of the respective delivery truck.

These tailored groupings are preserved for further optimization as we proceed to generate efficient delivery routes using Open Route services API (ORS) [27].

Thus, our comprehensive approach not only promises to minimize distances, reduce costs, and save time but also offers a tailored solution to the intricate challenges of modern distribution systems, ensuring that clusters are formed with precision based on both geographic proximity and the crucial constraint of truck capacity.

In the forthcoming sections, we will delve into the heart of our research. In Materials and Methods section, we will provide an in-depth exploration of our data collection process and unveil the inner workings of our custom clustering algorithm designed to address specific challenges in supply chain logistics while adhering to our constraints. The Results and Validation section will showcase the outcomes of our approach and reveal how we substantiated the effectiveness of our intelligent delivery systems through user applications and practical construction. Finally, in the Conclusion, we will draw together the threads of our study, highlighting its implications for modern delivery systems and supply chain management, all while offering a glimpse into potential future developments.

2. Method

2.1. Proposed research

The proposed research entails the development of an intelligent system leveraging the capabilities of unsupervised machine learning to optimize product distribution for large companies. The primary objective is to facilitate rapid, precise, and cost-effective product delivery. To realize this system, a series of well-defined steps (Figure 1) were meticulously followed.

Data Collection: The foundation of our optimization process lies in the acquisition of comprehensive data. This dataset encompasses crucial factors related to product delivery, including

client information (with an emphasis on client locations), order details (quantity and specifics), available delivery vehicles, and the respective capacities of these vehicles.

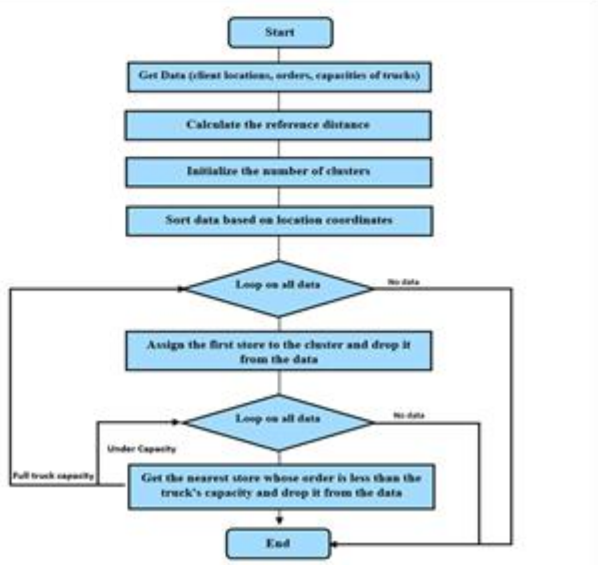

Fig 1 : Workflow of the proposed approach

Cluster Segmentation: To initiate the optimization process effectively, data segmentation was performed based on the types of delivery trucks available. Each truck type possesses its unique capacity constraints.

Reference Geodistance: A pivotal aspect of our methodology involves the calculation of the reference geodistance [28]. This metric represents the maximum geometric distance that two data points within the same cluster should not exceed. The geodistance metric, integral in the field of geometry, is determined using the geopy module and the longitude and latitude coordinates of each store (data point), and the mathematical formula as bellow:

$$a = sin^2\left(\frac{\Delta lat}{2}\right) + \cos(lat1).\cos(lat2).sin^2\left(\frac{\Delta lon}{2}\right) \quad (1)$$

$$c = 2.atan2(\sqrt{a}, \sqrt{(1-a)}) \quad (2)$$

$$d = R.C$$
(3)

Where: d is the geodistance (distance between points on the Earth's surface), R is the Earth's mean radius (approximately 6,371 km or 3,959 miles), lat1 and lat2 are latitudes of the two points (in radians), Δlat and Δlon are the difference in latitudes and longitude, c is the central angle between the two points (in radians) and atan2 is the arctangent function.

Clustering Algorithm: At the core of our approach lies a clustering algorithm, which plays a pivotal role in identifying groups with inherent similarities within the dataset. While we initially explored established unsupervised learning algorithms such as k-means and hierarchical clustering, we encountered limitations. These algorithms primarily base their clustering on inputs alone and cannot account for constraints during the process, which proved insufficient for our needs.

They did not adequately consider the quantities associated with each client, a critical aspect of our constraints.

To address this limitation, we embarked on the development of our proprietary clustering algorithm (Figure 1).

This algorithm is specifically designed to consider both geographical proximity and the quantities of products associated with each client. By harmonizing these essential factors, our custom clustering algorithm enables us to create clusters that optimize spatial efficiency while adhering to truck capacity constraints.

2.2. Data

The data used in this project pertains to a milk distribution company in Morocco. Given the perishable nature of milk, timely delivery is of utmost importance. Consequently, our system has been designed to prioritize robustness, efficiency, and speed in the delivery process.

The collected data encompasses various aspects, primarily focusing on the localization of shops (Figure 2). Additionally, it includes information regarding the types of vehicles employed for milk delivery, tailored to the specific requirements of each shop.

For instance, smaller shops necessitate compact vehicles capable of navigating narrow streets, while larger vehicles are allocated for delivering substantial quantities to larger outlets such as markets. In addition to shop information, the dataset includes daily order quantities for milk boxes and specifies the type of truck responsible for product delivery to each shop. Moreover, the maximum number of available delivery trucks, along with their respective capacities.

It also defines the maximum distance each vehicle must traverse from the factory to individual shops, considering the time required for unloading goods at each location.

Regarding data cleanliness, it's noteworthy that the dataset has undergone meticulous preparation and cleaning. Experts in the field of logistics and milk delivery have ensured that the data is devoid of missing values or outliers, as working with erroneous data can significantly impact the system's performance. Furthermore, it's important to acknowledge that the data is dynamic in nature, subject to modifications as the business environment evolves.

For instance, when a shop closes, it must be removed from the dataset. Conversely, when a new shop opens, it must be added as a new client. Similarly, any adjustments to order quantities or other relevant data are accommodated. These modifications can be seamlessly incorporated into the dataset, allowing our system to adapt and generate updated clusters with essential output information as needed.

Fig 1 : Customer locations – Casablanca Morocco

2.3. Store clustering algorithm

The clustering process began by computing the reference distance, determined as the maximum geodetic distance allowable between two data points within the same cluster, with technical guidance from my supervisor. Subsequently, we applied a straightforward clustering algorithm with key steps: copying data into a separate dataframe and sorting it by coordinates in ascending order, initializing cluster numbers, iterating through stores to allocate them based on distance and order constraints, updating the cluster count, and repeating these steps until the dataframe is empty. The algorithm's details are outlined algorithm 1.

<u>Algorithm 1:</u> Store Clustering for Efficient Product Delivery

---

Input (df):

Stores coordinates: *Latitude and longitude*

Reference distance: *The maximum allowed geodistance between stores in a cluster.*

Truck capacity: *The maximum number of orders a delivery truck can handle*

Output: ***Clusters of stores optimized for delivery*.**

<u>Algorithm Steps:</u>

Step 0: Initialize clusters
clusters = [ ], current_cluster = [ ]
Step 1: Sort data based on coordinates
sorted_df = sorted (df)
Step2: Initialize cluster variables
c = 1, cluster_c = [ ]
Step3: Iterate through stores
for store in sorted_df:
distance_to_cluster = [geodistance (store, s) for s in cluster_c]
if not cluster_c or all
d <= Reference_Distance for d in distance_to_cluster)
sum (store.orders for store in cluster_c) <= Truck_Capacity:
cluster_c.append(store)
else:
clusters. append(cluster_c) and cluster_c = [store]
Step4: Add the last cluster
clusters.append(cluster_c)
Step5: Dataframe Management and Cluster Increment
for cluster in clusters:
for store in cluster:
sorted_df.remove(store)
sorted_df.reset_index(drop=True)
c += 1

Step7: Compare geodistances and create clusters

```
while sorted_df:
    cluster_c = [ ]
    for store in sorted_df:
        distance_to_cluster = [geodistance(store, s) for s in cluster_c]
        if not cluster_c or all(d <= Reference_Distance for d in distance_to_cluster) and
         sum(store.orders for store in cluster_c) <= Truck_Capacity:
            cluster_c.append(store)
    clusters.append(cluster_c)
```

Step8: Dataframe Management and Cluster Increment

```
  for cluster in clusters:
      for store in cluster:
          sorted_df.remove(store)
    sorted_df.reset_index(drop=True)
  c += 1
```

2.4. Road mapping

This crucial step involves the establishment of optimized routes for the trucks to efficiently deliver products, specifically milk in our context. To accomplish this, we leveraged the Open Route service library, which harnesses real-time satellite data to provide valuable insights into directions, optimal routes, distances between points, and time-related information.

Our road mapping process commenced by charting out the routes for each cluster, commencing from the factory and navigating through the various shop locations before returning. These routes were generated as both visual maps, illustrating the path the trucks would follow, and detailed Excel files. These files include essential information such as the estimated arrival times at each store, unloading durations, and the respective quantities of orders for each store. This meticulous road mapping ensures timely and efficient product deliveries.

3. Results and discussion

To assess the efficacy of our solution, we conducted testing using a subset of the data comprising 55 stores for one type of truck and 24 stores for another, all located within the city of Casablanca in Morocco (Figure 2). This test scenario was characterized by specific constraints:

a maximum cluster size (representing each truck's capacity) of 20 units for the first type of truck catering to specific stores, and 30 units for the second type. Additionally, each delivery operation had a fixed duration of 180 minutes, which included the time required for unloading goods at each store.

Our system was rigorously evaluated based on its ability to cluster input data effectively, optimizing the utilization of available trucks' capacities, streamlining delivery routes, and providing precise delivery and turnaround times. This testing process aimed to validate the system's performance and its capacity to efficiently manage a complex distribution network.

Based on the simulation results presented in Figure 3, our system demonstrates impressive performance in optimizing the distribution process. Specifically, it successfully divides the delivery task into distinct clusters, tailored to the capabilities of two types of trucks employed in the operation. For the first type of truck, a total of eight clusters are formed, with seven of these clusters being fully occupied, accommodating precisely 20 orders each. The eighth cluster handles the remaining 15 orders, further showcasing the system's adaptability in addressing varying order quantities.

Fig 3 : Simulation results

Similarly, for the second type of truck, responsible for a total of 24 orders, the system efficiently creates three clusters, two of which operate at full capacity, each accommodating 10 orders, while the third cluster manages the remaining four orders. These results underscore the effectiveness of our clustering algorithm in optimizing truck assignment, minimizing empty spaces in delivery vehicles, and ensuring a resource-efficient distribution process.

These outcomes are particularly promising in terms of supply chain logistics, as they indicate that a significant portion of the delivery trucks are operating at full capacity. This not only enhances the efficiency of the distribution process but also contributes to notable cost reductions within the supply chain, ultimately translating into increased operational efficiency and customer satisfaction.

After obtaining the groups for each type of truck as an Excel file, we employed the Open Route Service (ORS) API, specifically utilizing the Folium mapping system, to construct the routes that drivers would follow for delivery. Figure 4 provides an example of one group, illustrating the mapped routes on the map, and the corresponding table. In this example, we set the starting time for the trajectories at 06:00, and the arrival time at each store is displayed in the table, along with the quantities of requested items. This approach greatly facilitates precise and efficient deliveries. These routes trace the paths originating from the factory, passing through all the shops, and returning upon the completion of stock deliveries.

Fig 2 : Example of truck trajectory

Finally, the obtained results were tested within the city of Casablanca, Morocco, over the course of one week in collaboration with our milk delivery partner. The results proved to be highly satisfactory, resulting in a significant reduction in customer complaints. Previously averaging 10 complaints per week, this number was reduced to just 3 complaints per week. The primary cause of these complaints was the timeliness of deliveries, a critical factor for our customers who prioritize receiving products promptly to facilitate early sales.

Furthermore, our optimization efforts have enabled the company to operate with a minimal fleet of trucks, reducing dependence on excess resources and positively impacting overall turnover. Additionally, our system provides decision-makers with a clear view of truck trajectories, allowing for the prompt detection of any anomalies related to the road, thereby facilitating swift decision-making. Moreover, it offers an intuitive graphical interface that enhances the satisfaction of drivers as they carry out their responsibilities.

4. Conclusion

In conclusion, our innovative intelligent system represents a significant leap forward in automating the order delivery process. Leveraging a groundbreaking algorithm developed through extensive research and utilizing data collected in Casablanca, Morocco, our system has redefined how deliveries are managed.

Our system streamlines the entire process, optimizing delivery routes, and minimizing costs. Through rigorous testing within the bustling city of Casablanca, our system has demonstrated remarkable efficiency, significantly reducing customer complaints from an average of 10 per week to a mere 3 per week. Timeliness in delivery, a critical concern for our customers, has been greatly enhanced, allowing them to access their products promptly, thereby facilitating early sales.

One of the standout advantages of our system is its ability to reduce the number of trucks required for deliveries while simultaneously optimizing the shipping and delivery processes. Additionally, it offers precise road tracking, ensuring that deliveries are made with utmost efficiency.

As we look to the future, our vision extends beyond the achievements of today. We are committed to integrating our system into an embedded platform with a user-friendly control interface for real-time truck management. This integration will revolutionize how delivery operations are monitored and controlled.

Furthermore, we envision creating a robust communication database that fosters seamless interaction between drivers and the supply chain department. This open line of communication will enable rapid response to any unforeseen challenges and enhance the overall efficiency of the system.

A particularly exciting prospect for the future is the integration of customer order requests directly into our system. By automating the process of collecting and incorporating customer orders before delivery, we aim to further streamline operations and provide an even higher level of service.

Finally, our system has already made substantial strides in modernizing the delivery process, and our commitment to ongoing innovation promises to shape the future of supply chain management, offering increased efficiency, customer satisfaction, and adaptability in a rapidly evolving business landscape.

Author contributions

Soufiane Reguemali: Conceptualization, Methodology, Software, Field study, Data curation, Writing-Original draft preparation, Software, Validation, Visualization, Investigation, Writing-Reviewing and Editing.

Abdellatif Moussaid: Conceptualization, Methodology, Software, Field study, Data curation, Writing-Original draft preparation, Software, Validation, Visualization, Investigation, Writing-Reviewing and Editing.

Abdelmajid Elouadi: Conceptualization, Methodology, Software, Field study, Data curation, Writing-Original draft preparation, Software, Validation, Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] M. H. Hugos, Essentials of supply chain management. John Wiley & Sons, 2018.

[2] F. R. Jacobs, R. B. Chase, et R. R. Lummus, Operations and supply chain management. McGraw-Hill/Irwin New York, NY, 2014.

[3] E. Muhalia, P. Ngugi, et M. Moronge, « Effect Of Transportation Management Systems On Supply Chain Performance Of Fmcg In Kenya », Am. J. Supply Chain Manag., vol. 6, no 1, p. 1-12, 2021.

[4] C. Lemardelé, M. Estrada, L. Pagès, et M. Bachofner, « Potentialities of drones and ground autonomous delivery devices for last-mile logistics », Transp. Res. Part E Logist. Transp. Rev., vol. 149, p. 102325, 2021.

[5] M. Salama et S. Srinivas, « Joint optimization of customer location clustering and drone-based routing for last-mile deliveries », Transp. Res. Part C Emerg. Technol., vol. 114, p. 620-642, 2020.

[6] F. Zhang et al., « Gastrointestinal tract drug delivery using algae motors embedded in a degradable capsule », Sci. Robot., vol. 7, no 70, p. eabo4160, 2022.

[7] R. De Souza, M. Goh, H.-C. Lau, W.-S. Ng, et P.-S. Tan, « Collaborative urban logistics–synchronizing the last mile a Singapore research perspective », Procedia-Soc. Behav. Sci., vol. 125, p. 422-431, 2014.

[8] G. Kovács et S. Kot, « Economic and social effects of novel supply chain concepts and virtual enterprises », 2017.

[9] D. I. Prajogo, « The strategic fit between innovation strategies and business environment in delivering business performance », Int. J. Prod. Econ., vol. 171, p. 241-249, 2016.

[10] V. L. Spiegler, M. M. Naim, D. R. Towill, et J. Wikner, « A technique to develop simplified and linearised models of complex dynamic supply chain systems », Eur. J. Oper. Res., vol. 251, no 3, p. 888-903, 2016.

[11] S. Arachchi et I. Perera, « Continuous integration and continuous delivery pipeline automation for agile software project management », in 2018 Moratuwa Engineering Research Conference (MERCon), IEEE, 2018, p. 156-161.

[12] D. Wealleans, The organizational measurement manual. Routledge, 2017.

[13] M. Hesse et J.-P. Rodrigue, « The transport geography of logistics and freight distribution », J. Transp. Geogr., vol. 12, no 3, p. 171-184, 2004.

[14] B. T. Hazen, C. A. Boone, J. D. Ezell, et L. A. Jones-Farmer, « Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications », Int. J. Prod. Econ., vol. 154, p. 72-80, 2014.

[15] Y. Yu, X. Wang, R. Y. Zhong, et G. Q. Huang, « E-commerce logistics in supply chain management: Practice perspective », Procedia Cirp, vol. 52, p. 179-185, 2016.

[16] J. An, H. Kwak, S. Jung, J. Salminen, et B. J. Jansen, « Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data », Soc. Netw. Anal. Min., vol. 8, no 1, p. 54, 2018.

[17] X. Wang, T. Kong, C. Shen, Y. Jiang, et L. Li, « Solo: Segmenting objects by locations », in Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16, Springer, 2020, p. 649-665.

[18] A. E. Ezugwu et al., « A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects », Eng. Appl. Artif. Intell., vol. 110, p. 104743, 2022.

[19] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, et B. D. Satoto, « Integration k-means clustering method and elbow method for identification of the best customer profile cluster », in IOP conference series: materials science and engineering, IOP Publishing, 2018, p. 012017.

[20] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, et J. Heming, « K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data », Inf. Sci., 2022.

[21] F. Murtagh et P. Contreras, « Algorithms for hierarchical clustering: an overview, II », Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 7, no 6, p. e1219, 2017.

[22] A. M. Ikotun et A. E. Ezugwu, « Boosting k-means clustering with symbiotic organisms search for automatic clustering problems », Plos One, vol. 17, no 8, p. e0272861, 2022.

[23] M.-Y. Tian et al., « Milk exosomes: An oral drug delivery system with great application potential », Food Funct., vol. 14, no 3, p. 1320-1337, 2023.

[24] T. Kim et H. Lee, « Delivery management system using the clustering based multiple ant colony algorithm: Korean home appliance delivery », Int. J. Ind. Eng., vol. 21, no 2, p. 1-13, 2014.

[25] T. G. Bosona et G. Gebresenbet, « Cluster building and logistics network integration of local food supply chain », Biosyst. Eng., vol. 108, no 4, p. 293-302, 2011.

[26] D. Prajapati, A. R. Harish, Y. Daultani, H. Singh, et S. Pratap, « A clustering based routing heuristic for last-mile logistics in fresh food E-commerce », Glob. Bus. Rev., vol. 24, no 1, p. 7-20, 2023.

[27] B. B. Susetyo, M. Ismail, M. I. Bagus, et F. Ramadani, « Monitoring Spatial Conditions and Road Traffic of the Tsunami Evacuation Route in Padang City Using Geographic Information System », Sumatra J. Disaster Geogr. Geogr. Educ., vol. 7, no 1, p. 24-34, 2023.

[28] J. Goldenberg et M. Levy, « Distance is not dead: Social interaction and geographical distance in the internet era », ArXiv Prepr. ArXiv09063202, 2009.