

Optimizing vehicle selection in supply chain management with data-driven strategies

Imane Zeroual, Jaber El Bouhdidi

SIGL Laboratory, ENSATE, Abdelmalek Essadi University, Tetouan, Morocco

Article Info

Article history:

Received Feb 2, 2025

Revised Jun 20, 2025

Accepted Jun 30, 2025

Keywords:

Artificial intelligence

City logistics

Decision trees

Logistic regression

Neural networks

Urban logistics

ABSTRACT

Logistics has undergone significant transformation to address the complex economic, social, and environmental challenges of the modern era. To maintain competitiveness, logistics providers have been compelled to optimize operations, meet increasing customer expectations, and improve satisfaction. Critical issues impacting logistics performance include traffic congestion, infrastructure limitations, rising demand, and the complexities of vehicle scheduling, coordination, and management. These challenges frequently disrupt delivery operations, undermining efficiency and overall system performance. This paper proposes the application of three machine learning models aimed at optimizing delivery processes, with a focus on improving vehicle assignment for order deliveries. By leveraging these models, logistics providers can enhance decision-making and operational efficiency. The study defines the core problem and evaluates several machine learning approaches to bolster logistics delivery systems.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Imane Zeroual

SIGL Laboratory, ENSATE, Abdelmalek Essadi University

Tetouan, Morocco

Email: imane.zeroual@etu.uae.ac.ma

1. INTRODUCTION

The supply chain industry encounters numerous challenges [1], with logistics being a central concern, especially for distributors and transporters. Although the logistics sector has experienced rapid growth, it has also exacerbated social and environmental issues, including severe traffic congestion and increased pollution in urban areas [2]. Addressing these challenges necessitates collaboration among all stakeholders and the implementation of effective strategies to mitigate their impacts. This paper proposes an innovative solution that leverages scientific research and artificial intelligence techniques to optimize urban freight transportation planning and infrastructure. Specifically, it introduces several machine learning models, including multinomial logistic regression, to predict the most suitable vehicle for order deliveries. Additionally, the study compares two other models to identify the best approach for the scenario. The paper emphasizes the social and environmental criticism surrounding delivery transport [3], often due to inefficient routes and lack of adaptability in planning. By focusing on selecting the appropriate vehicle for deliveries, this research aims to reduce financial burdens and environmental impacts, thereby improving the efficiency and precision of urban freight operations.

2. METHOD

2.1. Architectural overview of the proposed solution

The proposed architecture, illustrated in Figure 1, outlines several key stages for evaluating machine learning models in logistics delivery systems. These stages are designed to ensure effective data preparation,

model training, evaluation, and deployment. The architecture represents a structured pipeline that facilitates the extraction, processing, and analysis of historical data to optimize decision-making in delivery operations. Let's break down each step in detail:

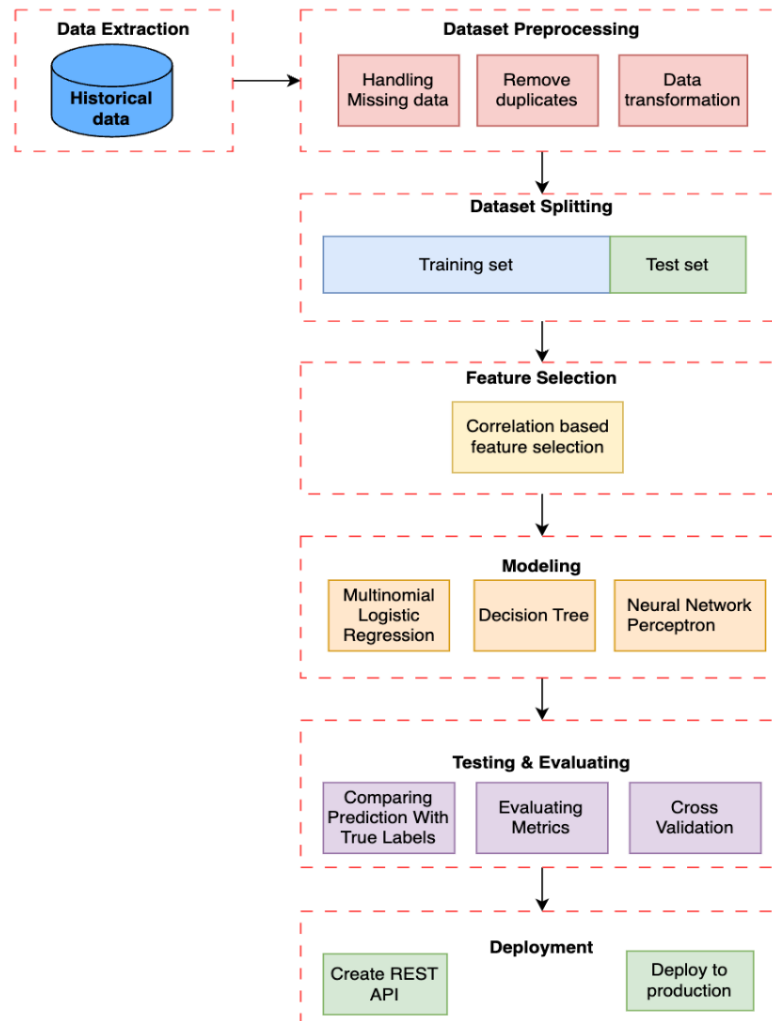


Figure 1. Machine learning-based architecture for optimizing logistics delivery systems

- a. Data extraction from historical data: The first step involves extracting relevant data from the historical logistics records database. This data serves as the foundation for model trains, including delivery times, vehicle performance, order characteristics, and other factors that influence logistics operations.
- b. Dataset preprocessing: Once the data is collected, it is processed to ensure its quality and consistency. This stage involves:
 - Handling missing data: Missing or incomplete entries are addressed through imputation or removal [4].
 - Removing duplicates: Duplicate records are eliminated [5] to prevent bias and ensure data integrity.
 - Data transformation: Such as normalization and encoding [6] is applied to make the data suitable for machine learning models.
- c. Dataset splitting: To evaluate model performance effectively, the dataset is split into two parts: a training set and a test set. The training set is used to train the machine learning models, while the test set is kept separate for evaluating the final model performance.
- d. Feature selection: Feature selection is performed to identify the most relevant features for model training. This step is applied exclusively to the multinomial logistic regression model, using correlation-based feature selection [7] to reduce dimensionality by selecting features that are most strongly correlated with the target variable (e.g., delivery duration, material shipped).

- e. Modeling: This stage involves training three different machine learning models to evaluate their performance in predicting optimal vehicle assignments for order deliveries. The models used are:
 - Multinomial logistic regression: A statistical model that predicts the probability of a vehicle being suitable for an order based on multiple features [8].
 - Decision tree: A tree-based algorithm that splits data into branches [9], predicting the best vehicle for delivery based on decision rules.
 - Neural network perceptron: A deep learning model that uses interconnected layers of neurons to learn complex patterns in the data [10].
- f. Testing and evaluating: After training the models, their predictions are tested against the actual outcomes (true labels) from the test set. The evaluation process involves:
 - Comparison of predictions with true labels: Assessing how well the models' predictions align with the actual data.
 - Evaluating metrics: Common performance metrics, such as accuracy, precision, recall, and F1-score, are used to assess model effectiveness.
 - Cross-validation: To ensure model reliability, cross-validation is performed to prevent overfitting [11] and assess the models' performance on unseen data.
- g. Deployment: The top-performing model is then deployed to a production environment, following these steps:
 - Creating a REST API: A RESTful API is developed to enable seamless communication between external systems and the model, allowing for real-time data input and instant output of predictions.
 - Production deployment: The model is integrated into the logistics system, facilitating real-time predictions and optimizing vehicle assignments for deliveries.

2.2. Data exploration and preprocessing

When an order is confirmed as feasible for delivery, it is important to define its profile. This profile serves as a detailed description of the order, helping to identify the most suitable delivery vehicle. Defining the order profile involves considering various factors, including the characteristics of the goods (such as volume, weight, and nature), the target customer, the expected delivery timeline, and the delivery typology (*i.e.*, the type of delivery service required, such as urgent, scheduled...). By understanding these factors, logistics professionals can better match the specific requirements of the order with the available vehicles during the delivery planning process. To identify the appropriate vehicle for the delivery, it is necessary to define both input and output variables that represent the relevant parameters and characteristics. These inputs form the basis for deciding which vehicle is most suited for transporting the goods. Some key parameters to consider include:

- a. Nature of the goods: The type of goods to be transported significantly impacts the vehicle selection. Certain goods, like hazardous materials, or fragile products, require specific storage, handling, and transportation conditions. A vehicle that does not meet these requirements could damage or even render the goods unfit for delivery.
- b. Delivery distance: The distance between the origin and destination directly affects vehicle capacity, fuel efficiency, and delivery time. Longer distances may require vehicles with higher fuel capacity or those more suited for long-haul trips.
- c. Delivery duration: The time required to transport the goods from the start point to the endpoint can influence the choice of vehicle. For instance, a time-sensitive delivery may require a vehicle that can travel faster or can navigate through areas with less congestion.

The input variables are the parameters and characteristics of the delivery mentioned above, while the output variable is the suitable vehicle for the delivery. The proposed model is applied to the "Delivery Truck Trip" dataset, which contains delivery trip information for each order, including the type of vehicle that is suitable for each delivery. The dataset is divided into two parts: the training set, used to train the model and minimize errors, and the test set, used to evaluate the performance of the trained model. Before feeding the dataset into the model, we noticed that some features were not normally distributed. As a result, standardization of the dataset is necessary [12]. We apply the StandardScaler preprocessing technique to standardize the data [13], using the (1):

$$x' = \frac{x - \mu}{\sigma} \quad (1)$$

Where:

x : is the training sample

σ : is the standard deviation of the training samples

μ : is the mean of the training samples

To focus on the most relevant data for our model, we selected several key columns, as shown in Table 1. These include “DeliveryID”, “BookingID”, “Date”, “Destination location”, “On-time delay”, “Trip start date”, “Trip end date”, “Transformation distance”, “Material shipped”, and “Vehicle type”. Additionally, we created a new column, “Transformation duration”, which calculates the time taken for each delivery by subtracting the “Trip start date” from the “Trip end date”. Before applying the model, we performed data cleaning by first identifying and removing any rows with missing or incomplete values. This step was crucial to ensure that the dataset used for training the model is both complete and reliable, preventing any issues that might arise from missing data [14]. Additionally, we filtered out negative values in the Transformation duration column, ensuring that it only contains valid, positive values. The “Material shipped” column consists of various categories, which we encoded into binary columns using the *OneHotEncoder* [15]. Similarly, the “Vehicle type” column includes multiple categories, which we converted to numerical labels using the *LabelEncoder* [16]. These numerical labels will be used as our target variable (y).

After cleaning and preprocessing the data, we split it into two sets: (X), which contains the feature variables (input variables), and (y), which represents the target variable “Vehicle type”. To avoid any overlap, we removed the “Vehicle type” column from (X), ensuring that it remains exclusively in the (y) set, which is used for predictions. Next, we used Scikit-learn’s *train test split* function to divide the dataset into two parts: a training set and a test set. The training set is used to fit the model, enabling it to learn the relationships between the features and the target variable. The test set is then used to evaluate the model’s performance by comparing its predictions to the actual values, providing a measure of how well the model generalizes to new, unseen data.

Table 1. Column value examples from the logistics dataset

Column name	Type	Values (example)	Description
Transportation distance	Numerical	8460.0	The distance traveled to move goods from the origin to the destination (km)
Transportation duration	Numerical	31.0	The time taken to move goods from the origin to the destination (seconds)
Material shipped	Categorical	TOOL KIT SET	The type of material that has been delivered
Vehicle type	Categorical	21MT, 40 FT 3XL Trailer 35MT	The type of vehicle used for transporting goods

3. RESULTS AND DISCUSSION

3.1. Multinomial logistic regression for delivery vehicle prediction

To make predictions, we decided to use multinomial logistic regression, a supervised machine learning model. This model is a classification technique that extends logistic regression to handle categorical dependent variables with more than two possible outcomes [17]. We selected this algorithm because our target variable, “Vehicle type”, consists of multiple classes.

After training the model, we assessed its performance by calculating the accuracy, which represents the proportion of correct predictions relative to the total predictions. In our case, the model achieved 76% accuracy on the training set and 60% accuracy on the test set. The significant difference (over 10%) between the training and test accuracies suggests potential overfitting. Overfitting occurs when the model becomes too specialized in memorizing the training data, leading to high performance on the training set but poor generalization to new, unseen data (*i.e.*, the test set). As a result, the model performs worse on the test set, which lowers the test accuracy.

3.2. Multi-layer perceptron for delivery vehicle prediction

Another machine learning algorithm we used to predict the suitable vehicle for delivery is the multilayer perceptron (MLP), a type of neural network model that is particularly well-suited for classification problems [18]. The preprocessing steps for our dataset, as described earlier, are the same as those used for the multinomial logistic regression model. For training, the MLP model utilizes two hidden layers, as illustrated in Figure 2. The first hidden layer contains three neurons, while the second hidden layer has two neurons. The algorithm processes the input data, applies transformations, and then passes the results to the output layer, which generates the final prediction.

To assess the performance of the multilayer perceptron model, we evaluated it using several classification metrics, as illustrated in Figure 3, for each of the three classes (30, 34, and 36). These metrics include precision, recall, F1-score, and support (the number of instances of each class in the dataset). These metrics provide a comprehensive understanding of how well the model predicts the correct vehicle type for delivery.

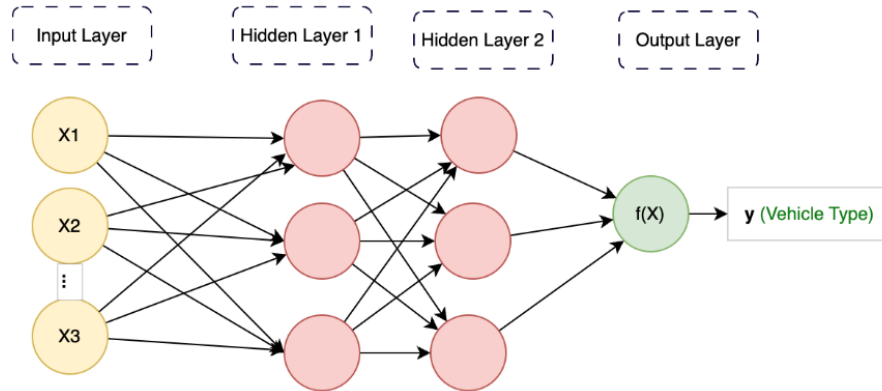


Figure 2. MLP classifier model for vehicle type prediction using transportation data

	precision	recall	f1-score	support
30	0.50	0.95	0.66	254
34	0.70	0.94	0.80	741
36	0.24	0.09	0.13	192
micro avg	0.62	0.80	0.70	1187
macro avg	0.48	0.66	0.53	1187
weighted avg	0.58	0.80	0.66	1187

Figure 3. Classification report for multi-layer perceptron

- a. Precision is the ratio of correctly predicted positive instances over the total predicted positive instance, as shown in (2):

$$\text{Precision: } \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2)$$

- b. Recall is the ratio of correctly predicted positive instances over the total actual positive instances [19], as shown in (3):

$$\text{Recall: } \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3)$$

- c. F1-score is the harmonic mean of precision and recall, and it provides a balance between precision and recall [20], as shown in (4):

$$\text{F1 - score: } \frac{2 * (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \quad (4)$$

- d. Micro average of precision, recall, and F1-score is a weighted average that accounts for class imbalance by considering the total number of instances across all classes [21]. In contrast, Macro average is an unweighted average of precision, recall, and F1-score, treating each class equally regardless of its size. In our case, the model's performance is as: weighted average precision: 0.58, recall: 0.80, and F1-score: 0.66.

These values indicate moderate performance overall. The relatively high recall of 80% suggests that the model is effective at correctly identifying positive instances. However, the lower precision of 58% indicates that only about 58% of the instances predicted as positive are correct, meaning the model has a relatively high false positive rate. For individual classes:

- Class 34 shows the highest precision and recall, indicating the model performs best for this class.
- Class 30 has the highest recall but the lowest precision, suggesting the model is more prone to false positives for this class.
- Class 36 exhibits both the lowest precision and recall, signaling poor performance for this class.

3.3. Decision tree algorithm for delivery vehicle prediction

The final model explored in this section is the decision tree, a supervised machine learning algorithm commonly used for classification tasks. This model is particularly useful for problems where the goal is to assign inputs to one of several predefined classes based on certain features [22]. We trained the decision tree on the same dataset as the previous models and generated the classification report, as shown in Figure 4. The results from the classification report, based on the test data, reveal the following performance metrics for the decision tree classifier: Precision: 66%, Recall: 82%, and F1-score: 68%.

Precision measures the accuracy of positive predictions, with the model correctly predicting 66% of positive instances. Recall reflects the model's ability to identify actual positives, correctly identifying 82% of them. The F1-score, the harmonic mean of precision and recall, is 68%, offering a balanced view of performance. While the model has high recall 82%, indicating it detects most positive instances, its lower precision 66% suggests it misclassifies some negatives as positives. In summary, the classifier has a weighted average precision of 66%, recall of 82%, and F1-score of 68%, indicating good detection of positives but with some false positives.

The three machine learning models tested in this study (multinomial logistic regression, decision tree, and multilayer perceptron) provide valuable insights into predicting the most suitable vehicle for delivery assignments. However, all models encountered challenges related to both underfitting and overfitting, largely due to the limited size and complexity of the dataset. Underfitting was observed when the models failed to capture the underlying patterns in the data [23], while overfitting occurred when the models learned the training data too well, including noise and irrelevant details [24]. This led to performance inconsistency, with the models performing significantly better on the training data than on the test data, which is a common symptom of overfitting.

To gain a more comprehensive understanding of model performance, we used precision, recall, and F1-score as additional evaluation metrics, alongside accuracy. Accuracy alone can sometimes be misleading, especially in multi-class classification tasks, where class imbalance may skew the model's true performance. To reduce overfitting, we adjusted the multinomial logistic regression model by applying cross-validation, which helps better assess its generalization to unseen data, as shown in Figure 5. Despite these adjustments, the performance of the models could be further improved by applying techniques such as hyperparameter tuning [25], ensemble learning, and training on larger, more diverse datasets. Future work could also explore methods like early stopping [26] and dropout during training [27] to enhance generalization and reduce overfitting.

	precision	recall	f1-score	support
0.0	0.94	0.78	0.85	175
19.0	0.10	0.11	0.10	28
31.0	0.40	0.97	0.56	184
34.0	0.83	0.19	0.31	105
36.0	0.66	0.95	0.78	508
micro avg	0.60	0.82	0.69	1000
macro avg	0.59	0.60	0.52	1000
weighted avg	0.66	0.82	0.68	1000

Figure 4. Classification report for decision tree model

Cross-Validation Accuracy Scores [0.63535912 0.63535912
0.63259669 0.6160221 0.62603878 0.62603878
0.63434903 0.62603878 0.62880886 0.63157895]

Figure 5. Cross-validation accuracy for multinomial logistic regression

4. CONCLUSION

The integration of artificial intelligence in logistics has the potential to significantly optimize processes by analyzing data and making real-time predictions and decisions. This capability can help carriers enhance the profitability of their operations in urban areas, improve overall performance, and ultimately increase customer satisfaction and loyalty. In this paper, we experimented with three machine learning

models to tackle a classification problem, aiming to identify the most suitable model based on various evaluation metrics. In addition to model selection, the quality of the data plays a crucial role. Effective data preprocessing such as cleaning, transformation, and preparation ensures that the data is properly structured for the algorithm, which helps improve the learning process. As a final note, future research should focus on incorporating environmental factors into logistics optimization models. This could include enriching the dataset with criteria like CO₂ emissions and fuel consumption, which are essential for promoting sustainable logistics practices.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Imane Zeroual	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓			
Jaber El Bouhdidi	✓	✓		✓	✓			✓		✓		✓	✓	

- C : Conceptualization
- M : Methodology
- So : Software
- Va : Validation
- Fo : Formal analysis
- I : Investigation
- R : Resources
- D : Data Curation
- O : Writing - Original Draft
- E : Writing - Review & Editing
- Vi : Visualization
- Su : Supervision
- P : Project administration
- Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author, IZ, on request.

REFERENCES

[1] P. Sirisawat, N. Hasachoo, and T. S. Rodbundith, "Sustainable supply chain management challenges analysis in local plastic recycling business," *Cleaner Logistics and Supply Chain*, vol. 13, p. 100188, Dec. 2024, doi: 10.1016/j.clscn.2024.100188.

[2] T. S. Tamir *et al.*, "Traffic congestion prediction using decision tree, logistic regression and neural networks," *IFAC-PapersOnLine*, vol. 53, no. 5, pp. 512–517, 2020, doi: 10.1016/j.ifacol.2021.04.138.

[3] J. Allen *et al.*, "Understanding the transport and CO2 impacts of on-demand meal deliveries: A London case study," *Cities*, vol. 108, p. 102973, Jan. 2021, doi: 10.1016/j.cities.2020.102973.

[4] X. Lai, X. Wu, L. Zhang, W. Lu, and C. Zhong, "Imputations of missing values using a tracking-removed autoencoder trained with incomplete data," *Neurocomputing*, vol. 366, pp. 54–65, Nov. 2019, doi: 10.1016/j.neucom.2019.07.066.

[5] L. Yue, "The comparison of storage space of the two methods used in SQL server to remove duplicate records," *Procedia Computer Science*, vol. 131, pp. 691–698, 2018, doi: 10.1016/j.procs.2018.04.313.

[6] M. Hu, C. Zhang, and W. Deng, "Constructed encoded data based coded distributed DNN training for edge computing scenario," *Physical Communication*, vol. 67, p. 102499, Dec. 2024, doi: 10.1016/j.phycom.2024.102499.

[7] S. Zhang, T. Wang, K. Worden, L. Sun, and E. J. Cross, "Canonical-correlation-based fast feature selection for structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 223, p. 111895, Jan. 2025, doi: 10.1016/j.ymssp.2024.111895.

[8] C. K. Gehringer, G. P. Martin, B. Van Calster, K. L. Hyrich, S. M. M. Verstappen, and J. C. Sergeant, "How to develop, validate, and update clinical prediction models using multinomial logistic regression," *Journal of Clinical Epidemiology*, vol. 174, p. 111481, Oct. 2024, doi: 10.1016/j.jclinepi.2024.111481.

[9] L. Bounia and I. Setitra, "Enhancing the intelligibility of decision trees with concise and reliable probabilistic explanations," *Data & Knowledge Engineering*, vol. 156, p. 102394, Mar. 2025, doi: 10.1016/j.datak.2024.102394.

[10] I. Betkier, "Estimating travel time in transport network with a combined multi-attributed graph convolutional neural network and multilayer perceptron model," *Engineering Applications of Artificial Intelligence*, vol. 142, p. 109898, Feb. 2025, doi: 10.1016/j.engappai.2024.109898.




[11] N. Acharki, A. Bertoncello, and J. Garnier, "Robust prediction interval estimation for Gaussian processes by cross-validation method," *Computational Statistics & Data Analysis*, vol. 178, p. 107597, Feb. 2023, doi: 10.1016/j.csda.2022.107597.

[12] I. K. Khan, H. Daud, N. Zainuddin, and R. Sokkalingam, "Standardizing reference data in gap statistic for selection optimal number of cluster in K-means algorithm," *Alexandria Engineering Journal*, vol. 118, pp. 246–260, Apr. 2025, doi: 10.1016/j.aej.2025.01.034.




- [13] R. C R and D. S. C P, "Evaluating deep learning with different feature scaling techniques for EEG-based music entrainment brain computer interface," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 7, p. 100448, Mar. 2024, doi: 10.1016/j.prime.2024.100448.
- [14] W.-T. Hong and H.-K. Yoon, "Estimating missing values in compressive strength of cementitious materials: A machine learning and statistical approach with irregular data," *Journal of Building Engineering*, vol. 101, p. 111797, May 2025, doi: 10.1016/j.job.2025.111797.
- [15] A. Alazeb, "Utilizing machine learning with unique Pentaplet data structure to enhance data integrity," *Computers, Materials & Continua*, vol. 77, no. 3, pp. 2995–3014, 2023, doi: 10.32604/cmc.2023.043173.
- [16] H. Assmi *et al.*, "A robust security detection strategy for next generation IoT networks," *Computers, Materials & Continua*, vol. 82, no. 1, pp. 443–466, 2025, doi: 10.32604/cmc.2024.059047.
- [17] X. Guo, A. Chauhan, J. Verschoor, and A. Margert, "A quality decay model with multinomial logistic regression and image-based deep learning to predict the firmness of 'Conference' pears in the downstream supply chains," *Journal of Stored Products Research*, vol. 109, p. 102450, Dec. 2024, doi: 10.1016/j.jspr.2024.102450.
- [18] N. Slimani, I. Slimani, N. Sbiti, and M. Amghar, "Traffic forecasting in Morocco using artificial neural networks," *Procedia Computer Science*, vol. 151, pp. 471–476, 2019, doi: 10.1016/j.procs.2019.04.064.
- [19] R. Salles *et al.*, "SoftED: Metrics for soft evaluation of time series event detection," *Computers & Industrial Engineering*, vol. 198, p. 110728, Dec. 2024, doi: 10.1016/j.cie.2024.110728.
- [20] W. Jia, Y. Qin, and C. Zhao, "Rapid detection of adulterated lamb meat using near infrared and electronic nose: A F1-score-MRE data fusion approach," *Food Chemistry*, vol. 439, p. 138123, May 2024, doi: 10.1016/j.foodchem.2023.138123.
- [21] A. Berger and S. Guda, "Threshold optimization for F measure of macro-averaged precision and recall," *Pattern Recognition*, vol. 102, p. 107250, Jun. 2020, doi: 10.1016/j.patcog.2020.107250.
- [22] A. Argente-Garrido, C. Zuheros, M. V. Luzón, and F. Herrera, "An interpretable client decision tree aggregation process for federated learning," *Information Sciences*, vol. 694, p. 121711, Mar. 2025, doi: 10.1016/j.ins.2024.121711.
- [23] M. Czajkowski and M. Kretowski, "Decision tree underfitting in mining of gene expression data. An evolutionary multi-test tree approach," *Expert Systems with Applications*, vol. 137, pp. 392–404, Dec. 2019, doi: 10.1016/j.eswa.2019.07.019.
- [24] J. L. Andrews, "Addressing overfitting and underfitting in Gaussian model-based clustering," *Computational Statistics & Data Analysis*, vol. 127, pp. 160–171, Nov. 2018, doi: 10.1016/j.csda.2018.05.015.
- [25] R. El Shawi, M. Bahman, and S. Sakr, "To tune or not to tune? An approach for recommending important hyperparameters for classification and clustering algorithms," *Future Generation Computer Systems*, vol. 163, p. 107524, Feb. 2025, doi: 10.1016/j.future.2024.107524.
- [26] S. Bös, "Statistical mechanics approach to early stopping and weight decay," *Physical Review E*, vol. 58, no. 1, pp. 833–844, Jul. 1998, doi: 10.1103/PhysRevE.58.833.
- [27] H. Wu and X. Gu, "Towards dropout training for convolutional neural networks," *Neural Networks*, vol. 71, pp. 1–10, Nov. 2015, doi: 10.1016/j.neunet.2015.07.007.

BIOGRAPHIES OF AUTHORS



Imane Zeroual    She is a data engineer currently pursuing a Ph.D. in computer science at the Research Laboratory of Information Systems and Software Engineering (SIGL), University Abdelmalek Essaadi, National School of Applied Sciences-Tetouan, Morocco. She holds an engineering degree in information systems and decision support from the National School of Applied Sciences in Morocco and a master's degree in big data and machine learning from the University of Lille, France. Her research interests include machine learning, internet of things (IoT) applications, big data analytics, data engineering, and scalable data processing. She can be contacted at: imane.zeroual@etu.uae.ac.ma.



Jaber El Bouhdidi    is an HDR Professor of Computer Science at the SIGL Laboratory, University Abdelmalek Essaadi, National School of Applied Sciences-Tetouan-Morocco. His research interests include web semantic, multi-agents' systems, e-learning adaptive systems and big data. He has several papers in international conferences and journals. He can be contacted at: jaber.elbouhdidi@uae.ac.ma.