Transportation Research Interdisciplinary Perspectives 24 (2024) 101052

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

Sensitivity evaluation of machine learning-based calibrated transportation mode choice models: A case study of Alexandria City, Egypt

Ahmed Mahmoud Darwish [a,*,1], Mohamed Almansour [b], Ayman Salah [a], Maged Zagow [c], Khaled Saeed [d], Ahmed Elkafoury [e]

[a] Transportation Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, Egypt
[b] University of Missouri, Columbia
[c] Department of Architecture Engineering, Faculty of Engineering, Tanta University, Tanta 3111, Egypt
[d] Civil Engineering Department, Faculty of Engineering, Aswan University, Egypt
[e] Department of Public Works Engineering, Faculty of Engineering, Tanta University, Tanta 3111, Egypt

ARTICLE INFO

Keywords:
Transport Mode Choice
Machine Learning (ML)
Multinomial Logit (MNL)
Revealed Preference (RP) Survey
Biogeme Software

ABSTRACT

Intelligent methods including Machine Learning (ML) techniques have been increasingly employed in transportation mode choice modeling, which is more complex than other demand models, since it has to reliably and accurately reflect a wide range of related categorical and continuous variables, concerning the travelers, transportation system, and trip characteristics. ML techniques can capture such complex relationships. So that, they can provide a more nuanced understanding of the travelers' decision process. Most research studies focused mainly on the evaluation of the model accuracy, where little has been done to evaluate the models' performance toward transport attributes. This research aims to calibrate the transportation choice models using ML techniques, then conduct a comparison with the Multinomial Logit (MNL) model to identify the impact on the model accuracy and performance and quantify to what extent the ML models are sensitive to transport policies compared to the traditional MNL model. To this end, eight ML classifiers were examined. As a case study, the models were calibrated to reflect the choice behavior of trip makers in Alexandria City, Egypt. The models were successfully calibrated with satisfying accuracy; however, the ML models have better calibration results in terms of predictability, outperforming the MNL model, where the GBDT classifier records the best prediction accuracy. Finally, the sensitivity analysis test was performed to quantify the elasticity of the models to transport policies. The results show the ML models' structure is more comprehensively and accurately built than the MNL model providing better indicative and reliable sensitivity analysis results.

1. Introduction

Calibration of transportation mode choice models is necessary for any area with satisfying accuracy and reasonable sensitivity to travel attributes that affect travelers' choice of the preferred transportation system. They help in deciding the appropriate transportation policies in any area whether urban or inter-urban. Mode choice significantly affects travel efficiency, spaces used, and availability of transportation choices (Heiss, 2016).

Over the last decades, several discrete choice models have been employed in different areas to reflect the determinants of transportation choice (Chu, 2018; Tan and Ma, 2021). Transportation applications utilized these models due to their ability to represent the fundamental factors that influence humans' decision-making, enable the estimate of parameters, and establish a defensible behavioral foundation, such as the principle of Random-Utility Maximization. Many discrete choice logit models are employed in the analysis of transportation mode selection, such as Binary Logit (BL), Multinomial Logit (MNL), Nested Logit (NL), Cross Nested Logit (CNL), and Mixed Multinomial Logit (MXL) models (Lu et al., 2015; Willumsen and Ortúzar, 2016; Newman et al., 2018).

Artificial Intelligence (AI) techniques are prominent approaches that

* Corresponding author.
E-mail addresses: Adarwish@alexu.edu.eg (A.M. Darwish), maay8f@umsystem.edu (M. Almansour), ayman.salah@alexu.edu.eg (A. Salah), mzagow@f-eng.tanta.edu.eg (M. Zagow), K.saeed@aswu.edu.eg (K. Saeed), kaforiengin@f-eng.tanta.edu.eg (A. Elkafoury).
[1] ORCID: 0000-0002-4041-932X. ResearcherID: GLN-3253-2022.

https://doi.org/10.1016/j.trip.2024.101052
Received 24 August 2023; Received in revised form 15 February 2024; Accepted 25 February 2024
Available online 10 March 2024

have been used nowadays to represent intelligent systems. They are widely applied in developing various travel demand models. They proved that they can capture complex relationships over the other traditional methods and provide superior predictability of mode choice models (Omrani, 2015; Sarker, 2022). One of the popular AI techniques is Machine Learning (ML).

In transportation choice modeling, the ML method is generally based on developing a model that comprehensively reflects the underlying structure of the entire data without pre-making assumptions of the data structure through adopting a fitting approach. For example, in a Decision Tree (DT) analysis, the algorithm recursively splits the dataset based on various features or attributes to create branches or nodes, aiming to maximize the homogeneity or purity of the samples. They provide a single model that accounts for different impeded segmented categories over the traditional logit models. Thus, they do not consume a huge time compared to logit models that apply an iterative process method to provide statistically significant parameters (Hussain et al., 2017; Richards and Zill, 2019; Abulibdeh, 2023).

ML models offer several advantages in modeling the choice of transportation mode. First, they can handle large and diverse datasets, facilitating the incorporation of a diverse array of factors that influence mode choice; thus, ML models have the potential to offer enhanced precision and dependability in contrast to conventional choice models. (Bhavsar et al., 2017). Second, ML models can capture non-linear relationships and interactions among variables, which are often present in transportation mode choice behavior. They can detect complex patterns that may not be readily apparent using traditional statistical techniques, allowing for a more nuanced understanding of the decision-making process (Bishop, 2006). Third, ML models have the potential to improve prediction accuracy by learning from large datasets and adapting to changing conditions. As transportation systems evolve emerged modes of transport, ML models can incorporate real-time data to update their predictions and reflect the dynamic nature of mode choice behavior (Abduljabbar et al., 2019; Khan et al., 2022).

ML techniques are introduced to model and formulate the choice of transportation mode. They work effectively in learning the data to reflect the complex relationships without making drastic assumptions about the data. They have been successfully applied in various transportation mode choice studies worldwide. These models have been used to explore the impact of factors such as travel time, travel cost, accessibility, land use patterns, built environment characteristics, and individual preferences on mode choice decisions (Shmueli et al., 1996). They have also been employed to forecast travel demand, evaluate the effectiveness of transportation policies, and support decision-making processes in urban planning and transportation infrastructure development (Xie et al., 2003).

The review of previous studies reveals that most research studies focused mainly on the evaluation of the model accuracy (e.g., prediction capability), where little has been done in the evaluation of the performance of the models in terms of the model sensitivity to transport attributes. This is essential to quantify to what extent the ML models are sensitive to transport policies compared to the traditional MNL model.

Thus, this research aims to develop transportation mode choice models using ML techniques and conduct a comparison with the Multinomial Logit (MNL) model to identify the impact of using ML on the model's accuracy and performance. To this end, the research tends to achieve the following objectives:

- Calibrating transportation mode choice models using various ML techniques.
- Developing a traditional choice model using the Multinomial Logit (MNL) method.
- Estimating the prediction accuracy of each method.
- Conducting a comparison with the Multinomial Logit (MNL) model to identify the impact of using ML on the model accuracy.
- Conducting a sensitivity analysis test to quantify the elasticity of each method to transport policies and assess the performance of each method from a practical point of view.

The structure of this research study is arranged in the following manner: following this introduction, Section 2 provides a comprehensive overview of the existing literature. Section 3 provides a comprehensive background about transportation mode choice modeling for both traditional MNL logit methods and ML techniques. Section 4 provides an overview of the research study area and the methods employed for data collection as well as the data analysis. In Section 5, the Multinomial Logit model's development is presented. Section 6 encompasses the basis of developing the ML models and discusses their main outputs. The comparison between the developed models is presented in detail in Section 7. The findings and conclusions of this study are presented in Section 8.

2. Literature review

Several studies have been prepared in different areas around the world that used machine learning techniques in modeling the choice of transportation mode. This section presents a brief literature review of research studies conducted either in Egypt or other countries. Table 1 shows a summary of the previous research studies. It clarifies the location, year, sample size, model type, and model accuracy.

In a study conducted by (Zhang and Xie, 2008), the use of the Support Vector Machine (SVM) classifier was investigated with the aim of calibrating a mode choice model in the San Francisco Bay Area, located in the United States of America. The analysis findings suggested that the utilization of the Support Vector Machine (SVM) model yields superior predictive accuracy compared to the conventional Multinomial Logit (MNL) model. The model prediction accuracy increased from 77.8 % using the SVM method compared to 78.4 % using the traditional MNL model.

(Omrani, 2015) applied three machine learning methods to calibrate individuals' choice models in Luxembourg City, Luxembourg. These methods were SVM, Multilayer Perceptron (MLP), and Radial Basis Function (RBF) neural network. The models' performance was compared with the MNL model. The findings indicated that ML methods achieved superior performance compared to the MNL model. Among the developed models, the MLP was the most accurate model with a prediction accuracy of 81 % compared to 64.7 for the MNL model.

(Sekhar et al., 2016) used the Random Forest (RF) method to model the choice of transportation mode in Delhi, India. A comparison was conducted using the MNL model. The results indicated that using RF highly enhanced the prediction accuracy of the choice model, where the model predicted correctly around 99 % of choices compared to 77 % for the MNL.

(Wang and Ross, 2018) evaluated the application of the Extreme Gradient Boosting (XGB) method in modeling transportation mode choice in Pennsylvania and New Jersey, USA. In the research, the prediction accuracy of the XGP is compared with the traditional MNL model. The results displayed that the XGB model achieved a prediction accuracy of 84.5 % compared to 81.70 % for the MNL model.

(Cheng et al., 2019) employed three ML classifiers (i.e., RF, SVM, and Adaptive Boosting (AdaBoost)) to simulate the selection of transportation modes in Nanjing, China. The models' prediction capability was compared with the traditional MNL model. The results indicated that the ML methods achieved superior performance compared to the MNL model, achieving the best accuracy for the RF method with a prediction accuracy of 85.36 % compared to 63.02 for the MNL Model.

2.1. Review of previous research studies in Egypt

Only one study was conducted in Egypt by (Elharoun, El-Badawy and Shahdah, 2023) in Mansoura City that applied Artificial Intelligence (AI)

Table 1
Review of previous studies.

SN	Reference	Location	Year	Discrete Choice Models	Model Prediction Accuracy %	ML models	Model Prediction Accuracy %
1	(Zhang and Xie, 2008)	San Francisco Bay Area, USA	2008	MNL	77.8	SVM	78.4
2	(Omrani, 2015)	Luxembourg City, Luxembourg	2015	MNL	64.68	MLP, RBF, and SVM	81 (MLP)
3	(Sekhar, Minal and Madhu, 2016)	Delhi, India	2016	MNL	77.31	RF	98.96
4	(Wang and Ross, 2018)	Pennsylvania and New Jersey, USA	2018	MNL	81.70	XGB	84.50
5	(Cheng et al., 2019)	Nanjing, China	2019	MNL	63.02	RF, Adaboost, and SVM	85.36 (RF)
6	(Elharoun, El-Badawy and Shahdah, 2023)	Mansoura City, Egypt	2023	MNL	82.55	DNN, KNN,Adaboost, SVM,NB,RF, GB, XGB,LDA,DT,and,QDA.	72.61 – 96.99 (DNN)
7	(El Esawey and Ghareib, 2009)	Cairo, Egypt	2009	MNL	NA		
8	(El-Bany et al., 2014)	Port Said, Egypt	2014	MNL, NL, and MXL	NA		
9	(Elharoun, Shahdah and El-Badawy, 2018)	Mansoura, Egypt	2018	MNL	85.1		
10	(Abdel-Aal, 2017)	Alexandria, Egypt	2017	BL	NA		
11	(Darwish, 2021)	Alexandria, Egypt	2021	MNL	60.2 ~ 90.2		

NA: Not Available.
MLP: Multilayer Perceptron; XGB: Extreme Gradient Boosting;
DNN: Deep Neural Network; GB: Gradient Boosting;
DT: Decision Tree; LDA: Linear Discriminant Analysis;
SVM: Support Vector Machine; KNN: K-Nearest Neighbor;
RBF: Radial Basis Function; QDA: Quadratic Discriminant Analysis Classifier;
NB: Naive Bayes; MNL: Multinomial logit Model;
RF: Random Forest; NL: Nested Logit;
Adaboost: Adaptive boosting; MXL: Mixed Multinomial Logit.

methodologies to forecast individuals' selection of transportation modes. The study examined the use of 10 different machine learning classifiers in addition to the Deep Neural Networks (DNNs) to conduct a benchmarking study with the traditional MNL model. The results indicated that the DNN scored the best performance with a prediction accuracy of 97.81 % followed by Gradient Booster (GB), RF, Decision Tree (DT), and XGB with a prediction accuracy of around 95 % compared to 82.55 for the MNL model.

(El Esawey and Ghareib, 2009) addressed the transport mode choice in the Greater Cairo Region (GCR), one of the most densely populated metropolitan areas in Africa and the Middle East. The study employed a multinomial logit (MNL) model on different subsets of the data based on trip purposes, leveraging the extensive dataset collected by the Cairo Regional Area Transportation Study (CREATS) conducted in the GCR in 2001. Experiments of this study revealed that high-income and female individuals are more likely to use private transportation modes such as private cars. Besides, the travel cost was shown to be a significant factor in choosing the mode of transportation.

(El-Bany et al., 2014) addressed the transportation choice in the city of Port Said by collecting the data using a Stated Preference (SP) questionnaire where the respondents chose between a private car, taxi, and bus rapid transit (BRT), where BRT is a hypothetical public service. This study used three discrete choice models: MNL, Nested Logit Model (NL), and Mixed Logit Model (MXL). Results revealed that income is the most effective factor in mode choice, where high-income individuals are more inclined towards cars than they are to buses and taxis. In addition, out-of-vehicle trip time was found more effective than in-vehicle trip time, which is rather unusual in most developing countries. Finally, reductions in travel costs and trip speed were found when using the new policy of BRT.

(Elharoun, Shahdah and El-Badawy, 2018) analyzed the behavior of travelers in choosing the transportation mode in Mansoura City. The collected dataset included 30,000 trips gathered through an online questionnaire, comprising five primary kinds of transportation often utilized inside urban areas conducting a sensitivity test of the dataset size vs. the model accuracy. The study showed that total trip time, trip cost, and car ownership are among the top factors affecting the MNL model decision.

(Abdel-Aal, 2017) developed a Binary Logit (BL) model for Alexandria city in Egypt to estimate the time value focusing on the behavior of travelers in choosing transportation mode. The data sample was collected between the spring of 2002 and the beginning of 2003 and included 2,366 trips. The transportation modes were combined into two alternatives: AUTO (car driver, car passengers, in addition to taxi) and TRANSIT (bus, tram, train, microbus, school bus, as well as work bus). Out of the estimated 20 models, only two models were considered successful regarding the parameters' sign and the magnitude of their significance (t-statistics value).

(Darwish, 2021) developed a transport mode choice model for Alexandria City to assess the impact of implementing transportation policies and measures on the travelers' behavior regarding selecting their transportation mode and develop a significant modal shift to public transport. The methodology was arranged to calibrate a separate MNL model with homogeneous preferences for each economic group. The results showed that the developed models were successfully calibrated. Each economic-based segmentation significantly improved the model accuracy, where the prediction capability increased from 60.2 % in the unified model to 90.2 % in the segmented model. However, searching for the most suitable and accurate segmented models consumed a huge time.

3. Transportation mode choice modelling

The field of transportation plays a vital role in our daily lives, facilitating the movement of people and goods. Understanding the factors influencing individuals' choice of transportation mode is important for transportation planners and policymakers to develop efficient and sustainable transportation systems. Recently, machine learning (ML) models have emerged as powerful tools for modeling and predicting human behavior, including transportation mode choice (Ortúzar and Willumsen, 2011).

The Multinomial Logit (MNL) models are commonly employed in the analysis of discrete choice behavior, which are characterized by their extensive usage and popularity due to their simple calibration and applicability.(O'Flaherty, 2018; Migliore and Ciccarelli, 2020; Harz and Sommer, 2022). The model structure is as follows (Elharoun, Shahdah and El-Badawy, 2018):

$$P_{ni} = \frac{e^{(V_{ni})}}{\sum e^{\left(V_{nj}\right)}} j \in Cn \tag{1}$$

Where:

V_{ni} is the deterministic part of the utility function of mode i for individual n;

V_{nj} is the deterministic part of the utility function of any mode j from the available transportation modes for an individual n;

P_{ni} is the probability of selecting mode i by an individual n among transportation modes; and

C_n is the set of all available transportation modes for an individual n.

Machine learning models have revolutionized various domains through their ability to analyze complex patterns and make accurate predictions (Bishop, 2006). These techniques leverage large datasets, including sociodemographic, spatial, and temporal variables, to uncover relationships and capture intricate decision-making processes related to transportation choices (Xie et al., 2007).

Machine learning algorithms may be classified into several categories. Supervised learning algorithms are designed to provide a mapping function that relates input data to corresponding target output data. Unsupervised learning involves the utilization of algorithms to construct models based on input data, without the presence of any predetermined intended output. The primary objective of this approach is to identify and discern patterns or clusters within the given dataset. Semi-supervised learning is an approach that integrates aspects of both supervised and unsupervised learning, leveraging both labeled and unlabeled data. Reinforcement learning is a machine learning algorithm that acquires knowledge by engaging in interactions within a given environment and afterward receives feedback about the efficacy of its actions. Inductive learning pertains to machine learning algorithms that acquire information and provide predictions by using past experiences and inherent biases (Khan et al., 2010).

In this research, eight classifiers, used in the literature in transportation choice modeling using ML techniques, were examined: Support Vector Machine (SVM); Multi-layer Perceptron (MLP); K-Nearest Neighbor (KNN); Adaptive Boosting (AdaBoost) Decision Tree (DT); Random Forest (RF), Gradient-Boosted Decision Trees (GBDT), and Extreme Gradient Boosting (XGB) (Arabameri et al., 2020; Sun et al., 2022; Elharoun, El-Badawy and Shahdah, 2023).

Support Vector Machines (SVM) algorithms are commonly used for binary classification tasks (Zhang and Xie, 2008). The Support Vector Machine (SVM) is a supervised learning method that uses a risk-based approach to categorize data patterns (Cortes and Vapnik, 1995). It establishes a decision boundary that maximizes the margin between examples belonging to the same class (Bhavsar et al., 2017). They have achieved notable success in pattern recognition applications (Ben-Hur et al., 2001). The SVM optimization problem is typically solved using quadratic programming techniques or optimization algorithms that handle convex optimization problems (Lasantha et al., 2023).

Multi-layer Perceptron (MLP) is a widely employed kind of Artificial Neural Network (ANN) that finds application in the field of transportation mode choice modeling (Assi et al., 2019). Every individual neuron receives input from the preceding layer, applies a nonlinear activation function to this input, and subsequently transmits the resulting output to the subsequent layer. This process allows the MLP to learn and extract high-level features from the input data, facilitating the modeling of mode choice behavior.

K-Nearest Neighbor (KNN) algorithm is a non-parametric and instance-based approach in machine learning. It is utilized for classifying new instances by measuring their closeness to existing examples inside the training dataset (Rahman et al., 2022). KNN operates by calculating the distances between instances in the feature space. These instances represent individuals or trips with associated characteristics and mode choice labels. KNN finds the k nearest neighbors to an instance based on the chosen distance metric and assigns a mode choice label to the new instance based on the main vote of the k nearest neighbors.

Adaptive Boosting (AdaBoost) is a machine learning ensemble algorithm that is commonly used in transportation mode choice modeling (Freund and Schapire, 1996). AdaBoost combines the predictions of multiple weak classifiers to create a strong and accurate model. AdaBoost operates by iteratively training weak classifiers on different subsets of the training data. Each weak classifier is a simple model, such as a decision tree or a logistic regression model, that performs slightly better than random guessing. During training, AdaBoost assigns higher weights to instances that are misclassified by previous weak classifiers, thereby focusing on the difficult examples and improving performance (Cheng et al., 2019).

Decision tree (DT) can be considered a powerful non-parametric algorithm that is applied in classification tasks. It is structured like flowcharts or trees, where nodes represent decisions and branches depict possible outcomes. DT is easily interpretable and requires minimal data preparation. However, they are prone to overfitting and sensitive to noisy data (Breiman, 2001; Bhavsar et al., 2017; Hagenauer and Helbich, 2017). Ensemble methods such as bagging and Random Forest (RF) have further enhanced DT's performance. Bagging involves training multiple DTs in parallel using bootstrap samples, while RFs employ random feature subsets for each split. Class assignments are determined by majority voting among the ensemble (Xie et al., 2003; Zhan et al., 2016).

The Random Forest (RF) technique is widely employed in transportation mode choice modeling because of its popularity in the field of machine learning (Hagenauer and Helbich, 2017). It is a collective method that combines many decision trees to create a robust and accurate model. The process involves the construction of several decision trees on various subsets of the training dataset via the Random Forest (RF) algorithm, and during prediction, the mode choice label is determined by combining the forecasts of all the individual trees. RF is robust against overfitting due to its ensemble nature. By constructing multiple decision trees on different data subsets, RF reduces overfitting risks by averaging biases and errors (Breiman, 2001).

Gradient-Boosted Decision Trees (GBDT) classifier combines the predictive strength of decision trees with gradient boosting, an ensemble learning technique, resulting in a robust and accurate model. In transportation modeling choice, GBDT constructs decision trees sequentially, with each tree aiming to rectify the mistakes made by its predecessors. GBDT classifier builds the ensemble by iteratively correcting the mistakes of the previous models (Breiman, 2001). GBDT can handle complex relationships in the data by combining weak learners (decision trees) in an additive manner.

Extreme Gradient Boosting (XGB) is a sophisticated machine learning method that is extensively employed in the field of transportation mode choice modeling. It is an optimized implementation of the gradient-boosting framework that leverages parallel processing and regularization techniques to deliver exceptional performance and accuracy (Wang and Ross, 2018). In the context of transportation modeling choice, XGB operates similarly to gradient-boosted decision trees by iteratively building a group of decision trees, which corrects the mistakes of previous trees. However, XGB introduces several enhancements to further improve the model's predictive power and efficiency.

4. Study area and data collection

4.1. Study area

Alexandria, the second biggest metropolitan region in Egypt behind Cairo, stands as one of the nation's prominent economic centers, boasting a population of over 5 million individuals as of 2021 (CAPMAS, 2021), and is expected to have 6.3 million by 2030 (United Nations, 2018). Fig. 1 shows the location of Alexandria City in Egypt. The current transportation system is unsustainable and inadequate to accommodate the transportation demand. The dramatic growth of motorized private travel demand in recent decades has given rise to several transportation challenges, notably the issue of traffic congestion (in terms of huge delays and long queues) and negative environmental impacts.

Alexandria is situated in the northern region of Egypt. The region under consideration is geographically demarcated by the Mediterranean Sea to the north and Maryout Lake, together with fertile agricultural area, to the south. The urban development of land use in Alexandria shows that the urban physical development is strongly influenced by its boundaries. The old city' width is limited to an average of 1 to 5 km, so the city development was initially to the east, while it turned to the west later to accommodate the growth of housing and industrial needs. Hence, the urban landscape predominantly exhibits a coastal strip configuration, wherein urbanization is primarily concentrated in the northwestern and southeastern regions. A huge number of developments were constructed during this period with a sprawling suburban pattern, and this resulted in a considerable shift from the central districts to the eastern and western areas of Alexandria (Darwish, 2021).

Alexandria's urban road network is developed with the same strip shape as the city. Five main arterial roads join the east and west areas with the city center (refer to Fig. 2), called Abu Qir Road, Mahmoudiya Road, Coastal Road, Al Cornish Road, and El Max Street. Public transportation combines three systems (refer to Fig. 3): Tram (City and Raml lines); Railway (Abu Qir and Borg El-Arab Lines) and Bus lines. In addition, private vehicles with a maximum of 14 seats are operated as collective taxis covering mostly the whole road network.

4.2. Data collection

For developing the mode choice model, a Revealed Preference (RP) survey was conducted to capture the travel behavior determinants in the study area.

The required sample size is estimated based on the empirical equation provided by Levy and Lemeshow (Levy and Lemeshow, 2013), as shown in Equation 1:

$$n = \frac{Z^2 xpx(1-p)}{E^2 + \frac{Z^2 xpx(1-p)}{N}} \tag{23}$$

Where:

- n is the sample size;

Fig. 1. Location of Alexandria City in Egypt.

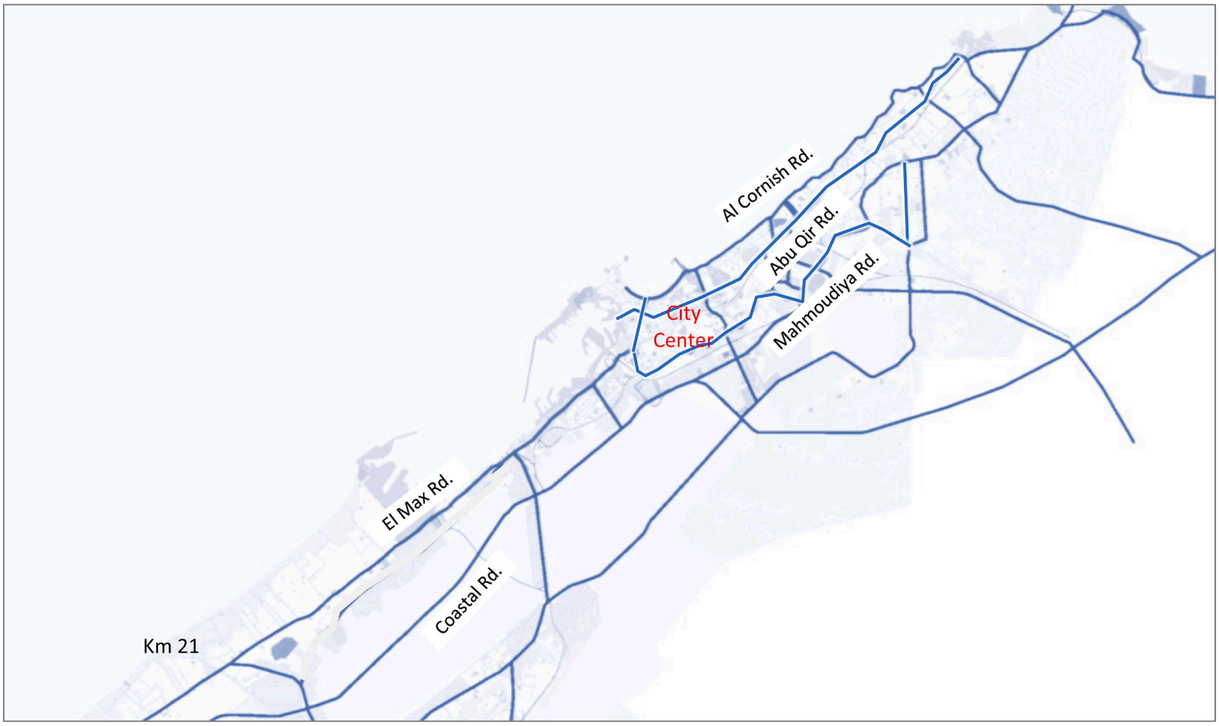

Fig. 2. Alexandria City Primary Road Network.

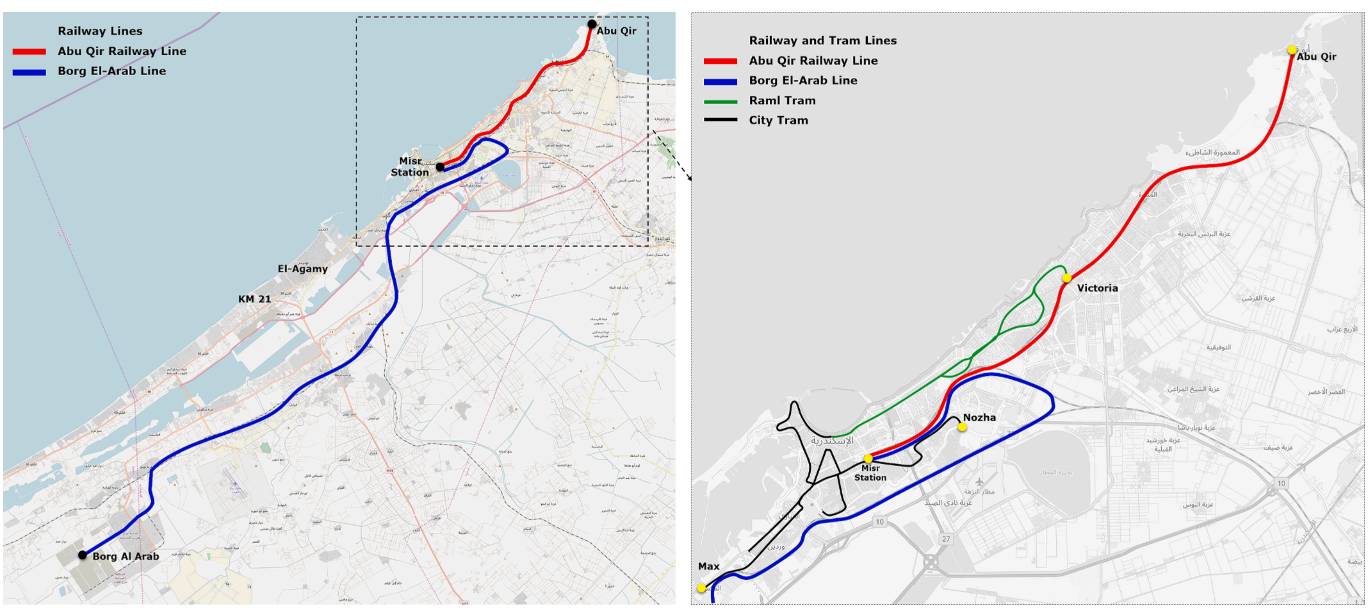

Fig. 3. Public transportation network in Alexandria City.

- P is the condition occurrence percentage, which can be considered 50 % according to (Taherdoost, 2017);
- E is the percentage of the maximum acceptable error, which can be considered 5 %;
- Z is the value for the determined confidence level, where the Z value is 1.96 for a 95 % confidence level; and
- N is the entire population

After applying Equation 1 using the above considerations, the minimum required sample size for Alexandria City is 385 observations. In this research, the data sample was collected between February 2023 and April 2023, where the dataset reached 2,498 successfully collected responses.

Table 2 shows the descriptive statistics of the collected samples where the main indicators are as follows:

- 53 % of the interviewees were male and 47 % were female.
- Work and educational trips constitute about 61 % of all trips.
- 82 % of the individuals do not own a private car, while 18 % own one or more cars.
- 54 % of the trips were conducted by the interviewees every day.
- The modal split between the four main transport modes is 7 %, 6 %, 61 %, and 27 % for private car, taxi, informal shared taxi, and public

Table 2
Descriptive statistics of the collected samples.

Criterion	Category	Frequency	Approximate Percentage (%)
Age	<20	365	15
	20–40	1,333	53
	40–60	671	27
	> 60	129	5
Gender	Male	1,313	53
	Female	1,185	47
Family size	1–2	375	15
	3–5	1,941	78
	>6	182	7
Trip Purpose	Work	776	31
	Education	740	30
	Shopping and Leisure	271	11
	Return Home	408	16
	Others	303	12
Car Ownership	0	2,042	82
	1	390	16
	≥2	66	3
Trip Frequency	Every working day	1,340	54
	Once-Twice a week	484	19
	Otherwise	674	27
Service quality	Good	479	19
	Fair	1,205	48
	Bad	814	33
Chosen Mode	Private Car	173	7
	Taxi	138	6
	Informal Shared Taxi	1,523	61
	Public Transport	664	27

transport modes, respectively. The trips conducted by other modes are minor, hence, neglected.

The samples obtained in the stated preference survey are subjected to analysis in order to determine the mean travel characteristics. These mean travel attributes are essential for calculating the qualities of the alternatives that were not chosen. Table 3 illustrates the average values of in-vehicle travel time (min.), out-of-vehicle travel time (min.), travel cost (Egyptian Pound "EGP"), and trip distance (km). The table indicates the following:

- The minimum travel distance and time are for taxi mode due to its high fare rate.
- The minimum travel cost is for public transport despite that its average travel time exceeds other modes' values indicating that most public transport users are not from the high-income level category.
- Since it is mostly available everywhere and has, a relatively low-cost rate in line with moderate travel time, compared to other modes, the informal shared taxi has become the dominant transport mode in Alexandria.

Table 3
Average travel attributes of the collected samples.

Attributes\Modes	Private Car	Taxi	Informal Shared Taxi	Public Transport
In-vehicle Travel Time (min.)	45.3	33.2	61.6	49.3
Out-of-vehicle Travel Time (min.)	8.1	14.3	35.6	28.5
Travel Cost (EGP)	34.5	30.8	3.4	6.9
Trip Distance (km)	17.7	11.4	12.1	12.4

5. Development of Multinomial Logit model

First, an initial model that includes all variables is calibrated to identify the significant attributes, where the non-significant attributes are excluded from the model. After excluding them, the remaining considered attributes of utility function include in-vehicle travel time (*IVTT*) expressed in minutes, out-of-vehicle travel time (*OVTT*) expressed in minutes, travel cost (*TC*) expressed in EGP, and mode-specific constants for the taxi, public transport, and informal shared taxi modes. The private car transportation mode is considered the reference mode. The model parameters are shown in Table 4. Travel time and travel cost parameters should have negative signs while mode-specific constants may have positive or negative signs. The model development is performed by splitting the dataset into a calibration set and a validation set of sizes of 85 % and 15 % of the total experiment dataset size, respectively.

The calibration procedure is executed with the utilization of Biogeme software developed by (Bierlaire, 2016). The program has two primary components: the model specification file and the data file. The data is prepared for the input file, which includes both the selected alternative and the non-selected alternatives. Additionally, the alternatives encompass factors such as service availability, in-vehicle travel time, out-of-vehicle travel time, and trip cost. The survey findings (Table 5) provide data on the average journey time rate and travel cost rate per distance. These rates are utilized to estimate the qualities of the alternatives that were not picked, taking into account the trip distance.

Upon completion of the preparation of the Biogeme model specification file and data file, the program executes the necessary runs, resulting in the acquisition of parameter values, which are subsequently displayed in Table 6. The computed parameters indicate that the constants associated with all modes have negative signs, suggesting that the private automobile (used as the reference mode) exhibits superior perception compared to other modes. The informal shared taxi is perceived more favorably compared to public transit and traditional taxi options. All parameters exhibit the anticipated logical signs. The negative signals of journey time and travel cost parameters are evident in this context.

All parameters exhibit statistical significance at the 5 % level of significance, as shown by the absolute value of the t-statistic for each parameter exceeding 1.96 and the corresponding p-value is less than 0.05. The likelihood ratio index ($\rho 2$) has a value of 0.311, while the adjusted likelihood ratio index ($\hat{\rho}$ 2) has a value of 0.309. These values are deemed to be adequate.

The calculation of the predictive capability of the calibrated model is performed based on the estimated parameters. The utilities associated with the four transportation modes are computed for the testing set of the entire sample. Next, the likelihood of selecting each method of transportation is computed, and the mode with the highest probability is regarded as the chosen mode. Subsequently, the mode that has been selected is compared to the initially selected mode. The aforementioned

Table 4
MNL model parameters.

Parameters	Nature	Expected Sign
In-vehicle travel time (β_{IVTT})	value	Negative
Out-of-vehicle travel time (β_{OVTT})	value	Negative
Travel Cost (β_{TC})	value	Negative
Taxi Constant (*ASC2*)	value	Negative or Positive
Public Transport Constant (*ASC3*)	value	Negative or Positive
Shared Taxi Constant (*ASC4*)	value	Negative or Positive

Table 5
Average rates of travel attributes per distance of travel.

Attributes	Private Car	Taxi	Public Transport	Informal Shared Taxi
In-vehicle Travel Time (Min./km)	2.6	2.9	5.1	4.0
Out-of-vehicle Travel Time (Min./km)	0.5	1.3	2.9	2.3
Travel Cost (EGP/km)	1.9	2.7	0.3	0.6

Table 6
MNL model calibration results.

Parameter	Value	Standard error	*t*-test	p-value
ASC1	–fixed–			
ASC2	−2.87	0.126	−22.72	< 0.01
ASC3	−2.15	0.135	−15.92	< 0.01
ASC4	−1.35	0.125	−10.77	< 0.01
IVTT	−0.00441	0.00148	−2.97	< 0.01
OVTT	−0.00989	0.00251	−3.94	< 0.01
TC	−0.0291	0.00354	−8.22	< 0.01
Null log-likelihood: −2981.904				
Initial log-likelihood: −2981.904				
Final log-likelihood: −2053.936				
Likelihood ratio test: 1855.937				
Rho-square: 0.311				
Adjusted rho-square: 0.309				

technique is repeated for the entire dataset. The capturing of the model's predictive power is ultimately achieved.

The findings shown in Table 7 demonstrate that the computed prediction value is 74.9 %. This indicates that the model possesses the ability to accurately forecast around 74.9 % of the choices made by individuals planning trips. The predicted predictive power value exhibits a lower magnitude than anticipated, where the average relative error demonstrates a notable value of 25.1 %.

6. Development of machine learning models

The transport mode choice model is developed as a multi-class classification problem where the label (transport mode) is chosen from four alternatives, i.e., private car, taxi, public transport, and informal shared taxi. The calibration of the ML model is conducted using Python. The model Python Script can be found in the following GitHub repository: https://github.com/Moccino17/Transport_Mode_Sklearn.git.

Eight classifiers, used in the literature in transportation choice modeling, are chosen for fitting and tuning: Support Vector Machine (SVM); Multi-layer Perceptron (MLP); K-Nearest Neighbor (KNN); Adaptive Boosting (AdaBoost) Decision Tree (DT); Random Forest (RF), Gradient-Boosted Decision Trees (GBDT), and Extreme Gradient Boosting (XGB). First, the Pearson Correlation matrix, shown in Fig. 4, is used to select the top significant attributes, which are found to be car ownership, in-vehicle trip time, trip cost, and number of transfers.

A sensitivity analysis is done to inspect the effect of changing the dataset size on the 10-fold cross-validation (CV) accuracy of the classifier (Fig. 5 displays the results). The results indicate the following:

- Most classifiers show positive responses with increasing the sample size on the model prediction accuracy, except for the AdaBoost classifier, where the model prediction accuracy decreased with increasing the sample size. Thus, the AdaBoost classifier is excluded from further experiments due to its apparent unstable accuracy response.
- For SVM; MLP, KNN, DT; RF, GBDT, and XGB classifiers, the prediction accuracy increased by about 10 % to 24 % when the sample size was increased from 2 % to 12 %.
- While, the increase in the model prediction accuracy can be considered minor with increasing the sample size above 12 % of the dataset (e.g., about 2 % to 3 % when the sample size is increased from 12 % to 50 %.

The best hyperparameter combination is obtained by performing a brute-force grid-search procedure on the SVM; MLP, KNN, DT; RF, GBDT, and XGB classifiers. The 10-fold cross-validation accuracies of the classifiers are displayed in Fig. 6 and Table 8. The results indicate the following:

- The RF, GBDT, and XGB classifiers have the best overall performances in terms of the stability of model accuracy after iterating the models' calibration.
- After iterating the calibration 10 times, the model prediction accuracy for RF, GBDT, and XGB classifiers changed by only 6 % (i.e., the difference percentage between the minimum accuracy and maximum accuracy scored among the 10 iterations).
- The DT, KNN, and SVM classifiers relatively achieved moderate stability, where the difference percentage between the minimum accuracy and maximum accuracy is 9 %, 10 %, and 16 % respectively.
- In contrast, the MLP classifier has the lowest overall performance, where the model accuracy dropped by 21 % (i.e., from 86 % to 71 %).

The classifiers are also tested by splitting the dataset into a training set and a testing set of sizes of 85 % and 15 % of the total experiment dataset size, respectively. The resulting prediction capability is noted in Table 9.

The results reveal that the GBDT, XGB, RF, and DT classifiers have the best overall performances, where the model prediction accuracy for them equals 95.5 %, 93.6 %, 93.3 %, and 91.2 %, respectively, followed by KNN, SVM, and MLP classifiers whose model accuracy reaches 87.2 %, 84.5 %, and 83.7 % for the three classifiers, respectively.

7. Results and discussion

The comparison between the multinomial logit (MNL) and the (ML) models reveals that the ML models exhibit a higher level of prediction, surpassing the performance of the MNL model. The findings are displayed in Fig. 7, illustrating that the MNL model exhibits a prediction accuracy of around 74.9 %, while the ML models achieve prediction accuracies ranging from 83.7 % to 95.5 %.

Table 7
MNL model prediction capability.

Modes	Private Car	Taxi	Public Transport	Informal Shared Taxi	Mode Accuracy (%)	Total Accuracy (%)
Private car	30	3	1	1	85.7%	74.9
Taxi	9	15	0	1	60.0%	
Public Transport	5	0	54	21	67.5%	
Informal Shared Taxi	22	2	29	182	77.4%	

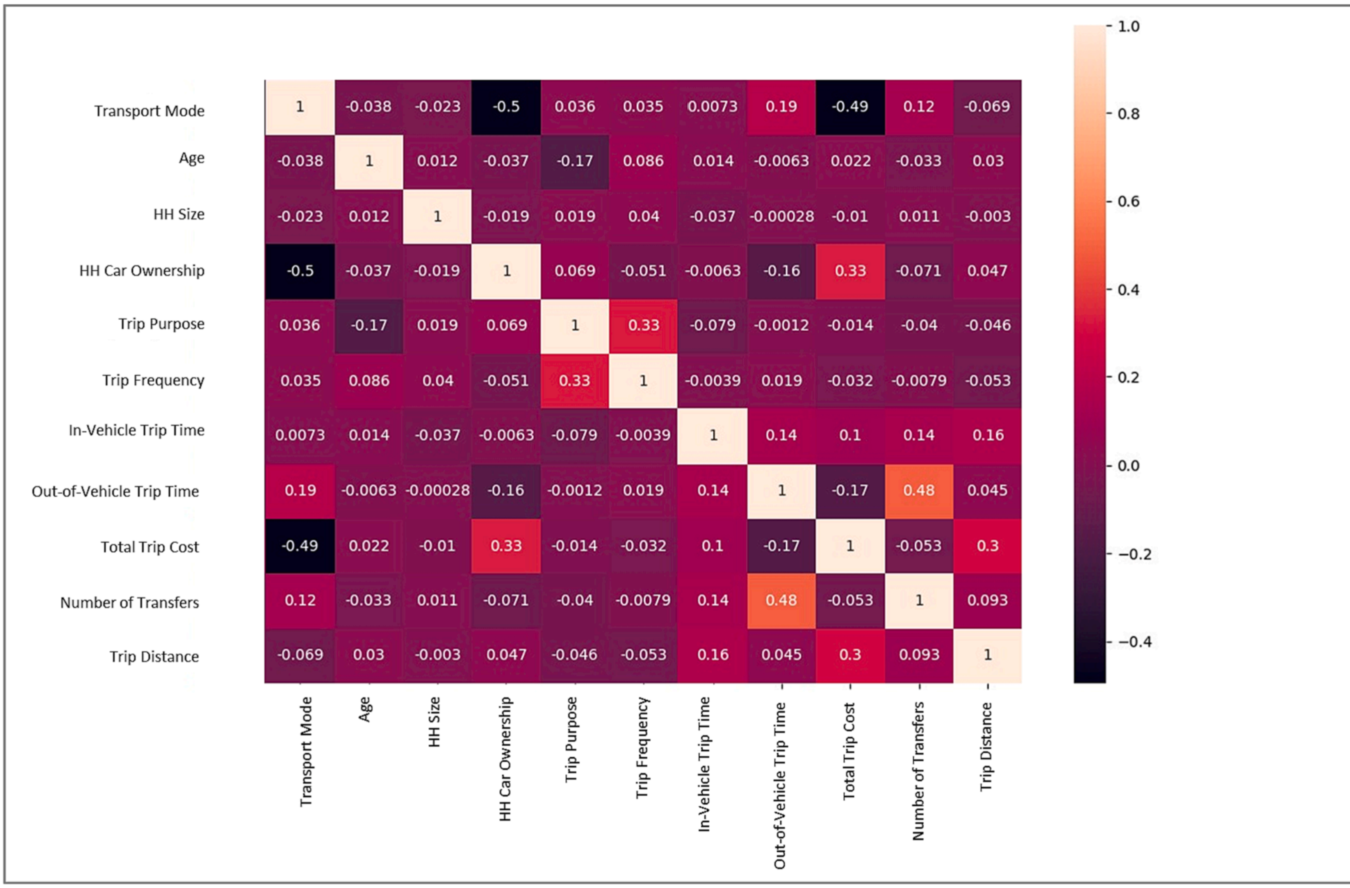

Fig. 4. Pearson correlation matrix.

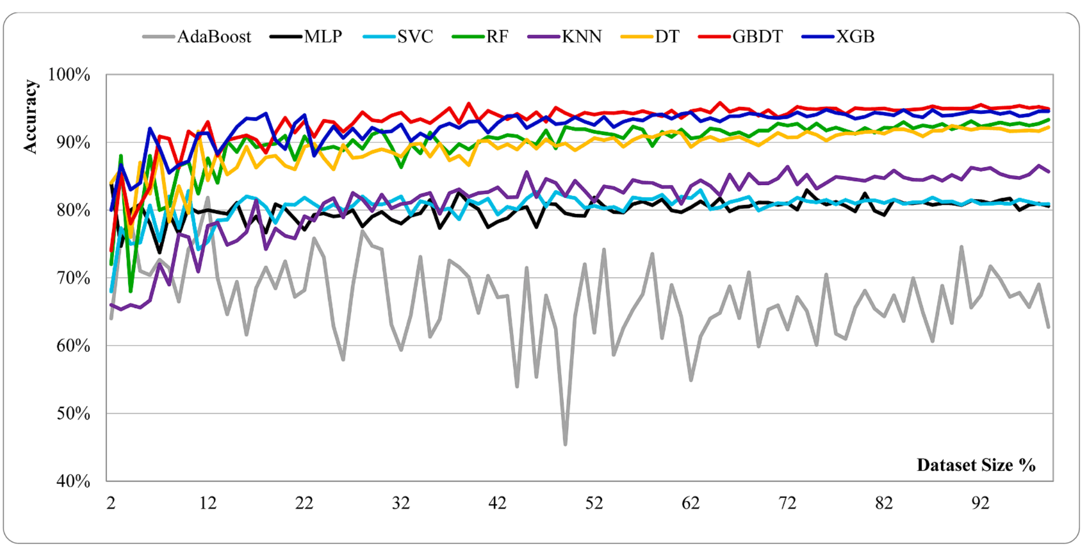

Fig. 5. Dataset size and accuracy for each considered classifier.

Among the ML models, the GBDT model has the highest level of accuracy. Therefore, the MNL model exhibits a prediction error of 25.1 %, but the ML-GBDT model has a significantly lower prediction error of just 4.5 %. The analysis of results reveals that using machine learning improves highly the prediction accuracy of the model, which proves that this intelligent approach can deeply understand and reflect the travelers' choice determinants and behavior over the traditional discrete choice models.

Practically, a developed model has to be accurate and robust, yet, sensitive to changes in travel attributes. Thus, this section aims to explore the sensitivity of the MNL model and the ML models (i.e., GBDT) to transportation policies.

Sensitivity is a measure employed to assess the extent to which the likelihood of selecting a certain mode of transportation changes in response to variations in the attribute value associated with that mode. The measurement of sensitivity analysis may be conducted using the

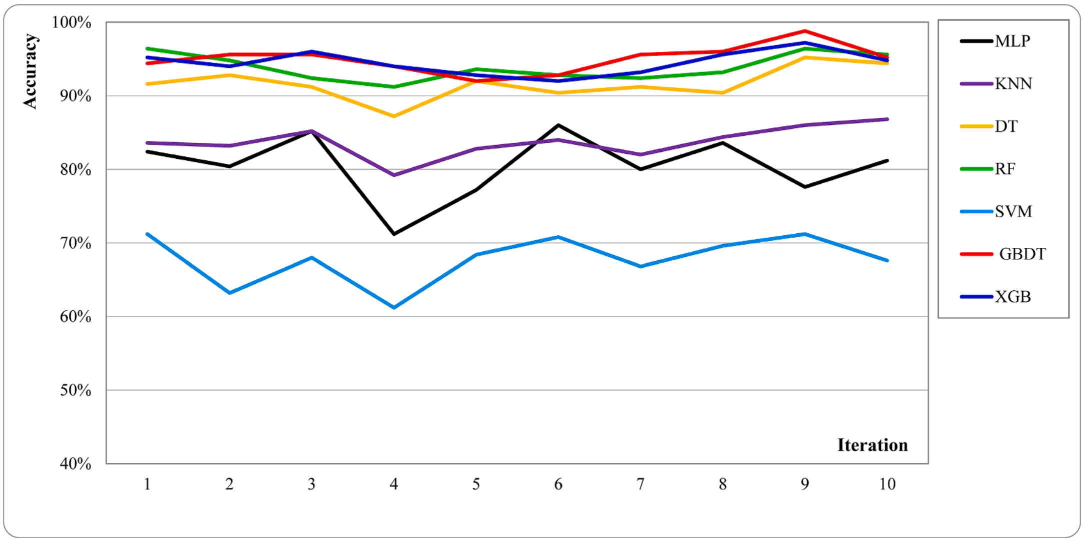

Fig. 6. 10-Fold cross-validation accuracy results.

Table 8
Main results of 10-fold cross-validation accuracy.

ML Models	MLP	KNN	DT	RF	SVM	GBDT	XGB
Minimum Accuracy %	71	79	87	91	61	92	92
Maximum Accuracy %	86	87	95	96	71	99	97
Difference %	21	10	9	6	16	6	6

Direct Elasticity test (Koppelman et al., 2006). The concept of Direct Elasticity pertains to the quantification of the percentage change in the probability of selecting a certain transportation mode in response to a modification in the value of the variable associated with it.

The input–output process is considered a major argument of ML models, where the statistical interpretation measures are not provided such as elasticity and sensitivity. However, the elasticity tests are performed by varying input variables, as in (Elharoun et al., 2023) and (Mohammadian and Miller, 2002).

To assess the sensitivity of the MNL model against the ML models, an application is performed relative to changes in the travel cost and travel time attributes of a transport mode. Consequently, the projected percentage change in the choice probability of the transport mode is calculated for different changes (i.e., increase and decrease) in the travel cost or the travel time within all models (i.e., MNL, GBDT, XGB, RF, DT, KNN, SVM, and MLP).

For instance, the application is applied to the travel cost of public transport mode. The sensitivity results for all models relative to changes (i.e., increase and decrease) are shown in Fig. 8. The results show that the elasticity in all models is negative in the case of increasing the travel cost attribute, where the choice probability of public transport decreases with an increase in its travel cost and vice versa.

However, the MNL model shows higher sensitivity than all ML models, especially the GBDT, GBDT, XGB, RF, DT, and KNN. For instance, the choice probability of choosing public transport mode after decreasing the travel cost by 75 % is increased by 31 % for the MNL model, while the same application in the ML models resulted in an increase ranging between 15.1 % and 23.1 %: 15.1 % in the ML-XGB and RF models, 15.6 % in the ML-GBDT model, 16 % in the ML-DT model, 17 % for the ML-KNN model, 22 % in the SVM model, and 23.1 % in the MLP model.

To validate the outcomes, another application is conducted on the informal shared taxi travel time, the models' elasticity relative to changes (i.e., increase and decrease) in its travel time attribute are shown in Fig. 9. Similar to the above application, the results show oversensitivity in the case of MNL model application compared to other ML models.

For illustration, the choice probability of choosing the informal shared taxi mode after increasing the travel time by 75 % is decreased by 25 % for the MNL model, while the same application in the ML models resulted a decrease ranging between 15.3 % and 19.5 %: 15.3 % in the ML-GBDT model, 16.4 % in RF model, 17.1 % in ML-XGB, 17.5 % in the MLP model, 18.7 % in the ML-DT model, 19.3 % for the ML-KNN model, and 22 % in the SVM model.

The comparison between the sensitivity of the two methods indicates that the MNL model can be considered oversensitive to changes in travel attributes. That is, the increase in the travel cost or time of a transport mode generates over modal shift to the other modes away from choosing it.

This can be explained as follows:

- For informal shared taxi and public transport modes, the ML models include the attribute of the number of transfers into the model, which plays an important role in choosing such collective transport modes. This attribute increases the accuracy of the model in terms of estimating the reasonable weight of other attributes in choosing these modes without overweighing them.
- For private cars, the MNL model does not include the car ownership attribute in the model contrary to the ML-GBDT model, disregarding the importance of this attribute in choosing the private car mode, which leads to an excessive increase of the impact of other attributes (e.g., travel cost) on the choice probability.
- For taxi mode, the MNL provides the worst prediction capability among other transportation modes (~60 %), while the prediction

Table 9
Prediction accuracy for each considered classifier.

Classifier	Modes	Private Car	Taxi	Public Transport	Informal Shared Taxi	Mode Accuracy (%)	Overall Accuracy (%)
MLP	Private car	22	2	3	8	62.9	83.7
	Taxi	9	13	0	3	52.0	
	Public Transport	0	0	58	22	72.5	
	Informal Shared Taxi	2	0	12	221	94.0	
KNN	Private car	18	4	2	11	51.4	87.2
	Taxi	6	16	0	3	64.0	
	Public Transport	0	0	70	10	87.5	
	Informal Shared Taxi	2	1	9	223	94.9	
RF	Private car	28	0	3	4	80.0	93.3
	Taxi	1	22	0	2	88.0	
	Public Transport	0	0	76	4	95.0	
	Informal Shared Taxi	2	0	9	224	95.3	
DT	Private car	26	1	0	8	74.3	91.2
	Taxi	1	23	0	1	92.0	
	Public Transport	0	0	70	10	87.5	
	Informal Shared Taxi	2	0	10	223	94.9	
SVM	Private car	23	0	3	9	65.7	84.5
	Taxi	8	15	0	2	60.0	
	Public Transport	0	0	57	23	71.3	
	Informal Shared Taxi	1	0	12	222	94.47	
GBDT	Private car	32	2	0	1	91.43	95.5
	Taxi	3	22	0	0	88	
	Public Transport	0	0	80	0	100	
	Informal Shared Taxi	2	0	9	224	95.32	
XGB	Private car	32	2	0	1	91.43	93.6
	Taxi	2	23	0	0	92	
	Public Transport	0	0	76	4	95	
	Informal Shared Taxi	4	0	11	220	93.62	

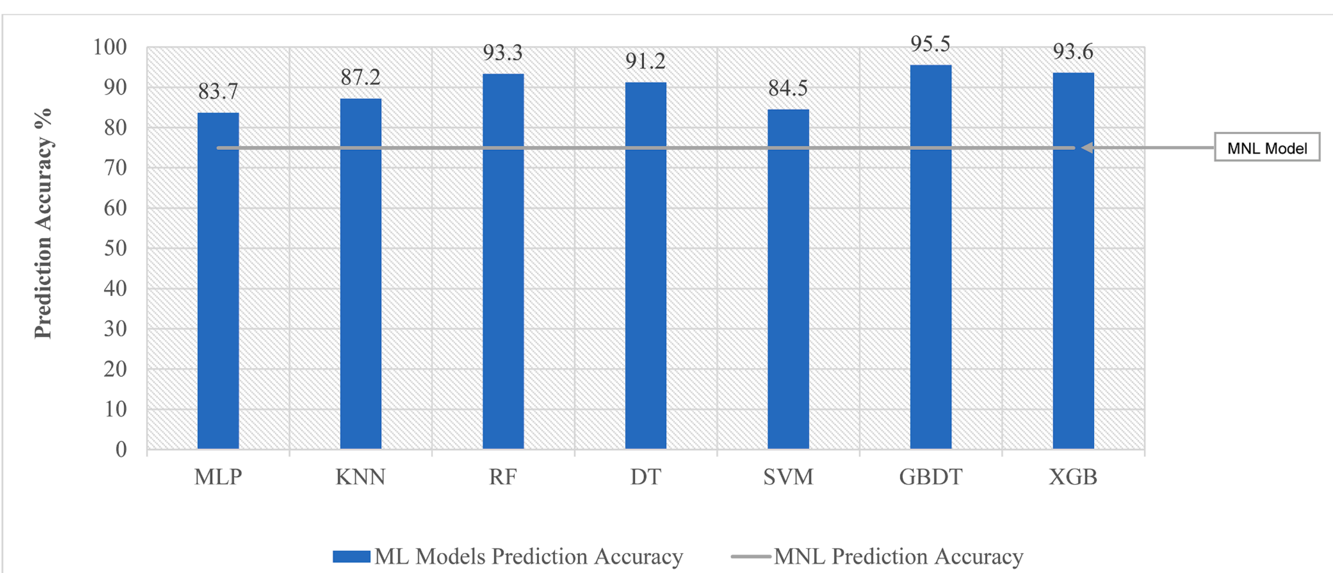

Fig. 7. Prediction accuracy comparison between MNL and ML.

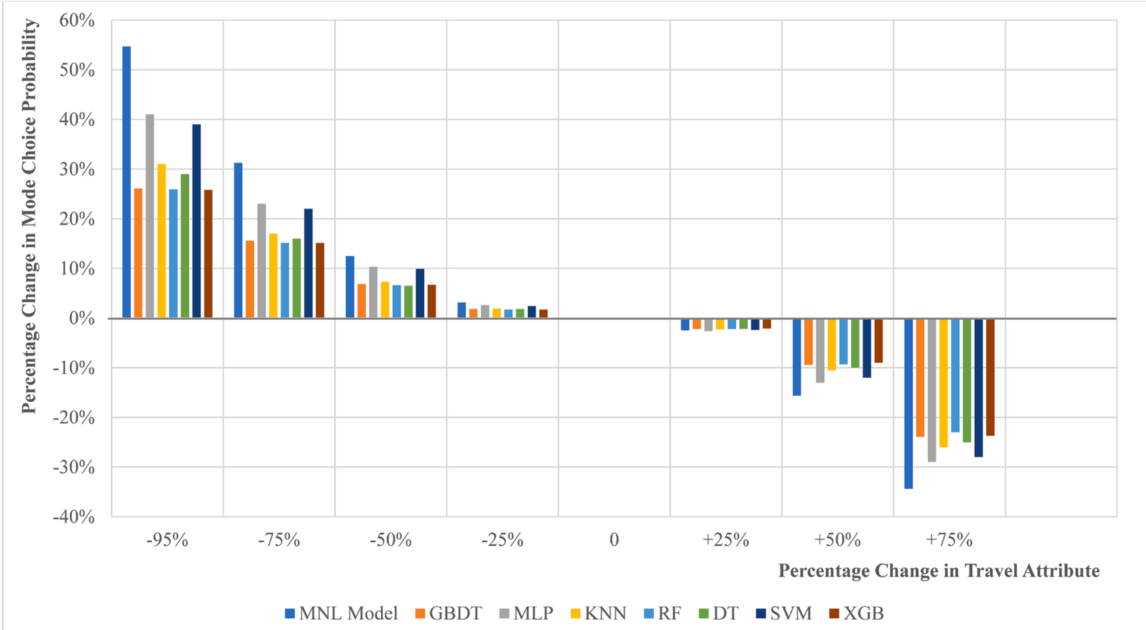

Fig. 8. Effect of Changing Public Transport Travel Cost on the Percentage Change of its Choice Probability.

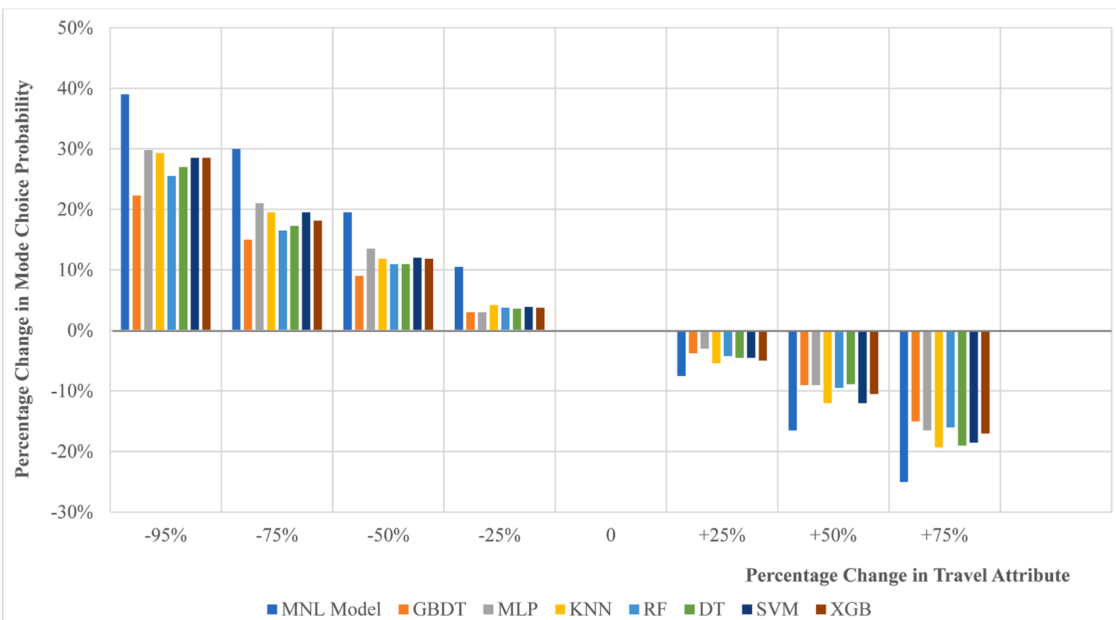

Fig. 9. Effect of Changing Informal Shared Taxi Travel Time on the Percentage Change of its Choice Probability.

capability of taxi mode in the ML-GBDT model is 92 %. This means that the error in taxi mode choice probability is about 40 %, which leads to an oversensitive condition to travel attributes.

Thus, it can be concluded that the ML models (especially the ML-GBDT model) can not only provide the best overall performance but also provide a better assessment of the model's sensitivity to travel attributes. ML-GBDT model reflects the travelers' choice determinants through a deep understanding of their travel behavior by considering important attributes that are disregarded in the MNL model, providing the most accurate weight of each transport attribute without overestimation. Thus, the ML-GBDT model is more indicative than the MNL model, and the sensitivity analysis results are more accurate.

8. Conclusions and future work

This article investigated the impact of using ML techniques in modeling the choice of transportation mode. A comparison was conducted with the Multinomial Logit (MNL) model to identify the impact on the model's accuracy. In addition, the performance of the models was evaluated to quantify to what extent the ML models are sensitive to transport policies compared to the traditional MNL model.

As a case study, the models were developed to reflect the choice behavior of travelers in Alexandria City, Egypt. To calibrate the models, a Revealed Preference (RP) survey was conducted for individuals from Alexandria of complete responses of 2,498 individual trips between February 2023 and April 2023.

According to the MNL model, successful calibration was achieved, with the parameters exhibiting the anticipated logical signs and demonstrating statistical significance at the 5 % level of significance. This is evidenced by the absolute value of the t-statistic for all parameters exceeding 1.96 and the P-value for all parameters being less than 0.05. Furthermore, the value of the likelihood ratio index (ρ2) is 0.311, and the adjusted likelihood ratio index ($\bar{\rho}$ 2) is 0.309. These results may be deemed as good. The predictive accuracy of the calibrated model was determined to be 74.9 % based on the estimated parameters.

For the ML model, eight classifiers were chosen for fitting and tuning: Support Vector Machine (SVM); Multi-layer Perceptron (MLP); K-Nearest Neighbor (KNN); Adaptive Boosting (AdaBoost), Decision Tree (DT); Random Forest (RF), Gradient-Boosted Decision Trees (GBDT), and Extreme Gradient Boosting (XGB). A sensitivity analysis was conducted to inspect the effect of changing the dataset size on the 10-fold cross-validation (CV) accuracy of the classifier, which indicated that all classifiers' accuracy responses were stable, except for the AdaBoost classifier, which was excluded from further experiments.

The 10-fold cross-validation accuracies of the classifiers were estimated, which show a good performance exceeding 90 % for the GBDT, XGB, and RF classifiers. Further examination was conducted by splitting the dataset into a training set and a testing set of sizes of 85 % and 15 % of the total experiment dataset size, respectively. The results indicate that the GBDT, XGB, RF, and DT classifiers have the best overall performances with model prediction accuracy of 95.5 %, 93.6 %, 93.3 %, and 91.2 %, respectively, followed by KNN, SVM, and MLP classifiers whose model accuracy reached 87.2 %, 84.5 %, and 83.7 %, respectively. Although the MNL model is easier than the ML model in terms of calibration, the comparison between the two models shows that the ML model notably has better prediction performance.

Finally, a sensitivity analysis test was conducted to quantify the elasticity of the MNL and ML models to changes in travel cost or travel time attributes. To this end, an application is performed relative to changes in the travel cost of public transport mode and travel time of informal shared taxi mode. Consequently, the projected percentage change in the choice probability of the transport modes is calculated for different changes (i.e., increase and decrease) in the travel attributes.

The results show that the elasticity is negative in the case of increasing the travel cost or time attributes in all models, where the choice probability of a transportation mode decreases with an increase in its travel cost or travel time. However, the MNL model shows higher sensitivity than the ML models, especially the GBDT, GBDT, XGB, RF, DT, and KNN.

The comparison between the sensitivity of the two methods shows that the MNL model overestimates the model sensitivity, while most ML models are more indicative and the sensitivity analysis results are more reliable.

The developed ML models comprehensively reflect the choice determinants, where the individuals' economic status is reflected through including car ownership attribute into the model beside the attributes included also in the MNL model. On the other hand, a main factor affecting transportation mode choice (i.e., number of transfers) is included. Accordingly, the ML interpretation of the choice determinants is more reliable.

Thus, it can be concluded that using ML highly improves not only the prediction accuracy but also the performance of the model indicating that this intelligent approach can reflect the travelers' choice determinants through an accurate understanding of their travel behavior over the traditional discrete choice models.

For future work, it is recommended to calibrate separate models for the groups of travelers rather than a single combined model to increase the MNL model accuracy. However, this may require a sufficient number of records. Thus, it is suggested to expand the size of the dataset in future studies to solve this problem. On the other hand, other discrete choice models such as Nested Logit, Cross Nested Logit, and Mixed Multinomial Logit can be calibrated and compared with ML models. Furthermore, other types of intelligent methods (e.g., ANN, and Fuzzy logic) can be calibrated and compared with the developed models.

Funding

The research conducted did not receive any financial support from external sources.

CRediT authorship contribution statement

Ahmed Mahmoud Darwish: Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Mohamed Almansour:** Writing – original draft, Software, Formal analysis, Data curation. **Ayman Salah:** Validation, Investigation, Conceptualization. **Maged Zagow:** Writing – original draft, Investigation, Data curation. **Khaled Saeed:** Writing – original draft, Methodology. **Ahmed Elkafoury:** Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Abdel-Aal, M., 2017. Value of time determination for the city of Alexandria based on a disaggregate binary mode choice model. Alex. Eng. J.

Abduljabbar, R., et al., 2019. Applications of artificial intelligence in transport: An overview. Sustainability 11 (1), 189. https://doi.org/10.3390/SU11010189.

Abulibdeh, A., 2023. Analysis of mode choice affects from the introduction of Doha metro using machine learning and statistical analysis. Transp. Res. Interdiscip. Perspect. 20, 100852 https://doi.org/10.1016/J.TRIP.2023.100852.

Arabameri, A., et al., 2020. Comparison of machine learning models for gully erosion susceptibility mapping. Geosci. Front. 11 (5), 1609–1620. https://doi.org/10.1016/J.GSF.2019.11.009.

Assi, K.J., et al., 2019. Travel-to-school mode choice modelling employing artificial intelligence techniques: A comparative study. Sustainability 11 (16), 4484. https://doi.org/10.3390/SU11164484.

Ben-Hur, A., et al., 2001. Support vector clustering. J. Mach. Learn. Res. 2, 125–137.

Bhavsar, P. *et al.* (2017) 'Machine Learning in Transportation Data Analytics', *Data Analytics for Intelligent Transportation Systems*, pp. 283–307. Available at: https://doi.org/10.1016/B978-0-12-809715-1.00012-2.

Bierlaire, M. (2016) *PythonBiogeme: a short introduction.*

Bishop, C.M. (2006) *Pattern Recognition and Machine Learning.*

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324/METRICS.

Capmas, 2021. Yearly Statistics Book. Central Agency for Public Mobilization and Statistics, Egypt.

Cheng, L., et al., 2019. Applying a random forest method approach to model travel mode choice behavior. Travel Behav. Soc. 14, 1–10. https://doi.org/10.1016/J.TBS.2018.09.002.

Chu, Y.L., 2018. Implementation of a new network equilibrium model of travel choices. J. Traffic Transp. Eng. (English Ed.) 5 (2), 105–115. https://doi.org/10.1016/J.JTTE.2017.05.014.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297. https://doi.org/10.1023/A:1022627411411/METRICS. Available at:

Darwish, A.M., 2021. Development of transport mode choice model between public and private transport in densely metropolitan areas, Alexandria as a case study. Alexandria University, Egypt.

El Esawey, M., Ghareib, A. (2009) 'Analysis of Mode Choice Behavior in Greater Cairo Region', in *Transportation Research Board 88th Annual Meeting.*

El-Bany, M.E.S., et al., 2014. Policy sensitive mode choice analysis of Port-Said City, Egypt. Alex. Eng. J. 53 (4), 891–901. https://doi.org/10.1016/J.AEJ.2014.09.006.

Elharoun, M., El-Badawy, S.M., Shahdah, U.E. (2023) 'Artificial Intelligence Techniques for Predicting Individuals' Mode Choice Behavior in Mansoura City, Egypt', *Transportation Research Record: Journal of the Transportation Research Board*, p. 036119812311625. Available at: https://doi.org/10.1177/03611981231162598.

Elharoun, M., Shahdah, U.E., El-Badawy, S.M., 2018. Developing a mode choice model for Mansoura city in Egypt. Int. J. Traffic Transp. Eng. (IJTTE). https://doi.org/10.7708/ijtte.2018.8(4).10.

Freund, Y., Schapire, R. (1996) 'Experiments with a New Boosting Algorithm'.

Hagenauer, J., Helbich, M., 2017. A comparative study of machine learning classifiers for modeling travel mode choice. Expert Syst. Appl. 78, 273–282. https://doi.org/10.1016/J.ESWA.2017.01.057.

Harz, J., Sommer, C., 2022. 'Mode choice of city tourists: Discrete choice modeling based on survey data from a major German city. Transp. Res. Interdiscip. Perspect. 16, 100704 https://doi.org/10.1016/J.TRIP.2022.100704. Available at:

Heiss, F., 2016. Discrete choice methods with simulation. Econ. Rev. 35 (4), 688–692. https://doi.org/10.1080/07474938.2014.975634.

Hussain, D., et al., 2017. Analysis of transportation mode choice using a comparison of artificial neural network and multinomial logit models. ARPN J. Eng. Appl. Sci. 12 (5).

Khan, A., et al., 2010. A review of machine learning algorithms for text-documents classification. J. Adv. Inf.

Khan, S., Adnan, A., Iqbal, N. (2022) 'Applications of Artificial Intelligence in Transportation', *International Conference on Electrical, Computer, and Energy Technologies, ICECET 2022* [Preprint]. Available at: https://doi.org/10.1109/ICECET55527.2022.9872928.

Koppelman, F.S. *et al.* (2006) *A self instructing course in mode choice modeling: multinomial and nested logit models*.

Lasantha, D., Vidanagamachchi, S., Nallaperuma, S., 2023. Deep learning and ensemble deep learning for circRNA-RBP interaction prediction in the last decade: A review. Eng. Appl. Artif. Intel. 123, 106352 https://doi.org/10.1016/J.ENGAPPAI.2023.106352.

Levy, P., Lemeshow, S. (2013) *Sampling of populations: methods and applications*.

Lu, X.-S., Liu, T.-L., Huang, H.-J., 2015. Pricing and mode choice based on nested logit model with trip-chain costs'. Transp. Policy.

Migliore, M., Ciccarelli, M.V., 2020. The role of the integrated public transport network for reducing the congestion in the city of palermo. Adv. Sci. Technol. Innov. 297–303. https://doi.org/10.1007/978-3-030-17308-1_28.

Mohammadian, A., Miller, E.J., 2002. Nested logit models and artificial neural networks for predicting household automobile choices: Comparison of performance'. Transp. Res. Rec. (1807), 92–100. https://doi.org/10.3141/1807-12.

Newman, J.P., Lurkin, V., Garrow, L.A., 2018. Computational methods for estimating multinomial, nested, and cross-nested logit models that account for semi-aggregate data. J. Choice Model. 26, 28–40. https://doi.org/10.1016/J.JOCM.2017.11.001.

O'Flaherty, C. (2018) *Transport Planning and Traffic Engineering*.

Omrani, H., 2015. Predicting travel mode of individuals by machine learning. Transp. Res. Procedia 10, 840–849. https://doi.org/10.1016/J.TRPRO.2015.09.037.

Ortúzar, J. de D., Willumsen, L.G. (2011) *Modelling transport*. 4th ed., *Modelling Transport*. 4th ed. Chichester: Wiley-Blackwell. Available at: https://doi.org/10.1002/9781119993308.

Rahman, M., et al., 2022. Application of machine learning classifiers for mode choice modeling for movement-challenged persons. Future Transportation 2 (2), 328–346. https://doi.org/10.3390/FUTURETRANSP2020018.

Richards, M.J., Zill, J.C., 2019. Modelling mode choice with machine learning algorithms. In *Australasian Transport Research Forum 2019 Proceedings*.

Sarker, I.H., 2022. AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. SN Comput. Sci. 3 (2), 1–20. https://doi.org/10.1007/S42979-022-01043-X.

Sekhar, C.R., Minal, Madhu, E., 2016. Mode Choice analysis using random Forrest decision trees. Transp. Res. Procedia 17, 644–652. https://doi.org/10.1016/J.TRPRO.2016.11.119.

Shmueli, D., Salomon, I., Shefer, D., 1996. Neural network analysis of travel behavior: Evaluating tools for prediction. Transp. Res. Part C: Emerg. Technol. 4 (3), 151–166. https://doi.org/10.1016/S0968-090X(96)00007-1.

Sun, Y., Fung, B.C.M., Haghighat, F., 2022. In-Processing, fairness improvement methods for regression Data-Driven building models: Achieving uniform energy prediction. Energy Build. 277, 112565 https://doi.org/10.1016/J.ENBUILD.2022.112565.

Taherdoost, H., 2017. Determining sample size; how to calculate survey sample size. Int. J. Econ. Manage. Syst. [preprint].

Tan, L., Ma, C., 2021. Choice behavior of commuters' rail transit mode during the COVID-19 pandemic based on logistic model. J. Traffic Transp. Eng. (English Ed.) 8 (2), 186–195. https://doi.org/10.1016/J.JTTE.2020.07.002.

United Nations (2018) Population Division, World Urbanization Prospects, the 2018 Revision, Department of Economic and Social Affairs.

Wang, F., Ross, C.L., 2018. Machine learning travel mode choices: Comparing the performance of an extreme gradient boosting model with a multinomial logit model'. Transp. Res. Rec.: J. Transp. Res. Board 2672 (47), 35–45. https://doi.org/10.1177/0361198118773556.

Willumsen, L., Ortúzar, J. de D. (2016) *Transport Planning*, *elgaronline.com*. Available at: https://doi.org/https://doi.org/10.4337/9781783471393.00028.

Xie, Y., Lord, D., Zhang, Y., 2007. Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. Accid. Anal. Prevent. 39 (5), 922–933. https://doi.org/10.1016/J.AAP.2006.12.014.

Xie, C., Lu, J., Parkany, E., 2003. Work Travel mode choice modeling with data mining: decision trees and neural networks'. Transp. Res. Rec. (1854), 50–61. https://doi.org/10.3141/1854-06.

Zhan, G., et al., 2016. Using hierarchical tree-based regression model to examine university student travel frequency and mode choice patterns in China'. Transport Policy 45, 55–65. https://doi.org/10.1016/J.TRANPOL.2015.09.006.

Zhang, Y., Xie, Y., 2008. Travel mode choice modeling with support vector machines'. Transp. Res. Rec.: J. Transp. Res. Board (2076), 141–150. https://doi.org/10.3141/2076-16.