Transportation Research Interdisciplinary Perspectives 20 (2023) 100852

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

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

Analysis of mode choice affects from the introduction of Doha Metro using machine learning and statistical analysis

Ammar Abulibdeh

Applied Geography and GIS Program, Department of Humanities, College of Arts and Sciences, Qatar University, Doha, Qatar

ARTICLE INFO

Keywords:
Machine learning
FIFA World Cup 2022
Travel behavior
Trip characteristic
Stated preference survey
Statistical analysis

ABSTRACT

The aim of this study was to investigate the possible influences of the operation of the new Doha Metro on the travel mode choice behavior in Doha City, Qatar. Revealed preference (RP) and stated preference (SP) survey questionnaires were designed to collect the necessary data. The questions considered different trip conditions and socioeconomic factors of travelers. Three different mode choices were considered in this study: private cars, taxi services, and metro. Two statistical models and one machine learning model were used to analyze the current and future mode choices: discrete choice binary logit (BL) and multinomial logit (MNL) models as well as extreme gradient boosting (XGBoost). Furthermore, the SHapley Additive exPlanations (SHAP) method was used to rank the input features based on their importance according to the mean SHAP value. The results showed that the XGBoost model outperforms the other two models in terms of predicting the travel mode choice as well as in terms of its accuracy. The results showed that various trip characteristics are significant in determining the mode choice, including the number of travelers and bags, journey time, and reimbursement of parking fees. Furthermore, different socioeconomic characteristics proved to be significant for the current and future mode choices, including nationality, income, age, employment status, and vehicle ownership.

1. Introduction

Large urban centers often face significant transport problems, such as traffic and road congestion, increased car ownership and automobile dependence, emission-based pollution, and decreased quality of life. Many cities across the globe have invested heavily in the introduction, extension, or redevelopment of old and new urban public transit systems with the aim of managing/mitigating automobile dependence and its associated problems. Many scholars view a rail-based system as a major public transport option that can comprise a crucial element in incapacitating the substantial problems raised by the use of private automobiles (Chen et al., 2016; Halse et al., 2019; Ibrahim et al., 2020; Zhu et al., 2020). Advocates of public transportation systems favor such systems based on perceived benefits that mainly include increased market share for public transportation, reduced environmental impact, reduced automobile dependence, and positive effects on urban development (Abulibdeh, 2017; Golias, 2002; Hawas et al., 2016).

Globally, most efforts to develop new transportation systems have been focused on metros (or subways). Implementing a rail-based transit system, particularly metro rail, is seen by many cities in developing countries as a promising approach to achieving sustainable transportation, and as a solution to the problems of urban traffic congestion and rapidly increasing travel demand. Metros offer unparalleled quality of service (in terms of speed, frequency of service, and travel time reduction) and can act as a better mitigation approach to automobile travel than purely bus-based systems (Golias, 2002; Zaidan & Abulibdeh, 2018). In many metropolitan areas, public transportation systems have attracted more attention in recent years at the expense of single-occupancy vehicles because of the restricted capacity of transportation infrastructure and the need to achieve sustainability (Zaidan and Abulibdeh, 2021; Moeckel et al., 2014). These metropolitan areas have aimed to develop and enhance the public transport network so as to increase ridership and at the same time, discourage individuals from using private vehicles for commuting. However, despite many efforts in this area, the share of single-occupancy vehicles is still much larger than that of public transport (Abulibdeh et al., 2015a; Abulibdeh et al., 2015b; Li et al. 2017; Ouda et al. 2013).

Travel demand forecasting for new transport systems, particularly for new metro and subway systems before their implementation, is a key undertaking in investigating travelers' attitudes to – and likely behaviors toward – new systems (Heinen et al., 2017; Liu et al., 2017; Sohoni et al., 2017). Several studies have examined the impacts of new public

E-mail address: aabulibdeh@qu.edu.qa.

https://doi.org/10.1016/j.trip.2023.100852
Received 20 March 2023; Received in revised form 6 May 2023; Accepted 25 May 2023
Available online 31 May 2023
2590-1982/

transport systems on mode choice in many cities around the world (e.g., (Birolini et al., 2019; Creemers et al., 2012; Golias, 2002; Hensher & Rose, 2007; Jou et al., 2011)). These studies aimed to determine the factors that encourage air travelers to use the metro to commute to airports. Some studies have found that these benefits are much smaller than those expected by actors fostering the systems. Conversely, other studies found that some attributes such as travel time saving, travel cost, trip purpose, number of luggage, and user-friendly nature of the modal, are significant in ground mode choice for air passengers (Jou et al., 2011). Some studies found that airport attractiveness is highly elevated by an extension of the rapid-transit link and increases the dominance of specific airports compared with others in the same area (Bergantino et al., 2020). Furthermore, Bergantino et al. (2020) found that introducing a new mode, such as a new express bus service, reduces the access time by 15 min (Tsamboulas & Nikoleris, 2008).

The present analysis builds on previous studies that provide a background on different facets and explanations relevant to the subject area and act as a starting point for the research. A review of the literature provided the following information:

First, previous studies on airport travel mode choice have categorized the explanatory variables into two groups to investigate the airport travel mode choice (e.g., (Birolini et al., 2019; Pasha et al., 2020; Reza Mamdoohi et al., 2012; Yaylali et al., 2016; Yazdanpanah & Hosseinlou, 2016)). One group consists of explanatory variables related to trip features (i.e., travel time and cost, number of travelers and their associated luggage, travel time to the airport from trip origin, airplane departure times, out-of-vehicle and in-vehicle travel times, and trip purpose). The other group consists of the socioeconomic features of the commuters, which include gender, profession, employment status, household income, age, and nationality.

Second, these studies are based on RP surveys conducted to amass data related to the explanatory variables to investigate the research question. In this type of survey, respondents indicated their mode of choice to travel to the target destination from their origin and provide information about the trip characteristics (such as the number of luggage items, trip cost, and time). They were also asked about their socioeconomic characteristics (such as household income, gender, age, employment status, etc.) (Abulibdeh, 2020; Cirillo et al., 2014; Hasnine et al., 2019; Ouda et al., 2013; Petrik et al., 2016).

Third, the collected data were analyzed using different discrete choice models (binary, mixed, and multinomial logit models). These models are characterized by their simple probabilistic choice function, clarity in algebraic manipulations related to the derivation of the final probabilistic choice function, and ease of interpretation of the estimation results (Gokasar & Gunay, 2017; Gunay & Gokasar, 2021; Jiang et al., 2021; Abulibdeh et al., 2018; Pasha et al., 2020).

Fourth, the literature presents compatible and key insights with respect to driving forces that influence the travelers' choice of specific travel mode to get to airports, particularly the time spent traveling and the cost of the trip (Birolini et al., 2019; Jiang et al., 2021).

Fifth, the latest computational progress has facilitated simpler implementation of machine learning models for the examination of travel behavior. However, studies conducted in this area are yet to be comprehensive or decisive. Statistical analysis and machine learning models were rarely employed together in the preceding research to scrutinize mode selection and contrast the outcomes of these models, along with their efficiency in prognosticating mode choice. The predominant advantage of many machine learning models lies in their lack of stringent statistical assumptions, which allows for their adaptable utilization across various data structures (Wang & Ross, 2018).

Completion of the Doha Metro influences the tendency of the travelers to commute within the city and to the airport. Therefore, it was deemed important to investigate the influence of the metro on mode choice — taking the airport as a case study — to determine the effectiveness of the metro in attracting new users. Therefore, this study aimed to investigate the existing mode choice behavior and future changes in this behavior after the operation of the new metro in Doha city. Furthermore, this study aimed to understand the modal preferences, trips, and demographic characteristics of airport users. This was done by focusing on airport trips in Doha city to Hamad International Airport (HIA) to gain a better understanding of the factors shaping the airport air travelers' current mode choice and how the introduction of the new mode will change the current choice. Specifically, the objective of this study was to enhance our perception of the travelers' requirements, anticipate the potential impact of the new metro, and to identify the driving forces that encourage travelers to use it as their primary choice to commute to the airport.

Previous studies on travel mode choice presented consistent and important insights into the variables that affect the travelers' mode choice of the existing access mode. Most of these studies were conducted in developed or developing countries where public transportation is an important element of the transportation system. Unlike these studies, the focus of this study was the use of public transportation in a developing country that is currently working on introducing the metro as a new mode of public transportation at the time this survey was conducted. Furthermore, this study aims to compare between two statistical models and one machine-learning model in terms of their performance in predicting future mode choice.

2. Study area

Qatar is located in the far eastern part of the Arabian Peninsula and has a geographical area of 11,437 square kilometers (Balakrishnan et al., 2023). Qatar is currently experiencing high rates of population growth as a result of massive urban development, rising government expenditures, and largescale investment projects (Al-Awadhi et al., 2022; Zaidan & Abulibdeh, 2020). Doha city is the capital of Qatar and its most populous city. The city has experienced significant economic growth and urban expansion over the past three decades (Abulibdeh and Zaidan, 2017). It is an urban primacy city in the Gulf region and is considered to be the hub of the state of Qatar in terms of economic development, population, and culture. It also serves as the basis for regional development. The location of the city has contributed to its becoming a junction of transport routes (Abulibdeh, 2022; Ghofrani et al., 2022).

The discovery of significant amounts of oil and gas, and population growth in cities, began to accelerate in the 1970 s, attracting a large number of expatriates to work in different fields. As a result, the urban area of the city expanded by approximately 640% at an annual average rate of 19.6% between 1990 and 2000, compared with 8.9% between 2000 and 2017 (Abulibdeh et al., 2019a; Abulibdeh, 2019b) as shown in Fig. 1.

Qatar is classified as an arid or semi-arid country with hot summers and mild winters (Abulibdeh, 2021a, 2021b; Timothy, 2018). Commuting using bicycles or by walking is difficult and unadvisable, particularly during the summer season. Since its establishment in 1970, Qatar has planned and built a transportation system around the movement of automobiles. In parallel with urban expansion, Doha city has experienced a rapid increase in the number of private vehicles in use. Owning a car is a trademark of this country, where a high percentage of households — particularly Qataris — own state-of-the-art cars. Qatari citizens consider owning more than one vehicle to be normal and a source of pride. This has influenced the use of public transport networks currently and will continue to do so in the future (Mohammed et al., 2023). Qatari and non-Qatari travelers in Qatar show a strong desire for private automobiles because of the lack of diverse transport options and inferior public transport services available in the city. Moreover, automobile ownership is often considered a status symbol that offers comfortable and safe travel opportunities for travelers. Automobile ownership is the primary mode of choice in travel-related decision-making processes in Qatar.

Public transportation was not popular in the country until it succeeded in getting selected to host the 2022 FIFA World Cup. The country

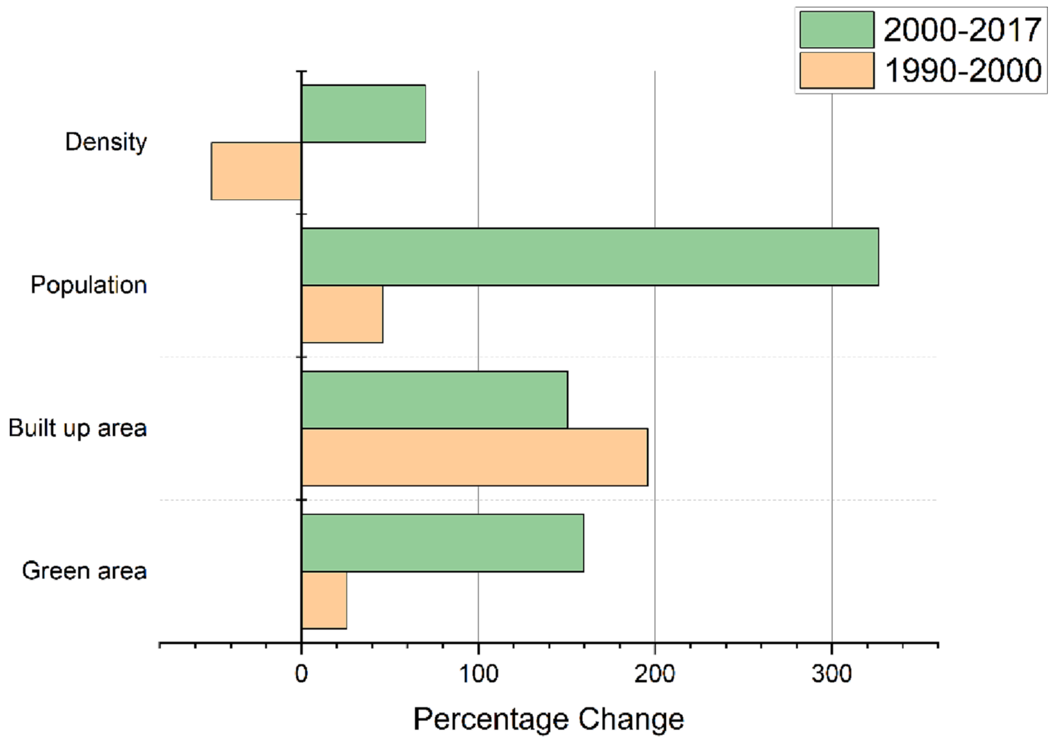


Fig. 1. Land-cover change and population growth in Doha (Abulibdeh et al., 2019).

spent and invested billions of dollars in developing its transportation infrastructure, including the public transportation network, to manage the growth in air travel demand during this mega event, which will put large pressure on the only airport in the country to improve the capacity and level of service of access and egress of the ground travel mode choice. The bus network was expanded with minimum fares ranging between QR 3 and 4 ($0.8–1) inside Doha and QR 4 – 9 ($1–2.5) outside the city (Mansour et al., 2022). However, this system is not popular in the country and is used mainly by low-income expats, particularly those from India, the Philippines, Bangladesh, and other Asian countries. The taxi system and Doha limousine service are options for travelers to commute to airports and other places within the country; however, the number of operated taxis and limousines is still lower than the actual needs of the residents. Therefore, proposing a new public transportation mode in a vehicle-oriented society requires thorough investigation of the transport user preferences for private vehicle ownership and future user preferences for a mode shift towards a new alternative.

Doha Metro is a state-of-the-art mass rapid transit system that is being built in phases. Phase 1 includes building 37 stations along three main lines: red, green, and gold lines, and over an operational length of approximately 76 km. The red line extends from south to north, while the green and gold lines extend from east to west in Doha city, as shown in Fig. 2.

3. Material and Methods

3.1. Questionnaire development and data collection

3.1.1. Survey design

An RP and SP survey was designed to gather information on travelers' travel behavior in choosing a mode choice to commute to HIA. SP and RP self-administered questionnaire surveys were used to collect primary data. This method is suitable for investigating ground access mode choices and enables us to reach a large number of people in a short period. The questions in the RP questionnaire survey focused on the current mode choice selection of air passengers, specifically their trip origin, trip purpose, travel time to the airport, mode selection, and parking reimbursement (Abulibdeh and Zaidan, 2018; Earnhart, 2002; Tseng et al., 2013). Conversely, the questions in the SP survey focused on hypothetical choice scenarios, such as using the Doha Metro in the future when it starts operating. These questions reflected what the commuters said rather than what they actually did. Therefore, the SP data were used in this investigation to predict the future mode choice after Doha Metro will start operating and to evaluate the critical factors (e.g., trip and socioeconomic attributes) that will affect the individual decisions regarding the mode choice.

One of the advantages of using the SP survey is that it allows us to investigate the possibility of using the metro to commute to the airport in a more comprehensive manner (Cherchye et al., 2015; Earnhart, 2002). Furthermore, the SP survey is more flexible than the RP survey (Cherchye et al., 2015; Earnhart, 2002; Tseng et al., 2013). However, one of the main criticisms of this technique is the lack of reliability, as the expressed preference of the travelers may not coincide with the actual behavior because the answers are related to a hypothetical situation. However, SP and RP self-administered questionnaire surveys have been used in many studies to collect primary data on ground access mode choices (Abulibdeh, 2018; Birolini et al., 2019; Ramsey et al., 2017). RP data are significant when analyzing the travel behavior of existing transport alternatives using a discrete choice model. Therefore, RP data were used to analyze the current mode choice (cars and taxi services) of travelers to HIA. Models estimated using RP data have the advantage of reflecting choices in real-world market settings. The results of the SP survey are significant for travel-demand forecasting for new alternatives, as RP data cannot exist before new modes are implemented.

In this study, the questions were divided into three categories. The first category was designed to gather information on travelers' trip characteristics. This part of the survey consisted of ten questions. Travelers were asked to state the number of people traveling, number of luggage trips, trip purpose, time of departure, arrival time to the airport, and class (economy, business, or first). The second set of questions was designed to gather information related to mode choice selection. This section consisted of 12 questions. Among these questions, respondents

Fig. 2. Qatar Metro lines and stations.

were asked to state their main mode of transport that they used to commute to the airport, whether they were drivers or passengers, if they always used the same mode to travel to the airport, how often they used public transportation, parking reimbursement, factors influencing their choice of mode of transport, and their intention to use Doha Metro once it begins to operate. The remaining questions were aimed at gathering information related to the socioeconomic characteristics of travelers, such as income, age, gender, household size, nationality, and education.

3.1.2. Data collection

To ensure a greater likelihood of a comprehensive response, the survey was randomly distributed only among air travelers who departed from the airport and waited in the boarding area, while connecting-flight air passengers were not included. The survey was conducted from January to March 2018 in both Arabic and English, as residents of Qatar are from different nationalities. The participants were selected based on a systematic approach: the first traveler sitting in the first row of the boarding waiting area was selected and given the survey; then, the third traveler was selected, and so forth. If any traveler refused to participate in the questionnaire, the next traveler was asked to participate. A total of 1546 air travelers were interviewed face-to-face to complete the survey. However, only 1247 of the surveyed households provided the complete information and were considered in the analysis.

3.2. Methodology

The impact of Doha Metro as a new mode of choice was assessed by analyzing the socioeconomic and trip characteristics of the travelers in Doha using binary logistic regression, multinomial regression analysis, and extreme gradient boosting (XGBoost). Furthermore, the SHapley Additive exPlanations (SHAP) method was used to rank the input features based on their importance according to the mean SHAP value. The explanatory variables were classified into two categories: i) trip characteristics, such as the number of luggage items and travelers, trip purpose, trip cost, and journey duration to the airport; and ii) socio-economic attributes of commuters, such as household income, nationality, age, gender, education, and employment status. A detailed flowchart is shown in Fig. 3.

In this study, airport travel mode choice was modeled according to the utility maximization principle combined with psychological choice behavior and the economic theory of consumer behavior. First, a chi-square test was used to examine the significant factors affecting the

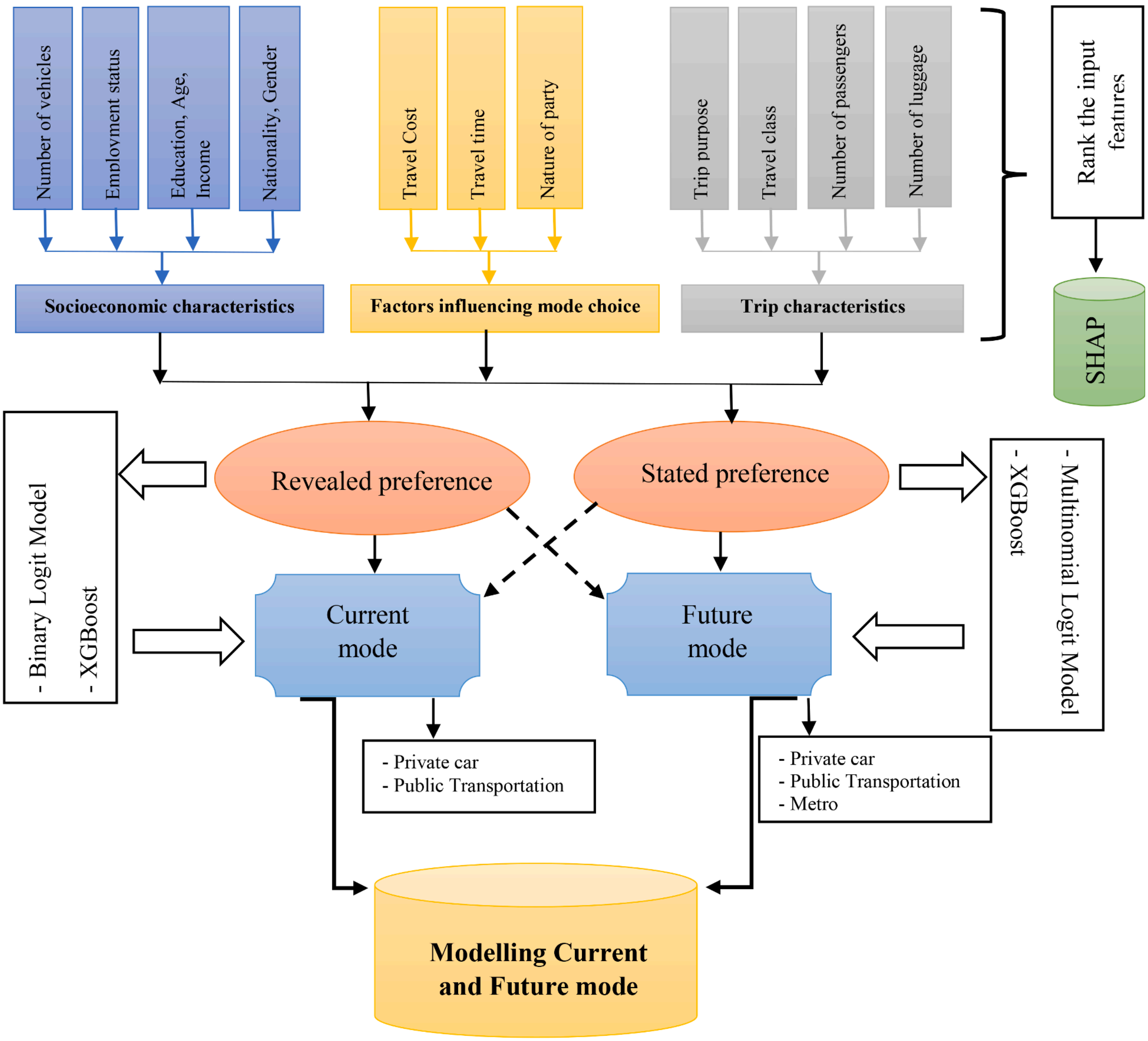


Fig. 3. The research flowchart.

current mode choice, in particular, the significant differences in the influence of trip characteristics and travelers' socioeconomic attributes on current and future mode choices.

3.2.1. Binary logit (BL) model

The significant explanatory variables resulting from the chi-square test were used to investigate the current airport travel mode choice using a BL regression model. This model has been used as a discrete choice model for mode choice studies owing to its capability to predict the possibility of occurrence of a specific event based on independent variables (see for example, (Alhussein, 2011; Reza Mamdoohi et al., 2012)). Utilizing this model deepens our perspective and understanding of the current mode choice in the country before introduction of the metro. The BL model was used to analyze the current mode choice and predict a categorical variable from a set of predictor variables based on the odds ratio between the variables. The dependent variable in the model was given a value of "0" for personal automobile use and "1" for taxi services. One of the advantages of using this model is that it controls for potential determinants, including the traveler's socioeconomic characteristics and trip conditions.

In this study, the probability of selecting a specific mode (*i*) for commuting was equal to the probability that the utility of mode (*i*) was equal to or greater than that associated with an alternative mode (*j*). Therefore, the traveler selects the mode of transportation that yields the maximum utility.

To formulate a BL model, probability is expressed as in the following Eq. (1):

$$P_{n1} = \frac{\exp(\beta X_{1n})}{\exp(\beta X_{1n}) + \exp(\beta X_{2n})} = \frac{1}{1+\exp(\beta X_{2n} - \beta X_{1n})} = \frac{1}{1+\exp(\Delta U)} \quad (1)$$

where,

P_{n1} is the probability that the traveler *n* selects the first mode.

βX_{n1} is the utility function in which the traveler *n* selects the first mode.

βX_{n2} is the utility function in which the traveler *n* selects the second mode.

$\Delta U = \beta X_{2n} - \beta X_{1n} = \sum (a_i - b_i) Z_i$, where Z_i is the i^{th} variable, a_i is the coefficient of the i^{th} variable in βX_{n1}, and b_i is the coefficient of the i^{th} variable in βX_{n2}.

3.2.2. Multinomial logit (MNL) model

The MNL model was utilized to assess the impact of introducing the Doha Metro as a mode of transportation to the airport. This model is based on the random utility theory. The concept underlying the model analysis is that each alternative in the choice set provides the travelers with some utility that can be expressed in terms of measurable or observable characteristics of both the traveler and alternative. The larger the difference in the utility between the two alternatives, the more likely the traveler is to choose the alternative with the higher utility. MNL can be expressed as

$$P(i) = \frac{e^{U_i}}{\sum_{j\in J} e^{U_i}}, \tag{2}$$

where,
P(i) is the probability of a decision maker choosing alternative i;
Ui and Uj are the utilities of alternatives i and j, respectively; and.
J is the number of alternatives.

3.2.3. Extreme gradient boosting (XGBoost)

In this study, the XGBoost (Chen & He, 2020) model was used to classify the travel mode choices according to the travelers' characteristics and trip condition variables. XGBoost is a framework of a gradient-boosted decision tree-based ensemble method based on the idea of additive training; it is used for both regression and classification, and is designed for speed and performance. In this model, each low-depth decision tree is built to minimize a defined loss function; however, more weight is allocated to the cases that are incorrectly predicted by the previously developed trees. Therefore, the final XGBoost model results are collectively determined by the results of all the developed trees. Furthermore, the XGBoost model enables an understanding of the significance of independent variables in explaining the dependent variable (Wang & Ross, 2018; Zaidan et al., 2022). The general unregularized XGBoost algorithm is as follows.

For a given dataset, let $x_i = \{x_{i1}, x_{i2}, \ldots, x_{ij}\}$ represent a vector of the observed values for the i^{th} observation on j features, where $i \in \{1,2,3,\ldots, I\}$ and $j \in \{1,2,3,\ldots,J\}$, and $\hat{y}_i$ is the value of the target outcome for the i^{th} observation. Using K as the number of trees, a tree ensemble model is established to predict the output using additive functions, as shown in Equation (3).

$$\hat{y}_i = \Phi(x_i) = \sum_{k=1}^{K} f_k(x_i), \qquad f_k \in \mathscr{F}, \tag{3}$$

where,
$\mathscr{F}$ = space of regression tree $\{f(x) = w_{q(x)}\}$ ($q: \mathbb{R}^J \rightarrow L$, $w \in \mathbb{R}^L$) and f_k corresponds to an independent structure of each tree q that maps an observation to the corresponding l^{th} leaf and leaf weight w, where $l \in \{1,2,3,..,L\}$.

A prediction output is obtained once q is developed and w is estimated. An observation of a leaf is first assigned to each tree to obtain the output based on the values of the feature set and the sum of the weights of the corresponding leaves. The following objective function is optimized by the algorithm to develop several trees:

$$\mathscr{L}(\Phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{4}$$

where $\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|^2$. Ω is a regularization term that penalizes and controls the complexity of the model and smoothens the final learned weights to prevent overfitting. l is the differentiable convex loss function. This function measures the difference between the observed y_i and predicted $\hat{y}_i$. T is the number of leaves in the tree structure, while the term $\lambda\|w\|^2$ denotes a form of L2 regularization on the leaf weights. For more information on XGBoost, readers can refer to Zopluoglu (2019) and Zaidan et al., (2022).

3.2.4. Shapley additive exPlanations (SHAP)

The SHAP method (Mangalathu et al., 2020) was employed in this study for an in-depth analysis to rank the primary factors that influence the mode choice before and after the introduction of the Doha Metro. SHAP is a coalitional game-theoretic method for describing the performance and output of any machine-learning model. It uses an additive feature attribution method and establishes a link between the optimal credit allocation and local explanations using game theory's traditional Shapley values and their related extensions. The SHAP method helps to explain different supervised learning models. Furthermore, this method assigns an important value to each explanatory variable for a specific prediction (Mangalathu et al., 2020). Mathematically, SHAP is expressed as follows:

$$g(x') = \varnothing_0 + \sum_{i=1}^{M} \varnothing_i x_i' \tag{5}$$

where, $g(x')$ is the explanation model, $x' \in \{0,1\}^M$ i is the coalition vector or simplified features, M is the maximum coalition size, $\varnothing_i \in \mathbb{R}$ is the feature attribution for the feature i; the Shapley values × are the input variable in the model, and × = (x1, x2, …,xp), where p is the number of input variables.

3.2.5. Elasticity analysis

The final step was to perform the direct and cross-elasticity calculations. In this study, elasticity represented the measures of socioeconomic characteristics and trip condition variable sensitivity to changes in mode choice. The aim of this step was to analyze the change in the probability of choosing a specific mode choice for the current and future mode choices based on the change in the percentage of significant independent variables. The direct elasticity analysis was related to the variables of the three transport modes under consideration (private car, taxi services, and metro). The cross-elasticity analysis was related to the significant independent variables considered in the analysis (travelers' socioeconomic characteristics and trip conditions). Direct elasticity and cross-elasticity were computed using the following expressions (Larranaga et al., 2017; Ortúzar, 2011):

$$E_{Piq,Xikq} = \theta_{ik}.X_{ikq}\left(1 - P_{iq}\right), \tag{6}$$

$$E_{Piq,Xjkq} = -\theta_{jk}.X_{jkq}.P_{jq}, \tag{7}$$

where $E_{Piq,Xikq}$ denotes the elasticity of the choice probability of the alternative i for the individual q (P_{iq}) of choosing the mode A_i, considering a marginal change (1% increase) in a given variable X_{ikq} with respect to the base situation. $E_{Piq,Xjkq}$ denotes the elasticity of the probability of choosing the mode A_i, considering a marginal change in the value of the k^{th} variable of the alternative A_j for the individual q, and θ is a constant.

4. Data analysis and results

4.1. Descriptive analysis

The entries in Table 1 show that the main modes of transportation used by the air travelers to arrive at the airport were cars (82.2%), followed by taxis (15.7%), limousines (a form of luxurious taxi) (2%), and buses (1%). For the purpose of analysis, these modes were grouped into two categories because of the insufficient sample size. Taxis, limousines, and the other modes were combined and identified as taxi services, despite the fact that these modes may differ in some of their characteristics. The percentage of passengers who commuted by bus to the airport was very low; hence, this mode was excluded from the model. The high percentage of car ownership may explain the high percentage of users using this mode to travel to the airports. The table shows that 82.6% of the respondents owned one car or more. Another interesting observation is that 67.1% of the travelers used the same mode of transportation each time they commuted to the airport. This implies that cars have been the major mode of transportation for a long time. Furthermore, parking and parking fees are not considered obstacles when using cars as the main mode choice. Although the majority of the

Table 1
Descriptive statistics of travelers' trip conditions.

Airport access mode		Number of bags		Parking passengers' cars		Number of times the passengers traveled from the HIA in the past 12 months	
Car	78.6%	1	16.3%	Airport long-term car park	16.7%	1	34.5%
Taxi	16.1%	2	21.4%	Private long-term car park	11.7%	2	29.2%
Limousine	2.5%	3	14.3%	Airport short-term car park	35.3%	3	18.4%
Bus	0.96%	4	14.5%	Hotel	2.1%	4	8.7%
Others	1.84%	5 or more	33.5%	others	34.2%	5 or more	9.2%
Primary factor influencing the choice of ground access mode		**Using the same mode to travel to HIA**		**Parking charges**		**Purpose of the trip**	
Cost	12.4%	Yes	68.8%	Reimbursed in full	15.4%	Holiday/leisure	63.7%
Journey time	26.5%	No	31.2%	Reimbursed partially	7.8%	Visit relatives/friends	13.1%
Parking charges	14.9%			None	76.8%	Business	12.6%
Number of bags amount	10.8%	**Passengers traveling on plane class**		**Resident or visitor**		Others	10.6%
Public transportation availability	5.6%	Economy class	81.8%	Residents	95.1%		
Nature of party	8.2%	Business class	10.3%	Visitors	4.9%		
Others	21.6%	First class	7.9%				

travelers (80%) are not reimbursed when they use the parking facilities at the airport, many of them still use these facilities owing to the low fees, which are QR 6 ($1.6) per hour for short-term and QR 45 ($12) per day for long-term parking. Table 2 shows the socioeconomic characteristics of travels that were contributed in this study.

4.2. Statistical analysis verses machine learning modeling

In this study, three models were used to investigate and predict the travel mode choices in Doha. To enhance the performance of the models in classifying unseen data and overcoming the overfitting and underfitting problems, a cross-validation method was applied. In this section, the RP data are used to compare the performance of the BL and XGBoost models, while the SP data are used to compare the performance of the MNL and XGBoost models. In K-fold cross-validation, the dataset gathered by both RP and SP questions in the survey was split into K subsets of equal size and then, the model was trained on all but one of the subsets and tested on the rest. This process was repeated until the model was trained on each instance of the given data. The evaluation metrics were then computed and averaged across all the iterations. In this study, cross-validation with ten-folds was conducted. The total errors of the two models and the classification error for each travel mode were averaged across the ten-fold range (Tables 3 and 4). Total error refers to the number of misclassified trips out of the total number of trips. The mode-specific classification error was computed as the number of misclassified trips out of the total number of trips made by the specified mode. Furthermore, a set of well-known evaluation metrics was applied to evaluate the model's classification performance (i.e., accuracy, precision, recall, and F1-score) (Tables 5 and 6). The mean accuracy (i.e., the ability of the classifier to correctly classify the unseen data points into different classes), precision, recall, and F1-score, which is the harmonic mean of precision and recall, were calculated across the ten cross-validation folds.

One challenge usually encountered in modeling travel mode choice is the issue of unbalanced datasets. If this issue is present, the estimation of the three models will be biased, which in turn will lead to a higher prediction error for the classes of mode choice with smaller shares. The majority of the trips are made by personal vehicles, whereas approximately 12.7%, 2%, and 2.1% of the trips are made by taxi, limousine, and other modes, respectively, and only approximately 1% of the trips are made by bus. The dataset is unbalanced and may affect the accuracy and performance of the models when predicting the mode choices with smaller shares, such as buses. Therefore, to reduce the imbalance effect, taxi, limousine, and the other modes were combined into one set, and the bus mode was excluded from the study, as the number of travelers who used this mode was very low.

The models were run 150 times to determine their average prediction accuracy and robustness to data changes. The dataset was randomly divided into training and testing subsets for each run. The training subset encompassed 90% of the data, while the testing subset encompassed the rest. The training and testing errors were averaged for the 150 runs, as shown in Tables 3 and 4. The error was calculated based on the number of journeys predicted to have the wrong travel mode choice out of the total number of journeys. The three models illustrated good overall prediction accuracy for the complete choice set of two or three travel mode choices, with the XGBoost model having lower errors than the other models. Table 3 shows that the XGBoost model has total training and testing errors of 5.74% and 16.73%, respectively, when

Table 2
Descriptive statistics of travelers' socioeconomic characteristics.

Age		Monthly household income (QAR)		Number of vehicles		Employment status	
18–24	18.7%	Less than 5000*	6.8%	None	15.1%	Full-time worker	53.6%
25–34	29.4%	5,000–14,999	38.6%	One car	29.8%	Part-time worker	15.7%
35–44	36.8%	15,000–24,999	18.3%	Two cars	28.3%	Not employed	30.7%
45+	15.1%	25,000 or more	36.3%	Three cars or more	26.8%		
Number of persons traveling		**Education**		**Nationality**		**Number of persons in household**	
1	14.3%	Did not finish high school	1.4%	Gulf Cooperation Council (GCC) including Qatari	30.3%	1	15.8%
2	22.5%	Finished high school	5.3%	North American and Europe	19.4%	2	16.7%
3	19.8%	College	23.8%	Arab (excluding GCC)	23.1%	3	15.7%
4	16.7%	University	64.8%	Asian	18.7%	4	22.5%
5 or more	26.7%	Higher education	4.7%	Others	8.5%	5 or more	29.3%
Gender		**Disability**					
Male	53.4%	Yes	4.2%				
Female	46.6%	No	95.8%				

* **$1 USA = 3.64 QR.**

Table 3
Average testing errors of BL and XGBoost models.

	BL Model			BL Model			XGBoost Model			XGBoost Model		
	Training Error			Testing Error			Training Error			Testing Error		
	Mean	Variance	STD	Mean	Variance	STD	Mean	Variance	STD	Mean	Variance	STD
Total	19.94%	0.001	0.0327	18.02%	0.0013	0.0364	5.74%	0.001	0.0471	16.73%	0.001	0.030
Private Car	12.02%	0.009	0.0315	12.13%	0.0013	0.0355	3.26%	0.005	0.0615	11.49%	0.001	0.037
Taxi services	28.12%	0.007	0.0822	28.50%	0.0073	0.0856	16.41%	0.015	0.0672	26.07%	0.002	0.047

Table 4
Average testing errors of MNL and XGBoost models.

	MNL Model			MNL Model			XGBoost Model			XGBoost Model		
	Training Error			Testing Error			Training Error			Testing Error		
	Mean	Variance	STD	Mean	Variance	STD	Mean	Variance	STD	Mean	Variance	STD
Total	17.84%	0.001	0.0327	19.68%	0.00031	0.0176	7.18%	0.0014	0.0471	12.46%	0.0023	0.0093
Private Car	2.92%	0.009	0.0315	4.06%	0.00174	0.0418	1.35%	0.0047	0.0615	2.76%	0.0002	0.0150
Taxi services	2.12%	0.007	0.0822	3.70%	0.00275	0.0525	1.61%	0.0148	0.0672	2.38%	0.0004	0.0187
Metro	32.76%	0.015	0.3280	34.27%	0.00335	0.0578	27.93%	0.0035	0.0498	32.26%	0.0012	0.0341

Table 5
Average of the evaluation metrics of the classification models BL vs. XGBoost.

	BL Model				XGBoost Model			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
Total	80.893%	81.97%	81.98%	81.77%		83.26%	83.27%	83.15%
Private Car		84.75%	87.88%	86.20%	82.25%	85.84%	88.52%	87.12%
Taxi services		77.04%	71.50%	73.89%		78.67%	73.93%	76.09%

Table 6
Average of the evaluation metrics of the classification models MNL vs. XGBoost.

	MNL Model				XGBoost Model			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
Total	80.32%	82.38%	80.32%	82.32%	87.54%	88.83%	87.53%	86.99%
Private Car		80.28%	85.93%	84.92%		83.07%	97.24%	89.58%
Public Transportion (Taxi, limousine)		84.59%	86.30%	89.49%		86.52%	97.62%	91.70%
Metro		90.26%	63.72%	73.56%		96.89%	67.74%	79.68%

considering two different modes, whereas the model has a total training and testing error of 7.18% and 12.46%, respectively, when considering three mode choices (see Table 4). The amount of these errors is less than that resulting from the other two models. Furthermore, the three models have a high total rate of accuracy when considering either two or three modes, and XGBoost performs better than the other two. This applies to each mode where the prediction accuracy is high, as shown in Tables 5 and 6. The BL and MNL models can replicate the shares of all the choices; hence, the predicted mode shares resulting from using these two models are similar to the observed ones, as shown in Tables 3 and 4. However, the XGBoost model performed well in predicting the mode choice, and it showed its capability to lower the overall prediction error. Although the total accuracy percentage was high in the three models, it was affected by unbalanced data.

4.3. Assessing current and future airport travel mode choice

The first step in assessing the current mode choice was to map the importance of the travelers' socioeconomic characteristics as well as the trip condition variables on the mode choice based on the rank of the answers to the survey questions, as shown in Table 7. The impact of socioeconomic characteristics and trip conditions on the mode choice can be further analyzed using the SHAP method. In this method, the order of the answers to each question of the data instances is presumed to behave as players in a coalition (Zaidan et al., 2022). Fig. 4 shows the mapping of these answers to understand how these variables influence the mode choice. The importance factors of the input variables, shown in Fig. 4, were calculated as the average of the absolute Shapley values per feature across the RP data. The input features were ranked based on their importance according to the mean SHAP value. The higher the mean SHAP value, the more important was the variable. The figure also shows the importance of each input variable for the mode choice, that is, private cars and taxi services. Each point in the Figure represents the Shapley value for the input variables. The y-axis represents the order of importance of the input variable from top to bottom. Furthermore, each point in the figure related to the input variable is colored by the value of the input variable from low (blue) to high (red). The density of the points indicates the distribution of the dots in the RP dataset. The number of times the public transport is used (taxi services) is the most important input variable in the mode choice. Fig. 4 shows that the higher is the "times using public transport" value, the higher is its SHAP value, and the larger is its impact on the mode choice selection. The respondents were asked how often they used public transportation (taxi services), and their answer choices are shown in Table 7. The figure shows that those who answered the first choices (i.e., daily/weekly) tended to use public transportation (blue color), while those who used it

Table 7
Some questions of the survey and their designated answers.

Questions	Respondents' answers choices
Times Using of public transportation	(1) Daily; (2) Weekly; (3) Monthly; (4) Annually; (5) Don't use.
Nationality	(1) Qatari; (2) Gulf Cooperation Council (GCC) citizens; (3) European; (4) North American; (5) Arab, (6) Asian; (7) Others
Number of persons in household	(1) 1; (2) 2; (3) 3; (4) 4; (5) 5 or more
Vehicle ownership	(1) 0; (2) 1; (3) 2; (4) 3 or more
Employment status	(1) Full-time; (2) Part-time; (3) Not employed
Number of travelers	Stated by the interviewed person
Age	(1) 18–24; (2) 25–34; (3) 35–44; (4) 45–54; (5) 55–64; (6) 65 or older
Reimbursed parking fees	(1) Full reimbursement; (2) Partial reimbursement; (3) No reimbursement
Airport access mode	(1) Car; (2) Taxi; (3) Limousine; (4) Bus; (5) Others
Trip purpose	(1) Vacation; (2) Visit family/friends; (3) Business; (4) Others
Monthly household income (QAR)	(1) Less than 5,000; (2) 5,000–9,999; (3) 10,000–14,999; (4) 15,000–19,999; (5) 20,000–24,999; (6) 25,000–29,999; (7) 30,000–34,999; (8) 35,000 or more
Occupation	(1) Official; (2) Manager/Specialist; (3) Sales/services; (3) Manufacturing; (4) Others
Using the same mode to travel to HIA	(1) Yes; (2) No
Number of luggage	Stated by the interviewed person
Gender	(1) Male; (2) Female
Class category	(1) Economy class; (2) Business class; (3) First class

less often tended to use the car as the main mode choice of travel (red color).

In terms of nationality, Qataris and Gulf Cooperation Council (GCC) citizens preferred to use private cars more than those belonging to the other nationalities did. This is due to their high level of income and their preference for using personal automobiles in their commuting activities. Asians are less likely to use cars, perhaps because of their low incomes. Furthermore, the figure shows that fewer travelers and younger travelers tend to use public transportation. However, as the number of owned vehicles increases, travelers tend to use their cars as their main mode choice. This figure only shows the tendency of the travelers to choose their mode but does not show which factors are significant in choosing the mode.

Fig. 5 shows the impact of introducing the Doha Metro as a new mode choice for feature importance. The metro will take shares either from those using cars or taxi services as the main modes of travel. The main socioeconomic features that encourage travelers to switch to using the metro include age, nationality, occupation, vehicle ownership, and income. Class category and trip purpose are among the most important trip condition variables that encourage travelers to use metro services. The figure shows that the importance of the features underwent some changes. The "times using public transportation" is still the most important feature in the mode choice. However, the Shapley value for the metro is very low compared with the use of private cars and taxi services (public transportation). The number of persons in the household is the second most important feature to determine the mode choice after introducing the metro, whereas it was the person's nationality prior to that. Although these figures show the importance of the input variables in mode choice selection, the SHAP method does not indicate whether these features are significant in mode choice selection. Therefore, there is a need for statistical analysis to determine the significant variables.

5. Statistical analysis

To assess the current mode choice, we considered two alternative airport travel modes: personal automobiles and taxi services. A BL model

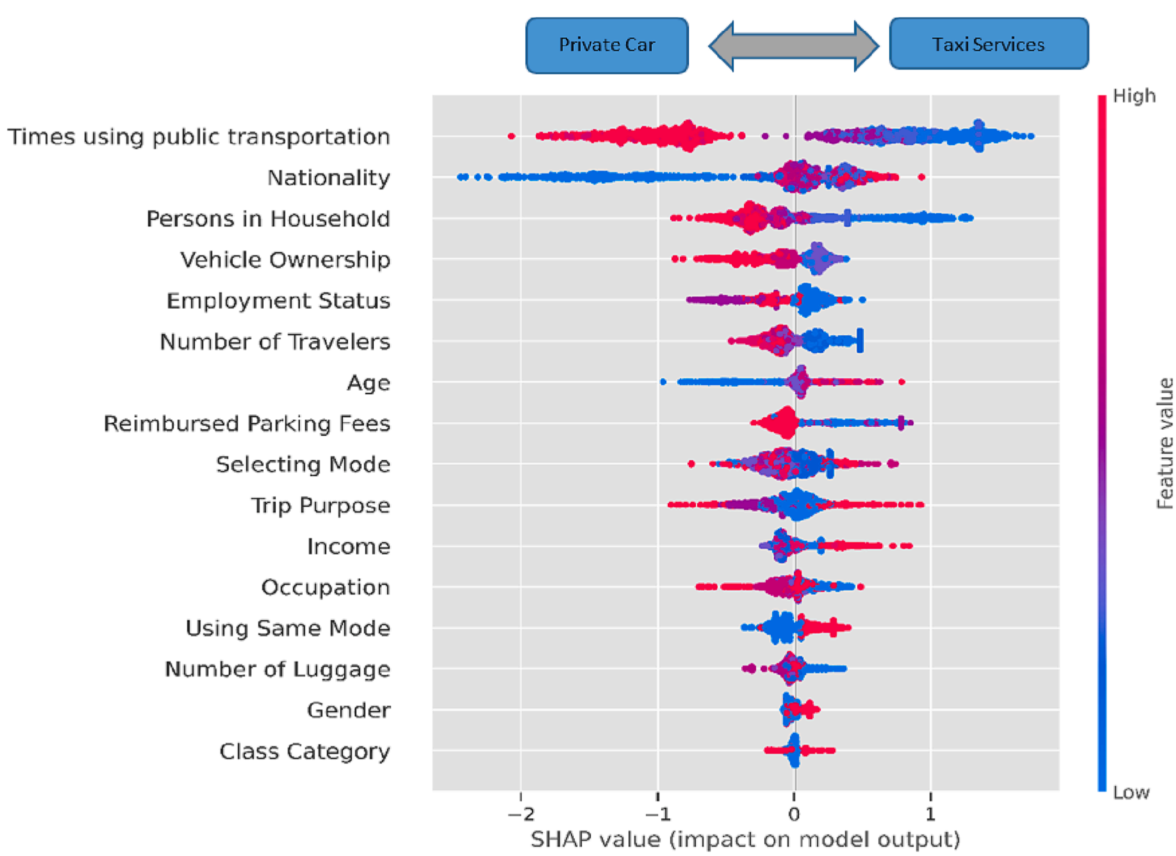


Fig. 4. Interpretation of features of binary classification.

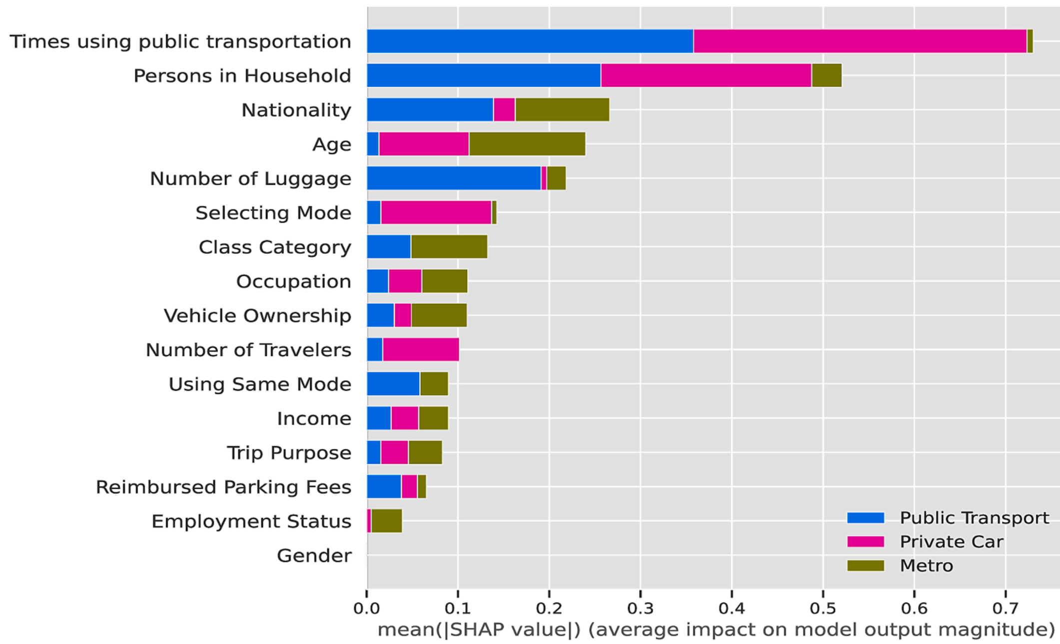


Fig. 5. Interpretation of features of multiclass classification.

was used in modeling this current mode choice. However, MNL and XGBoost models were designed and used for modeling the airport travel mode choice after introduction of the metro. The aim was to identify the elements that might influence car users to shift to this mode. In the model, the dependent variable was given a value of "1″ for cars, "2″ for taxi services, and "0″ for metro use. The variables used were the same as those used to assess the current mode choice using the BL model.

The explanatory variables identified in the survey were extensively evaluated to design an appropriate travel-mode choice model. These variables were assessed to determine and identify the significant variables that effectively augmented the data for the mode-choice model using the chi-square test. Some of the categories in these explanatory variables were combined to ensure reliability in conducting the chi-squared test by ensuring that each cell in the cross-tabulation had a count of five or more.

The chi-square test was used to compare the socioeconomic factors and trip characteristics that were significant in selecting the current or future mode choice. Based on the results of the chi-squared test, all the socioeconomic factors and trip characteristics were determined to be significant in explaining the current mode choice (Table 8). However, the results of the chi-square test showed that finding a place to park a car is not a significant trip characteristic feature in selecting the future mode choice. This also applies to gender as a socioeconomic factor, which affects the use of future mode choice. Based on the chi-square test results, all the variables found to be insignificant were excluded from the model. Other variables that had no direct influence on the selection of a specific mode were omitted from the model. One of these variables was the reimbursement of parking fees because those who used taxi services would not pay the parking fees. The number of respondents was considered a continuous variable, whereas the other variables were treated as categorical variables.

A comparison of the driving factors influencing the mode choice before and after implementation of the Doha Metro revealed that some of these factors were the same. The significant factors that negatively influenced the current mode choice were the number of travelers, full parking fee reimbursement, journey time, nationality, vehicle ownership (owning one or more cars), age (35 – 44 years), and full-time employment. However, the significant factors that positively influenced the current mode choice included the number of luggage, age (18 – 34 years), average household income, vehicle ownership (owning no

Table 8
Results of Chi-square test on main variables influencing current airport travel mode and future use of Doha Metro.

Factors influencing mode choice	Current mode choice (Personal vehicle vs. taxi services)		Future mode choice (Personal vehicle vs. taxi services vs. Doha Metro)	
	X^2 value	*P* value	X^2 value	*P* value
Trip characteristics				
Number of people traveling	57.835	0.000	76.257	0.000
Number of bags	18.853	0.008	42.386	0.000
Traveling class	42.491	0.000	14.926	0.008
Trip purpose	8.622	0.039	9.627	0.036
Place to park the car	93.738	0.000	4.217	0.269
Reimbursement of parking fees	48.249	0.000	8.470	0.009
Trip conditions (e.g., time)	85.545	0.000	24.546	0.000
Mode of transportation currently used	–	–	0.276	0.763
Socioeconomic characteristics				
Nationality	184.825	0.000	16.239	0.012
Age	116.653	0.000	25.224	0.000
Gender	64.194	0.000	0.827	0.485
Income	95.383	0.000	10.852	0.036
Vehicle ownership	204.639	0.000	34.898	0.000
Employment status	137.270	0.000	8.539	0.026
Disability	10.482	0.000	34.629	0.000

cars), employment status, and part-time employment (see Table 9). Upon introduction of the metro, some of these factors became significant in using the car but not in using the taxi services mode, and vice versa. Furthermore, the importance of these factors changed, as shown in Fig. 6. In that case, and in terms of using personal vehicles, the significant factors affecting the negative mode choice were the flight class category (economy class), nationality (North American and Europe), and car ownership. The significant factors affecting the positive mode choice were parking fee reimbursement (full and partial), nationality, and age (Table 10). In terms of taxi services, the significant factors affecting the negative mode choice were the number of luggage trips, journey time, age, car ownership (owning one car), and working part-time.

Table 9
Model parameter estimates for current mode choice (private car = base category).

	β	Std. error	Sig.	Exp (β)
Number of travelers (5 travelers or more = reference category)				
1 traveler	-0.841	0.465	0.023	0.737
2 travelers	-0.991	0.415	0.016	0.861
3 travelers	0.593	0.393	0.131	1.809
4 travelers	0.617	0.381	0.105	1.853
Number of luggage (5 luggage or more = reference category)				
1 luggage	-0.777	0.398	0.051	0.860
2 luggage	−1.351	0.357	0.000	0.659
3 luggage	−1.022	0.344	0.003	0.622
4 luggage	−1.065	0.327	0.001	0.445
Trip purpose (Business = reference category)				
Holiday/leisure	-0.720	0.326	0.075	0.912
Visit relatives/friends	−1.373	0.431	0.012	0.510
Class category (First class = reference category)				
Economy class	0.814	0.284	0.009	1.598
Business class	−1.062	0.390	0.078	0.646
Reimbursed parking fees (not reimbursed = reference category)				
Reimbursed in full	−1.052	0.240	0.000	0.564
Reimbursed partially	-0.781	0.319	0.044	0.784
Primary factor				
Cost	−1.280	0.493	0.000	0.455
Journey time	-0.550	0.235	0.019	0.577
Nationality (Others = reference category)				
Qatari and GCC	-0.959	0.346	0.000	0.526
North American and Europe	0.366	0.280	0.191	1.042
Arab (excluding GCC)	-0.611	0.279	0.028	0.543
Asian	1.294	0.297	0.000	1.574
Age (45 years or older = reference category)				
18 – 24 years	-0.789	0.387	0.128	0.955
25 – 34 years	-0.935	0.262	0.066	0.874
35 – 44 years	−1.201	0.245	0.001	0.670
Income (25,000 or more = reference category**)**				
Less than 5,000	1.253	0.421	0.003	3.500
5,000 – 14,999	0.377	0.209	0.072	1.457
15,000 – 24,999	-0.223	0.239	0.350	0.800
Vehicle ownership (Three vehicles or more = reference category)				
One Vehicle	−1.371	0.367	0.000	0.839
Two vehicles	−1.599	0.262	0.000	0.549
Employment status (Not employed = reference category)				
Full-time worker	-0.766	0.346	0.005	0.628
Part-time worker	0.952	0.406	0.075	1.386

*Cox and Snell R Square: 0.411, Nagelkerke R Square: 0.526.
The literature suggests that values of 0.2 to 0.4 for R^2 represent an excellent fit (Abulibdeh, 2018).

Table 10 shows the parameter estimates that indicate the influence of each factor for travelers' mode choice of Metro relative to private car and taxi services. The coefficient (β) values are the estimated MNL regression for the models. The negative values of the coefficient indicate that the factors decrease the likelihood of that response category with respect to the reference category. The table also shows the Exp(β) values, which represents the odds ratio for each category of the predictors.

Fig. 6 ranks the importance of the socioeconomic factors and trip characteristics in selecting the mode choice prior to and after introduction of the metro using the F-score. The XGBoost model consists of a number of boosted trees that represent the estimators. Each feature is represented by a node in the tree and the number of nodes split to make the final decision represents the F-score. The figure clearly shows that nationality and number of times the public transportation is used are the most important socioeconomic features and trip characteristic factors in selecting the mode choice prior to the introduction of the metro, respectively. However, age, number of persons in the household, and number of times that passengers used to use taxi services were the most important socioeconomic features and trip characteristics in selecting the mode choice. Introducing Metro as a mode choice shows that the variables that affect the mode choice include income as the most important variable, followed by nationality and the number of pieces of luggage, as shown in Fig. 6. This indicates that low-income travelers may choose Metro during their trip to the airport.

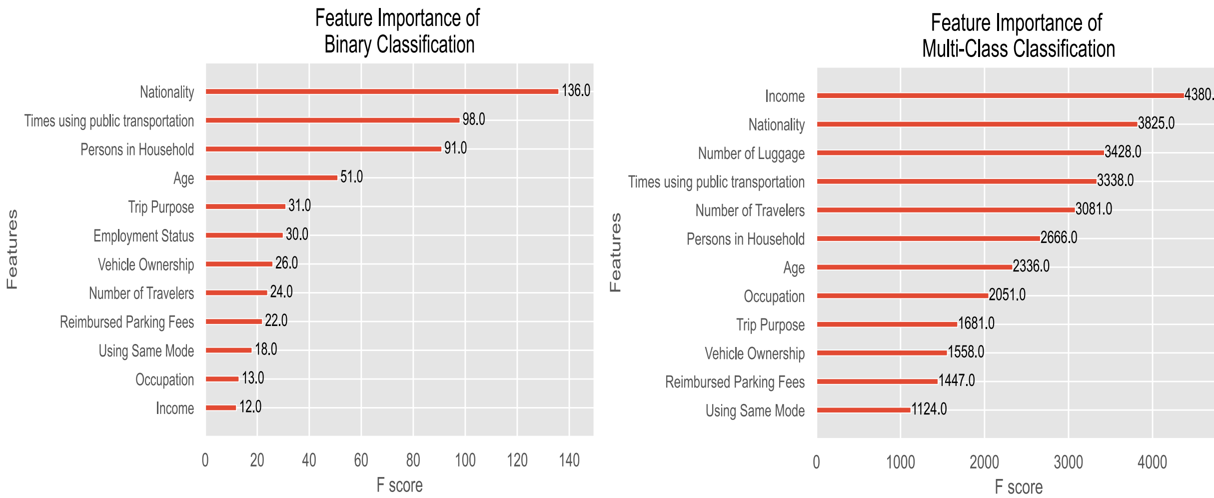
6. Trip characteristics

Some explanatory variables related to trip characteristics were found to be significant and influence the current and future mode choices as shown in Table 9 and Table 10. The number of travelers in the different groups was a significant explanatory variable influencing the current and future mode choices. The 5 or more travelers category is the reference category for the various number of travelers categories. Compared with this reference category, in the current mode choice, the coefficient (β) for the variable "number of travelers" was negative for 1 traveler and 2 travelers categories (β = − 0.841, β = − 0.991). This implies that a unit increase in the number of travelers who travel alone or travel in a group of two decreases the likelihood of choosing taxi by 26.3% and 13.9%, respectively. On contrast, the coefficient (β) is positive but not significant for the other categories of the "number of travelers" variable. Introducing the Metro attracts some travelers who used to use their cars or taxis. The entries in Table 10 shows that the coefficient is negative for 1 and 2 travelers categories and positive for 3 and 4 travelers categories. The table shows that a unit increase in the number of travelers who travel alone or in a group of 2 decreases the likelihood of choosing cars by 21.9% and 7.2% and of choosing taxi by 13.9% and 27.9%, respectively. Conversely, a unit increase in the number of travelers who travel in a group of 3 or group of 4 increases the likelihood of choosing cars by 28.1% and 42.3% and of choosing taxi by 21.5% and 42.6%, respectively.

Similarly, the number of luggage items was significant in determining the current and future mode choices. The 5 or more luggage category is the reference category for the various number of luggage categories. The coefficient of this variable was negative in determining the current mode choice for all categories as shown in Table 9. The entries in the table shows that a unit increase in the number of luggage decreases the likelihood of choosing taxi by 14%, 24.1%, 37.8%, and 55.5% for those who travel with 1, 2, 3, and 4 luggage, respectively. Introducing the Metro has no significant effect of travelers who travel with 1 luggage using either their cars or a taxi. On the contrary, traveling with more than one luggage increases the probability of choosing either a car or a taxi for traveling at the expense of the metro in the future. The entries in the table shows that a unit increase of carrying 3 or 4 luggage increase the likelihood of choosing car by 25.3% and 84.1% and taxi by 28.4% and 42.8%, respectively.

Trip purpose is an important and significant variable on determining the current and future mode choices. The business category was considered as the reference category for trip purpose. In the current mode choice, the coefficient (β) was negative for traveling for the purposes of holiday, leisure, and visiting relatives or friends. Table 9 shows that a unit increase of travelling for holidays or visiting relatives decreases the likelihood of choosing taxi by 8.8% and 49% for traveling for spending a holiday and visiting relatives, respectively. In contrast, introducing the Metro will not encourage travelers to use it in their commuting. A unit increase of travelling for holidays or visiting relatives increases the likelihood of choosing a car by 136.9% and 69.3% and choosing a taxi by 35.1% and 23% for traveling for holiday or visit relatives, respectively.

The flight class category was another variable related to trip characteristics that was significant and had influence on the current mode



(a) and (b).

Fig. 6. F-score of feature importance: (a) prior to introduction of the metro, and (b) after the introduction of the metro.

Table 10
Model parameter estimates for using the Doha Metro, future mode choice, (Metro is the reference).

	Cars*				**Taxi services***			
	β	**S.E.**	**Sig.**	**Exp(β)** (odds ratio)	**β**	**S.E.**	**Sig.**	**Exp(β)** (odds ratio)
Number of travelers (5 travelers or more = reference category)								
1 traveler	-0.867	0.275	0.012	0.781	-0.936	0.326	0.083	0.861
2 travelers	-0.558	0.286	0.047	0.928	-0.852	0.241	0.062	0.721
3 travelers	0.747	0.670	0.009	1.281	0.639	0.451	0.023	1.215
4 travelers	1.153	0.266	0.004	1.423	0.955	0.379	0.011	1.426
Number of luggage (5 luggage or more = reference category)								
1 luggage	-0.753	0.266	0.184	0.923	-0.641	0.236	0.213	0.741
2 luggage	0.958	0.286	0.021	1.148	-0.859	0.341	0.113	0.852
3 luggage	1.126	0.225	0.015	1.253	0.798	0.459	0.036	1.284
4 luggage	1.440	0.306	0.002	1.841	1.153	0.396	0.004	1.428
Class category (First class = reference category)								
Economy class	−1.136	0.368	0.007	0.695	−0.720	0.321	0.019	0.526
Business class	1.366	0.456	0.004	1.572	1.643	0.428	0.000	1.736
Trip purpose (Business = reference category)								
Holiday/leisure	1.237	0.257	0.003	2.369	0.993	0.461	0.015	1.351
Visit relatives/friends	0.926	0.361	0.034	1.693	0.837	0.285	0.003	1.23
Reimbursed parking fees (not reimbursed = reference category)								
Reimbursed in full	0.510	0.248	0.005	1.948	–	–	–	–
Partially	0.461	0.313	0.094	1.631	–	–	–	–
Primary factor								
Cost of the trip	-0.735	0.368	0.012	0.564	-0.604	0.361	0.021	0.753
Journey time	−1.152	0.274	0.003	0.726	−0.497	0.260	0.036	0.608
Nationality (Others = reference category)								
Qatari and GCC	0.867	0.275	0.002	1.381	−0.787	0.338	0.034	0.629
North American and Europe	0.757	0.308	0.005	1.124	−0.899	0.295	0.036	0.805
Arab (excluding GCC)	0.692	0.279	0.018	1.096	1.211	0.436	0.006	1.250
Asian	−1.169	0.293	0.044	0.584	1.357	0.329	0.002	1.393
Age (45 years or older = reference category)								
18 – 24 years	-0.807	0.248	0.080	0.793	0.694	0.358	0.042	1.152
25 – 34 years	0.864	0.373	0.014	1.166	-0.874	0.438	0.032	0.891
35 – 44 years	1.294	0.243	0.033	1.516	−0.658	0.280	0.019	0.518
Income (25,000 or more = reference category)								
Less than 5,000	-0.751	0.358	0.081	0.673	0.692	0.362	0.008	1.352
5,000 – 14,999	-0.560	0.200	0.057	0.942	0.854	0.239	0.036	1.138
15,000 – 24,999	1.281	0.217	0.007	1.324	0.546	0.451	0.002	0.731
Vehicle ownership (Three vehicles or more = reference category								
One Vehicle	0.792	0.250	0.002	1.453	-0.672	0.269	0.026	0.881
Two vehicles	1.544	0.250	0.000	2.580	−0.874	0.422	0.000	0.460
Employment status (Not employed = reference category)								
Full time	1.197	0.365	0.022	1.682	-0.873	0.337	0.005	0.840
Part time	-0.626	0.247	0.089	0.732	-0.632	0.469	0.076	0.535

* Reference is the Metro
* Cox and Snell R Square: 0.453, Nagelkerke R Square: 0.518, McFadden: 0.184.

choice and on the future mode choice. Regarding the current mode, the coefficient (β) is positive for economy class and negative for business class. This implies that a unit increase in traveling on economy class increases the likelihood of choosing taxi by 59.8%, and decreases the likelihood of choosing taxi by 35.4% when traveling on business class. Introducing the Metro decreases the likelihood of using the car and taxi when traveling on the economy class and increases the likelihood of using the car and taxi when traveling on business class.

6.1. Reimbursement of parking fees

Another significant factor that influences the current mode choice is the reimbursement of parking fees. The coefficient of full reimbursement for the parking fees variable had a negative sign prior to introduction of the Doha Metro and a positive sign after its introduction. The no reimbursement category is the reference category for the various reimbursed parking fees categories. Prior to the introduction of the metro, the results show that full reimbursement for the parking fees variable has an odd ratio of − 1.052, meaning that a unit increase in reimbursing the parking fees decreases the likelihood of choosing taxi by 43.6%. After the introduction of the metro, full parking fee reimbursement was still significant, and the coefficient was positive with an odds ratio of 1.948 for reimbursed in full and 1.631 for partial reimbursement. This means that a unit increase in full or partial parking reimbursement increases the likelihood of using the car by 94.8% and 63.1%, respectively. These results indicate the importance of this variable in selecting the travel mode choice. Although full parking fee reimbursement is significant before and after the introduction of the metro, this feature is less important than the other features, as shown in Fig. 6.

6.2. Trip journey cost and time factors

Journey cost and time was a significant factor in determining the current mode choice. An increase in journey cost and time results in a decrease in public transportation usage. The coefficient of the journey cost and time variables was negative with an odds ratio of − 1.280 and -0.550, respectively. This indicates that a unit increase in the cost and time decreases the likelihood of choosing taxi by 54.5% and 42.3%, respectively. The journey cost and time variables were also significant in influencing the future airport mode choices, including metro. Concerning personal vehicle usage, the impact of journey cost and time was the same as that of the current airport mode choice.

6.3. Socioeconomic characteristics

Personal socioeconomic characteristics also influence the current and future airport travel mode choices. Nationality was one factor that significantly affected current and future airport mode choices. Qatari and GCC nationals currently use private automobiles to travel to airports more than those belonging to other nationalities do. Being a Qatari or GCC resident reduces the probability of using taxi services in the present analysis. The entries in Table 9 shows that a unit increase of being Qatari or from the GCC countries decreases the probability of choosing taxi by 47.4%, while being from Arab nationalities decreases the likelihood of choosing taxi by 15.7%. On the contrary, being from East or Southeast Asia increases the likelihood of choosing taxi by 57.4%. Thus, having the metro available will not encourage Qataris and GCC residents to switch from using private automobiles. For these nationalities (Qataris and GCC), the odds ratios are 1.381 and 0.629 for using cars and taxi services, respectively. This means that a unit increase in being Qatar or from the GCC countries increases the probability of choosing a car by 38.1% and decreases the probability of choosing a taxi by 37.1%. The same trend applies for the North American and Europe nationalities. Conversely, Asians are use taxis more than private cars and they are more willing to use the Metro. In terms of the current mode choice, a unit increase of being Asian increases the probability of choosing a taxi by 57.4% and by introducing the Metro this probability decreases to 39.3%, while the probability of choosing a car decreases by 41.6%.

Age was a significant explanatory variable in explaining the current travel mode choice, but it also had influence on the future travel mode choice. For the current mode choice, as age increased, the probability of selecting taxi services decreased (Table 9). The 45 years or older age category is the reference category for the various age categories. Air passengers aged 25–44 years old are less willing to choose tax over cars. A unit increase in travelers in 25–34 age group decreases the likelihood of choosing taxi by 12.6%, while a unit increase in 35–44 age group decreases the likelihood of choosing taxi by 33%. When assessing the future mode choice, the 35–44 year age cohort variable was significant. However, air passengers aged 35–44 years preferred to switch from using taxi services to using the metro or their private cars more than the other age groups (Table 10).

Household monthly income was another significant indicator of the current airport travel mode choice and had significant influence on the future choices. In terms of the current situation, as income increases, the probability of selecting personal automobiles increases. Low-income travelers (Income < QR 5000; 1 USD = 3.68 QR) were more willing to use taxi services than high-income travelers (Income > QR 25,000), with an odds ratio of 1.781. As household income increases, the tendency to use taxi services decreases. For example, the coefficient of the air passengers with an average monthly income between QR 15,000 and QR 24,000 was negative (β = − 0.223), with an odds ratio of 0.738, meaning that a unit increase in being within this income category decreases the probability of choosing taxi services by 26.2%. The same pattern is found when introducing the Metro whereby as income decreased, the tendency to use personal vehicles decreased.

In a wealthy developing country such as Qatar, owning more than one vehicle is normal. This may have influenced the current and future use of taxi services. In assessing the influence of the car ownership variable on the airport mode choice, the results indicated that this variable was significant in selecting the current mode (see Table 9). As the number of vehicles per household increases, the tendency to use taxi services decreases. For example, the sign of the coefficient of owning one vehicle was negative, which meant that a unit increase of owing one vehicle decreases the likelihood of choosing taxi by 16.1%. Vehicle ownership is still a significant factor in selecting future airport travel mode choices. Introducing the Metro will influence the taxi services share of mode choice. For example, a unit increase in owing two cars decreases the probability of choosing taxi by 54%, where the in the current mode choice owing two cars decreases the probability of choosing taxi by 45.1%. This implies that some travelers who used to select taxi to travel to the airport will choose the Metro once it is introduced.

Employment status was another significant variable ($P = 0.05$ and $P = 0.1$) for the current airport travel mode choice and for the future choice after Doha Metro came into service. In terms of current mode choice, full-time employees were less likely to use taxi services, as the coefficient of this variable was negative, while part-time employees preferred to use taxi services, with an odds ratio of 1.386. However, the availability of the metro encourages part-time employees to switch to using this new mode to travel to the airport.

6.4. Elasticity analysis

Direct and cross-elasticity analyses were performed to detect and understand the changes in the probabilities of selecting a specific transport mode choice based on the occurrence of percentage changes in the independent variables. Table 11 shows the results of the elasticity of mode choice probability calculations based on the significant variables presented in Table 10, excluding full and partial parking reimbursements, because these variables are not applicable when using taxi services or the metro. The direct and cross-elasticity values in Table 11 show that travelers using private cars are highly sensitive to the

Table 11
Elasticity analysis.

Variable		Private car	Taxi services	Metro
Number of luggage	Private car	1.54	−0.39	−0.28
	Taxi services	−0.14	0.94	−0.28
	Metro	−0.14	−0.39	1.32
Economy class	Private car	−0.22	0.63	0.47
	Taxi services	0.16	−0.15	0.47
	Metro	0.16	0.63	−0.61
Journey time	Private car	3.79	−2.13	−1.18
	Taxi services	−3.39	1.81	−1.18
	Metro	−3.39	−2.13	3.37
Qatari and GCC	Private car	2.83	−1.36	−1.53
	Taxi services	−1.47	1.88	−1.53
	Metro	−1.47	−1.36	1.76
North American and Europe	Private car	−0.23	0.47	0.52
	Taxi services	0.13	−0.64	0.52
	Metro	0.13	0.47	−0.64
35 – 44 years	Private car	0.59	−0.26	−0.36
	Taxi services	−0.35	0.48	−0.36
	Metro	−0.35	−0.26	0.58
1 car	Private car	−0.65	0.12	0.15
	Taxi services	0.23	−0.35	0.15
	Metro	0.23	0.12	−2.21
2 cars	Private car	3.45	−2.27	−1.34
	Taxi services	−1.42	2.49	−1.34
	Metro	−1.42	−2.27	0.63
Part time	Private car	0.21	−0.16	−0.13
	Taxi services	−0.09	0.33	−0.13
	Metro	−0.09	−0.16	0.25

number of pieces of luggage, the duration of the trip to the airport, nationality, and car ownership (owning two cars). These variables appear to be the most important in the mode of choice to travel to the airport. For example, an increment of 1% in the time of the trip to the airport represents a 3.79% higher market share for those who use private cars. Conversely, a 1% reduction in the trip time to the airport should increase the demand for the other alternatives by 3.39%. Changes in the remaining variables affect the probability of selecting the corresponding mode to a lesser extent. The elasticity results represent an important tool for analyzing demand response and model competition when introducing a new mode of transportation. The elasticity analysis in this study allows transport planners to assess the demand response to the different mode choices considered in this study.

7. Conclusion and policy implications

Modeling travel mode choice is a dynamic and significant step in travel demand forecasting in Doha city due to the transport infrastructure development in preparation for the FIFA 2022 World Cup and the introduction of the metro as a new mode of transportation. In this study, three models were used to investigate and predict travel mode choices in Doha. These models are the extreme gradient boosting (XGBoost), binary logit (BL), and multinomial logit (MNL) models. The results of these models were examined by comparing their average multiclass prediction errors. The factors affecting the travel mode choice were categorized into two groups: travelers' characteristics and trip conditions. A set of independent variables were selected when developing the three models using the entire dataset. These variables were selected by developing a chi-square test to determine the statistically-significant variables to improve the goodness-of-fit of the models. MNL maintained a high consistency between the training and testing errors because of its capability to avoid overfitting. The XGBoost and MNL models explain the relationships between the independent variables and travel mode choices. Different factors were found to influence the current mode choice, and many of these factors continued to influence future mode choice after introduction of the Doha Metro. Furthermore, this study is among the studies that use statistical analysis and machine-learning techniques to investigate travel mode choice, and it presents a relatively comprehensive range of independent variables that are ready for practical use.

Analysis of the driving forces influencing airport travel mode choice before and after introduction of the Doha Metro revealed many common forces. The significant factors influencing both the current and future mode choices were the number of travelers, number of bags, trip purpose, flight class category, parking fees, full parking fee reimbursement, cost of the trip, journey time, nationality, age, average household income, vehicle ownership, and employment status. Personal automobiles were used extensively by travelers who traveled to airports more frequently and for different trip purposes. However, other modes of transportation were used by those traveling for holidays or leisure purposes, and hence, they will be more interested in using the metro in the future. Another key factor in determining the travel mode choice to the airport was parking charges — mainly long-term parking charges — where many travelers were unwilling to take their personal car and park it at the airport and therefore, preferred to use other modes to get to the airport. However, parking charge reimbursement was a key factor in encouraging the commuters to use their personal vehicles and park them at the airport. Travel time was another significant factor that influenced mode choice. The results revealed that commuting time influenced approximately 23% of the commuters to use their personal automobiles and avoid using other modes, particularly buses. These findings can help public authorities to develop transportation-related policy measures. Of course, the policies adopted by the Ministry of Transportation in Qatar must be consistent with other facets of the government transportation policy. However, the findings in this study can help the Ministry design incentives to encourage travelers to use the Doha Metro when it comes into operation and discourage the use of private cars. The models used in this study show high prediction accuracy for travel mode choices in Doha city before and after the introduction of the new metro. The performance of the XGBoost model substantially exceeds that of both the BL and MNL models in predicting and improving the accuracy of predicting the mode choice.

Future work should take place after the metro becomes fully operational and real data on metro ridership are gathered. Furthermore, it is necessary to investigate metro accessibility for both residents and visitors to the country. Once operated, it is recommended that the authorities undertake measures to augment the metropolitan metro system to cover additional areas within the country, predominantly residential areas, with the aim of enhancing commuter convenience and mobility. Such measures may include intensifying the frequency of metro services, particularly during peak hours, to minimize wait times for travelers. Strategies and policies that are deemed important for improving connectivity with other modes of public transportation are important and should be pursued in future research. Investigating these factors will contribute to the success of metro attractiveness.

CRediT authorship contribution statement

Ammar Abulibdeh: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing.

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

The authors do not have permission to share data.

Acknowledgement

Open Access funding provided by the Qatar National Library.

References

Abulibdeh, A., 2018. Implementing congestion pricing policies in a MENA region city: analysis of the impact on travel behaviour and equity. Cities 74, 196–207. https://doi.org/10.1016/j.cities.2017.12.003.

Abulibdeh, A., 2021a. Analysis of urban heat island characteristics and mitigation strategies for eight arid and semi-arid gulf region cities. Environ. Earth Sci. 80 (7), 259. https://doi.org/10.1007/s12665-021-09540-7.

Abulibdeh, A., 2021b. Spatiotemporal analysis of water-electricity consumption in the context of the COVID-19 pandemic across six socioeconomic sectors in Doha City, Qatar. Appl. Energy 304, 117864. https://doi.org/10.1016/J.APENERGY.2021.117864.

Abulibdeh, A., 2022. Time series analysis of environmental quality in the state of Qatar. Energy Policy 168, 113089. https://doi.org/10.1016/J.ENPOL.2022.113089.

Abulibdeh, A., Andrey, J., Melnik, M., 2015a. Insights into the fairness of cordon pricing based on origin-destination data. J. Transp. Geogr. 49, 61–67. https://doi.org/10.1016/j.jtrangeo.2015.10.014.

Abulibdeh, A., Al-Awadhi, T., Al-Barwani, M., 2019a. Comparative analysis of the driving forces and spatiotemporal patterns of urbanisation in Muscat, Doha, and Dubai. Dev. Pract. 29 (5), 606–618. https://doi.org/10.1080/09614524.2019.1598335.

Abulibdeh, A.O., Zaidan, E.A., Alkaabi, K.A., 2018. Empirical analysis of the implementation of cordon pricing: Potential impacts on travel behaviour and policy implications. Transport. Res. F: Traffic Psychol. Behav. 53, 130–142. https://doi.org/10.1016/j.trf.2018.01.006.

Abulibdeh, A., Zaidan, E., 2018. Analysis of factors affecting willingness to pay for high-occupancy-toll lanes: Results from stated-preference survey of travelers. J. Transp. Geogr. 66, 91–105. https://doi.org/10.1016/j.jtrangeo.2017.11.015.

Abulibdeh, A., Zaidan, E., Abuelgasim, A., 2015b. Urban form and travel behavior as tools to assess sustainable transportation in the greater Toronto area. J. Urban Plann. Dev. 141 (3) https://doi.org/10.1061/(ASCE)UP.1943-5444.0000227.

Abulibdeh, A., Zaidan, E., 2017. Empirical analysis of the cross-cultural information searching and travel behavior of business travelers: a case study of MICE travelers to Qatar in the Middle East. Appl. Geogr. 85, 152–162. https://doi.org/10.1016/j.apgeog.2017.06.001.

AbuLibdeh, A. (2017). Traffic Congestion Pricing: Methodologies and Equity Implications. *Urban Transport Systems*. https://doi.org/10.5772/66569.

Abulibdeh, A. (2019). Water-energy nexus challenges and opportunities in Qatar. In Qatar: Political, Economic and Social Issues (pp. 209–227). nova. https://novapublishers.com/shop/qatar-political-economic-and-social-issues/.

Abulibdeh, Ammar. (2020). Planning for congestion pricing policies in the middle east: public acceptability and revenue distribution. Transportation Letters. https://doi.org/10.1080/19427867.2020.1857908.

Al-Awadhi, T., Abulibdeh, A., Al-Masri, A.N., Bin Touq, A., Al-Barawni, M., El Kenawy, A.M., 2022. Spatial and temporal changes in electricity demand regulatory during pandemic periods: The case of COVID-19 in Doha, Qatar. Energy Strategy Reviews 41, 100826. https://doi.org/10.1016/J.ESR.2022.100826.

Alhussein, S.N., 2011. Analysis of ground access modes choice King Khaled International Airport, Riyadh, Saudi Arabia. J. Transp. Geogr. 19 (6), 1361–1367. https://doi.org/10.1016/J.JTRANGEO.2011.07.007.

Balakrishnan, P., Abulibdeh, A., Abul Kasem Kabir, T., 2023. Assessment of the impact of anthropogenic evolution and natural processes on shoreline dynamics using multi-temporal satellite images and statistical analysis. Water 15 (8), 1440. https://doi.org/10.3390/W15081440.

Bergantino, A.S., Capurso, M., Hess, S., 2020. Modelling regional accessibility to airports using discrete choice models: An application to a system of regional airports. Transp. Res. A Policy Pract. 132, 855–871. https://doi.org/10.1016/J.TRA.2019.12.012.

Birolini, S., Malighetti, P., Redondi, R., Deforza, P., 2019. Access mode choice to low-cost airports: evaluation of new direct rail services at Milan-Bergamo airport. Transp. Policy 73, 113–124. https://doi.org/10.1016/J.TRANPOL.2018.10.008.

Chen, T., & He, T. (2020). xgboost: eXtreme Gradient Boosting.

Chen, X., Liu, Z., & Currie, G. (2016). Optimizing location and capacity of rail-based Park-and-Ride sites to increase public transport usage. Http://Dx.Doi.Org/10.1080/03081060.2016.1174366, 39(5), 507–526. https://doi.org/10.1080/03081060.2016.1174366.

Cherchye, L., Demuynck, T., De Rock, B., Hjertstrand, P., 2015. Revealed preference tests for weak separability: An integer programming approach. J. Econ. 186 (1), 129–141. https://doi.org/10.1016/J.JECONOM.2014.07.001.

Cirillo, C., Maness, M., & Serulle, N. U. (2014). Measuring value of travel time and travel time variability in the presence of managed lanes: results from a pilot stated preference survey on the Capital Beltway. Http://Dx.Doi.Org/10.1179/1942787513Y.0000000008, 6(1), 23–35. https://doi.org/10.1179/1942787513Y.0000000008.

Creemers, L., Cools, M., Tormans, H., Lateur, P.-J., Janssens, D., & Wets, G. (2012). Identifying the Determinants of Light Rail Mode Choice for Medium- and Long-Distance Trips: Results from a Stated Preference Study. Https://Doi.Org/10.3141/2275-04, 2275, 30–38. https://doi.org/10.3141/2275-04.

Earnhart, D., 2002. Combining revealed and stated data to examine housing decisions using discrete choice analysis. J. Urban Econ. 51 (1), 143–169. https://doi.org/10.1006/JUEC.2001.2241.

Ghofrani, A., Zaidan, E., Abulibdeh, A., 2022. Simulation and impact analysis of behavioral and socioeconomic dimensions of energy consumption. Energy 240, 122502. https://doi.org/10.1016/J.ENERGY.2021.122502.

Gokasar, I., Gunay, G., 2017. Mode choice behavior modeling of ground access to airports: a case study in Istanbul, Turkey. J. Air Transp. Manag. 59, 1–7. https://doi.org/10.1016/J.JAIRTRAMAN.2016.11.003.

Golias, J.C., 2002. Analysis of traffic corridor impacts from the introduction of the new Athens Metro system. J. Transp. Geogr. 10 (2), 91–97. https://doi.org/10.1016/S0966-6923(01)00033-3.

Gunay, G., Gokasar, I., 2021. Market segmentation analysis for airport access mode choice modeling with mixed logit. J. Air Transp. Manag. 91, 102001 https://doi.org/10.1016/J.JAIRTRAMAN.2020.102001.

Halse, A. H., Østli, V., & Killi, M. (2019). Revealed and stated preferences for reliable commuter rail in Norway. Https://Doi.Org/10.1080/19427867.2019.1586088, 12 (4), 288–292. https://doi.org/10.1080/19427867.2019.1586088.

Hasnine, M. S., Rashedi, Z., & Habib, K. N. (2019). Making use of elicited confidence ratings of the SP survey in a joint RP-SP model: a case study of commuting mode choices for TDM evaluations. Https://Doi.Org/10.1080/19427867.2019.1585687, 12(4), 274–287. https://doi.org/10.1080/19427867.2019.1585687.

Hawas, Y.E., Hassan, M.N., Abulibdeh, A., 2016. A multi-criteria approach of assessing public transport accessibility at a strategic level. J. Transp. Geogr. 57, 19–34. https://doi.org/10.1016/j.jtrangeo.2016.09.011.

Heinen, E., Harshfield, A., Panter, J., Mackett, R., Ogilvie, D., 2017. Does exposure to new transport infrastructure result in modal shifts? Patterns of change in commute mode choices in a four-year quasi-experimental cohort study. J. Transp. Health 6, 396–410. https://doi.org/10.1016/J.JTH.2017.07.009.

Hensher, D.A., Rose, J.M., 2007. Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: a case study. Transp. Res. A Policy Pract. 41 (5), 428–443. https://doi.org/10.1016/J.TRA.2006.09.006.

Ibrahim, A.N.H., Borhan, M.N., M.d. Yusoff, N.I., Ismail, A., 2020. Rail-based public transport service quality and user satisfaction – a literature review. Promet - Traffic & Transportation 32 (3), 423–435. https://doi.org/10.7307/PTT.V32I3.3270.

Jiang, Y., Yu, S., Guan, W., Gao, S., Feng, T., 2021. Ground access behaviour of air-rail passengers: a case study of Dalian ARIS. Travel Behav. Soc. 24, 152–163. https://doi.org/10.1016/J.TBS.2021.03.006.

Jou, R.C., Hensher, D.A., Hsu, T.L., 2011. Airport ground access mode choice behavior after the introduction of a new mode: a case study of Taoyuan International Airport in Taiwan. Transportation Research Part E: Logistics and Transportation Review 47 (3), 371–381. https://doi.org/10.1016/J.TRE.2010.11.008.

Larranaga, A.M., Arellana, J., Senna, L.A., 2017. Encouraging intermodality: a stated preference analysis of freight mode choice in Rio Grande do Sul. Transp. Res. A Policy Pract. 102, 202–211. https://doi.org/10.1016/J.TRA.2016.10.028.

Li, T., Guan, H., Ma, J., Zhang, G., & Liang, K. (2017). Modeling travel mode choice behavior with bounded rationality using Markov Logic Networks. Https://Doi.Org/10.1080/19427867.2017.1342945, 11(6), 303–310. https://doi.org/10.1080/19427867.2017.1342945.

Liu, S., Yao, E., Cheng, X., Zhang, Y., 2017. Evaluating the impact of new lines on entrance/exit passenger flow of adjacent existing stations in urban rail transit system. Transp. Res. Procedia 25, 2625–2638. https://doi.org/10.1016/J.TRPRO.2017.05.316.

Mohammed, S., Alkhereibi, A.H., Abulibdeh, A., Jawarneh, R.N., Balakrishnan, P., 2023. GIS-based spatiotemporal analysis for road traffic crashes; in support of sustainable transportation Planning. Transp. Res. Interdiscip. Perspect. 20, 100836.

Mamdoohi, A.R., Saffarzade, M., Taherpour, A., Yazdan Panah, M., 2012. Modeling air passengers' ground access mode choice a case study of IKIA feasibility study of congestion pricing of current tehran odd and even zone view project. Article in International Journal of Modeling and Optimization 147–152. https://doi.org/10.7763/IJMO.2012.V2.101.

Mangalathu, S., Hwang, S.-H., Jeon, J.-S., 2020. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. Eng. Struct. 219 https://doi.org/10.1016/J.ENGSTRUCT.2020.110927.

Mansour, S., Alahmadi, M., Abulibdeh, A., 2022. Spatial assessment of audience accessibility to historical monuments and museums in Qatar during the 2022 FIFA World Cup. Transp. Policy 127 (August 2021), 116–129. https://doi.org/10.1016/j.tranpol.2022.08.022.

Moeckel, R., Fussell, R., & Donnelly, R. (2014). Mode choice modeling for long-distance travel. Http://Dx.Doi.Org/10.1179/1942787514Y.0000000031, 7(1), 35–46. https://doi.org/10.1179/1942787514Y.0000000031.

Ortúzar, J. de D., & Willumsen, L. (2011). Modelling transport. https://books.google.com/books?hl=en&lr=&id=qWa5MyS4CiwC&oi=fnd&pg=PT7&ots=twYlYe0CSC&sig=-UNBchao6uHrtrVUoabd8CraGUM.

Ouda, O.K.M., Shawesh, A., Al-Olabi, T., Younes, F., Al-Waked, R., 2013. Review of domestic water conservation practices in Saudi Arabia. Appl. Water Sci. 3 (4), 689–699. https://doi.org/10.1007/s13201-013-0106-1.

Pasha, M. M., Hickman, M. D., & Prato, C. G. (2020a). Modeling Mode Choice of Air Passengers' Ground Access to Brisbane Airport: Https://Doi.Org/10.1177/0361198120949534, 2674(11), 756–767. https://doi.org/10.1177/0361198120949534.

Petrik, O., Silva, J. de A. e, & Moura, F. (2016). Stated preference surveys in transport demand modeling: disengagement of respondents. Http://Dx.Doi.Org/10.1179/1942787515Y.0000000003, 8(1), 13–25. https://doi.org/10.1179/1942787515Y.0000000003.

Ramsey, E., Berglund, E.Z., Goyal, R., 2017. The impact of demographic factors, beliefs, and social influences on residentialwater consumption and implications for non-price policies in urban India. Water (Switzerland) 9 (11), 844. https://doi.org/10.3390/w9110844.

Sohoni, A.V., Thomas, M., Rao, K.V.K., 2017. Mode shift behavior of commuters due to the introduction of new rail transit mode. Transp. Res. Procedia 25, 2603–2618. https://doi.org/10.1016/J.TRPRO.2017.05.311.

Timothy, D. (2018). Routledge handbook on tourism in the Middle East and North Africa. https://books.google.com/books?hl=en&lr=&id=XRypDwAAQBAJ&oi=fnd&pg=PT12&dq=Routledge+Handbook+on+Tourism+in+the+Middle+East+and+North+Africa&ots=wG-wqScdKK&sig=ZSX5Ozh2x8DoKNLFq_uLOR4YEhY.

Tsamboulas, D.A., Nikoleris, A., 2008. Passengers' willingness to pay for airport ground access time savings. Transp. Res. A Policy Pract. 42 (10), 1274–1282. https://doi.org/10.1016/J.TRA.2008.03.013.

Tseng, Y.Y., Knockaert, J., Verhoef, E.T., 2013. A revealed-preference study of behavioural impacts of real-time traffic information. Transport. Res. Part C: Emerg. Technologies 30, 196–209. https://doi.org/10.1016/J.TRC.2011.11.006.

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: Https://Doi.Org/10.1177/0361198118773556, 2672(47), 35–45. https://doi.org/10.1177/0361198118773556.

Yaylali, M., Çelik, A. K., & Dilek, Ö. (2016). Analyzing key socio-economic and socio-demographic drivers of domestic passengers' airline choice behavior in Turkey using multinomial and mixed logit models. Http://Dx.Doi.Org/10.1179/1942787515Y.0000000014, 8(3), 121–130. https://doi.org/10.1179/1942787515Y.0000000014.

Yazdanpanah, M., Hosseinlou, M.H., 2016. The influence of personality traits on airport public transport access mode choice: a hybrid latent class choice modeling approach. J. Air Transp. Manag. 55, 147–163. https://doi.org/10.1016/J.JAIRTRAMAN.2016.04.010.

Zaidan, E., Abulibdeh, A., 2018. Modeling ground access mode choice behavior for Hamad International Airport in the 2022 FIFA World Cup city, Doha, Qatar. J. Air Transp. Manag. 73, 32–45. https://doi.org/10.1016/j.jairtraman.2018.08.007.

Zaidan, E., Abulibdeh, A., 2021. Master planning and the evolving urban model in the gulf cities: principles, policies, and practices for the transition to sustainable urbanism. Plan. Pract. Res. 36 (2), 193–215. https://doi.org/10.1080/02697459.2020.1829278.

Zaidan, E., Abulibdeh, A., Alban, A., Jabbar, R., 2022. Motivation, preference, socioeconomic, and building features: New paradigm of analyzing electricity consumption in residential buildings. Build. Environ. 219, 109177. https://doi.org/10.1016/J.BUILDENV.2022.109177.

Zhu, P., Ho, S.N., Jiang, Y., Tan, X., 2020. Built environment, commuting behaviour and job accessibility in a rail-based dense urban context. Transp. Res. Part D: Transp. Environ. 87, 102438 https://doi.org/10.1016/J.TRD.2020.102438.

Zopluoglu, C. (2019). Detecting examinees with item preknowledge in large-scale testing using extreme gradient boosting (XGBoost): Https://Doi.Org/10.1177/0013164419839439, 79(5), 931–961. https://doi.org/10.1177/0013164419839439.