big data and cognitive computing

Article

Exploring Public Health Perspectives on Travel Behavior Using a Machine Learning Approach: Thailand Case Study

Manlika Seefong [1], Panuwat Wisutwattanasak [2], Kestsirin Theerathitichaipa [3], Pattarawadee Prasomsab [4], Nisa Dackuntod [5], Thanapong Champahom [6] and Rattanaporn Kasemsri [7,*]

[1] Faculty of Industrial Technology, Pibulsongkram Rajabhat University, Phitsanulok 65000, Thailand; manlika.s@psru.ac.th
[2] Institute of Research and Development, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; panuwat.w@g.sut.ac.th
[3] Faculty of Logistics and Digital Supply Chain, Naresuan University, Phitsanulok 65000, Thailand; kestsirint@nu.ac.th
[4] School of Transportation Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; d6700959@g.sut.ac.th
[5] Suranaree University of Technology Hospital, Nakhon Ratchasima 30000, Thailand; nisa.da@sut.ac.th
[6] Department of Management, Faculty of Business Administration, Rajamangala University of Technology Isan, Nakhon Ratchasima 30000, Thailand; thanapong.ch@rmuti.ac.th
[7] School of Civil Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand
* Correspondence: kasemsri@sut.ac.th

Abstract

Hospital transport services represent a vital alternative for addressing inequities in access to medical care, particularly in countries where public transportation systems are inadequate, such as Thailand. This approach enables equitable and widespread access to healthcare services for residents in underserved areas. The objective of this study is to analyze the factors influencing the choice of hospital transport travel mode by comparing various machine learning algorithms. The findings reveal that the categorical boosting model outperformed the other models across all performance metrics. The model results indicate that waiting time, travel time, travel cost, and comfortability significantly influence the decision to use hospital transport services. Furthermore, demographic data analysis highlights critical factors such as age, gender, income, travel frequency, occupation, and time of travel, all of which significantly affect the choice of hospital transport service. To maximize the practical implications of this study, policy recommendations and implementation strategies are proposed to support decision-makers in promoting equitable travel options and eliminating barriers to fair access to healthcare services.

Keywords: machine learning; travel behavior; travel mode choice; categorical boosting; extreme gradient boosting; Shapley additive explanations

Academic Editor: Giuseppe Maria Luigi Sarnè

Received: 14 January 2026
Revised: 22 February 2026
Accepted: 2 March 2026
Published: 5 March 2026

1. Introduction

Transportation is recognized as a critical factor influencing access to medical services, particularly in remote areas distant from healthcare facilities. High-quality healthcare institutions are typically concentrated in urban centers, which affects individuals' choice of travel modes and may lead to social exclusion [1,2]. The provision of mobile hospital services represents a vital alternative for addressing healthcare disparities, enhancing accessibility and equity in medical treatment for populations in underserved areas. Hospital transportation not only reduces transportation barriers but also influences future travel decisions to healthcare centers. Recently, they have become a viable alternative to traditional travel modes such as motorcycle taxis, public buses, and private vehicles. This emerging competition holds the potential to shift patients' travel behaviors, as hospital transportation are capable of reaching remote locations, operating 24 h a day, and is equipped with essential medical equipment. Furthermore, they can deliver various forms of emergency medical care in out-of-hospital settings prior to patient arrival at a medical facility, thereby enabling more rapid medical intervention and improving patient outcomes [3–5]. These advantages are key determinants in the shift toward alternative transportation choices among patients, especially in countries like Thailand, where public transportation systems are limited in coverage. As a result, individuals increasingly prefer travel options that offer greater convenience, speed, and safety [6].

In 2024, the Ministry of Public Health approved a draft proposal for a transportation service system aimed at facilitating access to public health services. The initiative is intended to ensure equitable access to healthcare for all individuals. To reduce spatial disparities in healthcare accessibility, health service strategists must establish precise and robust measures [2]. Understanding travelers' attitudes and behaviors is crucial for formulating effective policies and assessing future impacts [6–8]. However, prior studies examining attitudes or behavioral patterns related to mode choice in healthcare-related travel remain limited. Therefore, gaining in-depth insights into such attitudes and behaviors is essential for policymakers and relevant agencies. This understanding can help promote fairness, equality, and efficiency in healthcare service accessibility.

A review of the existing academic literature reveals the absence of empirical studies specifically examining the factors influencing hospital transportation mode choice. Only a limited number of related studies have explored factors affecting travel mode decisions for healthcare-related activities. For instance, one study investigated travel characteristics and influencing factors on the mode choice of elderly individuals in both core area and suburban areas of Fuzhou, China, focusing on travel distance, convenience, and personal attributes [1]. Another study analyzed healthcare-related travel decisions of elderly people in urban Kunming, China, considering personal attributes, family attributes, and travel attributes [9]. Additionally, a study conducted in Beijing, China examined determinants of healthcare transportation mode choice by incorporating attributes such as age, household income, land use, urban development, and parking availability [10]. Furthermore, in the context of Thailand, no studies have been found that investigate the factors influencing hospital transportation mode choice. This significant research gap highlights the novelty and contribution of the present study.

Machine learning techniques are increasingly utilized in travel behavior analysis. For instance, Abulibdeh [11] examined subway mode choice prediction. Wang and Ross [12] explored household travel surveys in the Delaware Valley region. Li and Shi [13] investigated high-speed rail mode choice. Ganmin Yin et al. [14] examined active travel behavior in Beijing, China, with the objective of predicting active travel volume and its associated probability. Chinnakrit Banyong et al. [6,15,16] investigated the impacts of High-Speed Rail (HSR) on the public transport market and evaluated the performance of machine learning techniques in predicting travel mode choice. as well as examined travel mode choice behavior in the context of Thailand's HSR development. Jenny Díaz-Ramírez et al. [17] analyzed transport mode choice preferences in a university district. And Dahmen et al. [18] developed a mode choice model based on tracking-based revealed preference data. These studies provide strong evidence that machine learning techniques often outperform traditional models in terms of accuracy and effectiveness when analyzing travel mode choice behavior. However, challenges remain in addressing non-linear relationships and handling large-scale datasets [6]. Techniques such as Categorical Boosting, Extreme Gradient Boosting, and Light Gradient Boosting Machine have demonstrated the ability to manage high-dimensional data, reduce data loss, and mitigate overfitting. Moreover, previous research has shown that these methods are highly effective in accurately predicting travel behavior [6]. However, a review of the existing literature reveals a lack of empirical studies specifically focusing on the factors influencing transportation mode choice for accessing hospital services an issue critical to promoting equity in healthcare accessibility.

Therefore, this study focuses on integrating a public health perspective with travel behavior by examining the determinants influencing individuals' decisions regarding hospital transport mode choice, an area that has received limited attention in previous research. The study conducts a comparative evaluation of various machine learning algorithms alongside a traditional statistical model to identify the most appropriate and effective modeling approach. Furthermore, the research integrates Machine Learning with Explainable AI (Shapley Additive Explanations—SHAP) to interpret the influence of key factors affecting transport mode choice. This approach provides both high predictive accuracy and strong policy interpretability. The proposed framework aims to generate reliable and meaningful insights that reflect real travel needs and the barriers individuals face in accessing healthcare services. In addition, the study seeks to provide policy recommendations and practical guidelines to support policymakers and public health agencies in promoting equity and reducing disparities in healthcare accessibility. Ultimately, the goal is to facilitate sustainable and inclusive planning and decision-making that better responds to population needs.

The rest of this paper is organized as follows. Section 2 presents the Materials and Methods, including the stated preference survey, the statistical design of the choice experiment, establishing the attributes, modeling approaches (Binary Logit, CatBoost, and XGBoost), SHAP analysis, and measures for Performance Evaluation. Section 3 presents the results and discussion. Section 4 provides policy recommendations and implementation implications. Section 5 concludes the study, and Section 6 outlines the limitations and directions for future research.

2. Materials and Methods

2.1. Stated Preference Surveys

The exploration of factors influencing the choice of hospital transportation services will be conducted in areas with varying levels of infrastructure, hospital distribution, and public transportation systems. A representative example is the Nakhon Chaiburin region, which includes the provinces of Nakhon Ratchasima, Chaiyaphum, Buriram, and Surin each exhibiting distinct economic conditions and infrastructural characteristics. In Thailand, even provinces located within the same region often show significant disparities in infrastructure development [6]. Examining areas with such diversity allows for a more nuanced understanding of travel behavior and contributes to deeper and more contextually relevant research findings.

The survey will be conducted among individuals aged 18 and above who have previously utilized transportation services to access healthcare facilities. All participants will receive comprehensive information about the objectives of the study from field surveyors to ensure accurate and complete responses. This study has obtained informed consent from all participants, all of whom voluntarily provided written consent prior to their participation. However, a total of 1200 questionnaires were distributed, proportionally allocated based on the population size in each province as follows: 468 in Nakhon Ratchasima, 204 in Chaiyaphum, 288 in Buriram, and 240 in Surin. This sample size is intended to ensure the adequacy of data for effective machine learning model analysis. According to established guidelines, the sample size should be at least 50 to 1000 times the number of prediction classes [6]. Given that this study involves two prediction classes, the proposed sample size is deemed sufficient for machine learning analysis.

2.1.1. Statistical Design of the Choice Experiment

This study evaluated two modes of transportation: public buses and hospital vehicles. The evaluation was based on four key attributes: travel time, waiting time, travel cost, and comfortability. A full factorial design would yield 243 choice sets. The number of profile combinations is $3 \times 3 \times 3 \times 3 = 81$, and the total number of choice sets is $81 \times 3 = 243$ sets. However, presenting all of these to respondents may be overwhelming. To ensure comprehensive data collection while minimizing respondent burden, an orthogonal fractional factorial design was employed. This involved generating 243 choice sets, which were subsequently grouped into 27 questionnaires, with each questionnaire containing 9 choice sets. Consequently, each respondent was required to evaluate only 9 out of the total 243 choice sets during the survey, in order to obtain comprehensive data while minimizing respondent burden and time commitment [6,19].

2.1.2. Establishing the Attributes

The travel behavior assessment in this study was based on four key attributes: travel time, waiting time, travel cost, and comfort, as presented in Table 1. These attributes are considered influential factors in individuals' decision-making regarding mode choice. The hypothetical scenarios used in the survey were based on a 50 km travel distance, which is the minimum threshold set by national health security office for providing transportation services to access public healthcare facilities [20]. As these transport services have not yet been fully implemented, empirical data on actual travel time, cost, and waiting time are currently unavailable. Therefore, the attribute values used in the study were derived from national health security office sources, including estimations published through official media and websites. Nonetheless, under the assumed scenarios, these estimations are expected to align closely with projected real-world figures.

Table 1. Typical cards from stated choice (SC) sets.

Attribute/Levels	Mode Choice	
	Bus	**Hospital Car**
Travel time (min)	70	50, 60, 70
Waiting time (min)	60	50, 60, 70
Travel cost (baht)	45	300, 400, 500
Comfort (3 level)	3, 2, 1	3, 2, 1

Note: 3 = High; 2 = Medium; 1 = Low. Hospital car refers to a hospital-operated patient transport vehicle that provides supportive facilities and basic assistance equipment for non-emergency travel.

The total travel cost for hospital vehicles was estimated based on criteria set by the national health security office and private transport service providers. According to these criteria, the cost of a 50 km journey ranges from 300 to 500 Baht [20–22]. In contrast, the travel cost for public buses was derived from actual user data, reflecting real-world expenditures [23].

The travel time for hospital vehicles was estimated based on Thailand's highway speed limit, set at 90 km/h. For a 50 km journey on a highway with minor delays, the estimated travel time is approximately 50 min [24]. In contrast, the travel time for public buses was estimated using empirical data from actual users and guidelines from the Department of Land Transport. Slight delays are expected due to scheduled stops for passenger pick-up and drop-off along the route [23].

The waiting time for hospital vehicles was estimated similarly to the travel time, assuming that users may wait for the vehicle to arrive from a 50 km distance, which takes approximately 50 min. In contrast, the waiting time for public buses was based on the official timetable provided by the Department of Land Transport, which indicates a service frequency of one trip per hour [23].

The level of comfort was determined based on reasonable assumptions drawn from real-world conditions. Comfort was categorized into three levels: Level 3 represents sitting in an air-conditioned environment; Level 2 refers to either sitting in a non-air-conditioned environment or standing in an air-conditioned environment; and Level 1 indicates standing in a non-air-conditioned environment [25].

2.2. Methods

The research process is detailed as shown in Figure 1. The process begins with data collection, including socioeconomic data, travel behavior, and decision-making situations related to travel. The data is then prepared for model training, where the dependent variable, representing the chosen travel mode, is assigned to y_train, while the independent variables, consisting of socioeconomic data and travel behavior, are assigned to x_train. Next, the data is analyzed using traditional statistical models, such as the Binary Logit Model, and machine learning algorithms, including the CatBoost Model and XGBoost, by using Python 3.12.7 via Jupyter Lab 4.x from Anaconda 2024.02. The are two part of evaluating the model and need to do it simultaneously: (1) prevent overfitting or underfitting (2) test the train model with best parameter on test set. During the hyper parameter tuning, we train our model and evaluate it with the test set. First, we need to make sure the accuracy (or loss) of on the train data and test data are close to each other, thus making our model not overfitting or underfitting. After that, we can record the model performance in the table below. Additionally, the performance of the model is evaluated using six metrics: (1) Accuracy (ACC), (2) Sensitivity (SNS), (3) Specificity (SP), (4) Precision (PRC), (5) F1-score, and (6) AUC. Finally, the key factors influencing travel mode choice are identified using SHAP analysis, making it easier to interpret and rank the importance of the factors that influence travel mode choice from the most effective model.

2.2.1. Binary Logit Model

The binary logistic model is one of the discrete choice models [6] commonly used in situations involving only two alternatives [26]. Its primary objective is to predict the probability that an individual will choose one option over the other. The model is based on the principle that decision-makers (e.g., travelers) are assumed to select the alternative that provides the highest utility. The utility of each alternative is determined by various independent variables, such as travel cost, travel time, income, and other relevant factors. The probability of an individual selecting a particular option is calculated using the logistic function, as shown below:

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

where

P_{n1} represents the probability that traveler n chooses the first alternative.

βX_{1n} and βX_{2n} denote the utility functions associated with traveler n for the first and second alternatives, respectively.

The difference in utility (ΔU) between the two alternatives is given by:

$$\Delta U = \beta X_{2n} - \beta X_{1n} = \sum (a_i - b_i) Z_i$$

where

Z_i is the ith explanatory variable,

a_i is the coefficient of the ith variable in βX_{1n},

b_i is the coefficient of the ith variable in βX_{2n}

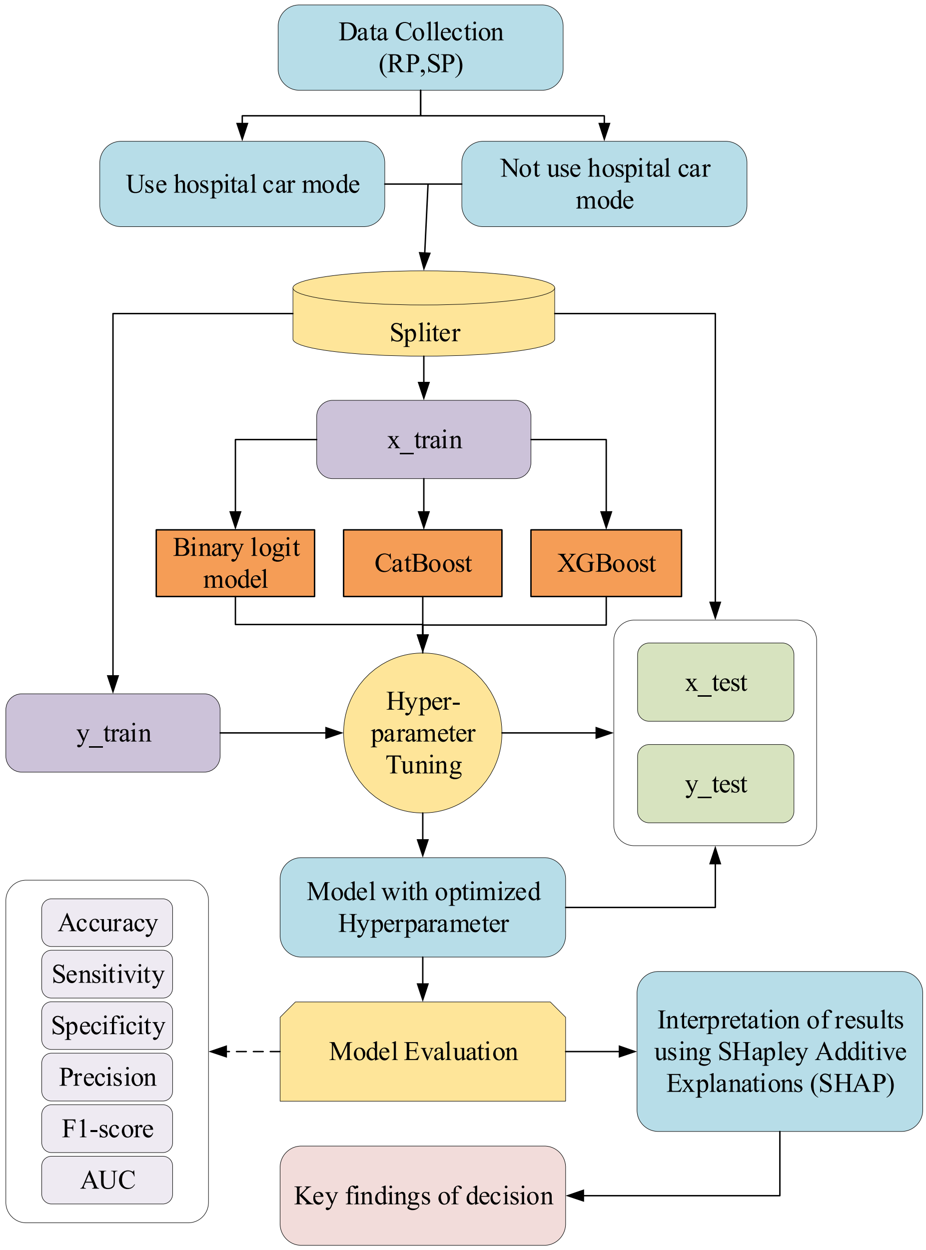

Figure 1. Model Development and Optimization Framework.

2.2.2. Categorical Boosting (CatBoost)

CatBoost is an advanced Gradient Boosting Decision Tree (GBDT) algorithm that has been enhanced to outperform traditional GBDT methods in both efficiency and predictive accuracy [15,27]. One of its key strengths lies in its effective handling of categorical features and its built-in mechanisms to prevent target leakage. It employs novel techniques to improve learning performance and mitigate the risk of overfitting [28,29]. The CatBoost algorithm initiates by assigning equal weights to all training examples, then gradually adjusts the weights to place more emphasis on misclassified instances. It leverages random permutations of the data and constructs multiple weak models to correct gradient bias. Additionally, it is well-suited for handling imbalanced datasets, with key parameters such as learning rate, tree depth, and number of iterations being tunable to optimize performance [6,30]. A distinctive feature of CatBoost is its use of target-based statistical metrics, such as mean target values per class, which enhance prediction quality. Each decision tree is built based on the residual errors of the preceding tree, progressively improving the model's accuracy. The mathematical formulation of the algorithm is presented as follows:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(X_i) \quad (2)$$

where

η is the learning rank.

2.2.3. Extreme Gradient Boosting (XGBoost)

XGBoost is a Gradient Boosting Decision Tree (GBDT) algorithm that has been optimized for efficiency and scalability [31,32]. Developed by Chen and Guestrin, it builds upon the foundational concepts introduced by Friedman [33]. A key characteristic of XGBoost is its ensemble learning approach, which combines multiple decision trees through a boosting process [34]. In this approach, each new model is trained to predict the residual errors of the previous models, thereby incrementally correcting mistakes and improving overall accuracy. XGBoost is designed to make efficient use of hardware resources, incorporating techniques such as data compression, optimized memory usage, and cache-aware computation. This allows it to process large-scale datasets, including those with billions of samples [13,31,35]. The model construction begins with a regularized learning objective, which balances model complexity and training loss to prevent overfitting [36]. Given a dataset D consisting of k instances and m features, denoted as $D = \{(X_i, Y_i)\}(\lceil D \rceil = k, X_i \epsilon R^m, Y_i \epsilon R)$ XGBoost predicts the outcome using an ensemble of N additive functions, where each function corresponds to a decision tree [32]. The overall prediction is the sum of the outputs of these individual trees, optimized through iterative learning.

$$\hat{Y}_i = \phi(X_i) = \sum_{n=1}^{N} f_n(X_i),\ f_n \in F, \quad (3)$$

where

$F = \left\{f(x) = \omega_{q(x)}\right\}\left(q : R^m \rightarrow T, \omega \epsilon R^T\right)$ represents the space of regression trees, particularly Classification and Regression Trees (CART).

N denotes the number of trees in the ensemble.

F defines the function space encompassing all possible regression trees.

q represents the tree structure, which maps an input x to a specific leaf index in the tree.

T is the number of leaves in a tree.

Each function f_n in the ensemble is defined by a specific tree structure q and a set of leaf weights ω.

$q(x)$ corresponds to the leaf index that input x falls into based on the structure of the tree.

The regularized objective function can be approximated and simplified as follows:

$$\Lambda(\theta) = \sum_i l\left(\hat{Y}_i, Y_i\right) + \sum_n \Omega(f_n) \quad (4)$$

where $\Omega(f_n) = \gamma T + \frac{1}{2}\lambda\|\omega\|^2$.

l is a differentiable convex loss function that measures the discrepancy between the predicted value $\hat{Y}_i$ and the actual target Y_i.

The second term Ω penalizes the model's complexity and consists of a function over the regression trees.

The parameters γ and λ are regularization terms used to control the model's complexity and prevent overfitting.

This additional regularization helps adjust the final learned weights to prevent overfitting. However, the optimization of the tree ensemble model cannot be performed using traditional Euclidean space optimization techniques. Instead, the model is trained in an additive manner, following the structure of the objective function as described below.

$$\Lambda^{(t)} = \sum_{i=1}^{k} l\left(Y_i \hat{Y}_i^{\,t-1} + f_t(X_i)\right) + \Omega(f_t) \tag{5}$$

The equation described adds the function f_t iteratively, which helps improve the model according to Equation (5).

2.2.4. Shapley Additive Explanations (SHAP)

SHAP is a tool designed to help interpret predictions made by complex models, such as machine learning models. Although these models may perform well during training, their internal workings can be difficult to interpret. As models become more complex, the need for effective interpretation increases [37,38]. SHAP is a combination of various interpretability methods that are linked to Shapley values. The SHAP value is a global measure of the importance of a feature for a given prediction, quantifying how much a feature contributes to increasing or decreasing the predicted value compared to the model's average prediction. This approach aids in ranking the importance of features to explain the model's behavior. Consequently, this study applies SHAP to interpret the machine learning model. Typically, the results of a SHAP analysis are visualized as a feature importance ranking plot, where the horizontal position of each feature indicates its importance and its impact on the prediction. The interpretation uses color coding to show the strength of the relationship: positive observations (red) increase the predicted value, while negative observations (blue) decrease the predicted value [6,37,39]. The SHAP value computation can be expressed through the following equation:

$$\varphi_k(\mathrm{val}) = \sum_{s \subseteq N \setminus \{i\}} \frac{|s|!(n-|s|-1)!}{n!}\left(\mathrm{val}(s \cup \{i\}) - \mathrm{val}(s)\right) \tag{6}$$

where

val represents the relevance of the feature to the algorithm's target. $\varphi_k(\mathrm{val})$ is the weighted sum of the feature's contribution to the model's target outcome, including the aggregation of features across all subsets $|s|$, where $|s|$ and $(n - |s| - 1)!$ represent the weights of $|s|$, val(s) is the expected value of $|s|$, and n is the total number of features in the dataset. The set s represents a subset of the model's features, and the letter i denotes the vector of feature values for the portion requiring interpretation.

$$g(x) = l(x_8) = \varphi_0 + \sum_{t=1}^{n} \varphi_t x_8^i \tag{7}$$

where

φ_0 is the baseline value for features that are not assigned,

n is the total number of features in the dataset. For a feature t_t,

φ_t represents the SHAP value, and X_8 is the vector of input variables that has been reduced [6,40].

2.2.5. Measures for Performance Evaluation

This study evaluates the performance of machine learning models using various metrics, including Accuracy (ACC), Sensitivity (SNS), Specificity (SP), Precision (PRC), F1-score, and AUC [41–43]. Among the evaluation criteria mentioned earlier, Accuracy is considered the overall performance measure of the classifier. Sensitivity assesses the classifier's ability to correctly identify positive labels, while Specificity measures its ability to accurately recognize negative labels. The F1-score provides insights into the relationship between the actual positive labels in the dataset and the labels predicted by the classifier. To improve the evaluation process, the AUC metric offers a robust evaluation by assessing the trade-off between true positive rate and false positive rate across different thresholds, helping to highlight the classifier's ability to minimize misclassification [44,45]. These metrics can be calculated using the following equations:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \times 100 \quad (8)$$

$$\mathrm{SNS} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (9)$$

$$\mathrm{SP} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \quad (10)$$

$$\mathrm{PRC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad (11)$$

$$\mathrm{F}_1 = 2 \times \frac{\mathrm{SNS} \times \mathrm{PRC}}{\mathrm{SNS} + \mathrm{PRC}} \quad (12)$$

$$\mathrm{AUC} = \frac{1}{2}(\mathrm{SNS} + \mathrm{SP}) \quad (13)$$

The confusion matrix, which consists of: TP (True Positive) is the true positive rate, TN (True Negative) is the true negative rate, FP (False Positive) is the false positive rate, FN (False Negative) is the false negative rate [41].

3. Result and Discussion

3.1. Descriptive Analysis

This study collected data from 1200 travelers visiting healthcare facilities using a self-administered questionnaire. This sample size was selected to ensure that it accurately represents the target population of travelers. During the survey process, the enumerators explained the study's objectives to participants to promote understanding of the research details and to obtain accurate and complete responses. Upon completion of data collection, the completeness of the dataset was assessed prior to analysis to ensure data integrity. Basic statistical measures such as frequency, percentage, mean, skewness, and standard deviation were computed to detect any irregularities in the data distribution that could potentially affect the modeling analysis. The results of the data quality assessment based on these statistical indicators are presented in Table 2. Prior to model development, Pearson's correlation coefficient was applied to examine the pairwise relationships among the 21 variables. The results suggest that the correlations among variables are not sufficiently high to indicate potential multicollinearity. The correlation heatmap (Figure 2) illustrates these relationships, where darker shades denote stronger positive correlations and lighter shades represent negative correlations. To further assess multicollinearity, the Variance Inflation Factor (VIF) was calculated for all independent variables. As presented in Table 7, the VIF values range from 1.00 to 3.28, which are well below the commonly accepted threshold of 10 reported in previous studies [46–48]. These results confirm the absence of multicollinearity, and all variables were therefore retained for inclusion in the model.

Table 2. A summary of the demographic characteristics of the respondents, based on fundamental statistical measures.

Item		Frequency	Percent	Skewness	Kurtosis
Age	(1) 18–25 years	655	19.38	0.197	−1.032
	(2) 26–40 years	1278	37.82		
	(3) 41–60 years	812	24.03		
	(4) More than 60	634	18.76		
	Total	3379	100.00		
Sex	(1) Female	1831	54.19	0.168	−1.973
	(2) Male	1548	45.81		
	Total	3379	100.00		
Income	(1) ≤10,000	1589	47.03	1.351	1.952
	(2) 10,000–20,000	1296	38.35		
	(3) 20,001–30,000	330	9.77		
	(4) 30,001–40,000	116	3.43		
	(5) ≥40,000	48	1.42		
	Total	3379	100.00		
Occupation	(1) Student	371	10.98	0.048	−1.339
	(2) Government/State enterprise officer	696	20.60		
	(3) Private company	664	19.65		
	(4) Private business	416	12.31		
	(5) Farmer	449	13.29		
	(6) Laborer	783	23.17		
	Total	3379	100.00		
Travel frequency	(1) Every month	285	8.43	−0.782	−0.615
	(2) Every 3 months	617	18.26		
	(3) Every 6 months	847	25.07		
	(4) Once a year	1630	48.24		
	Total	3379	100.00		
Time	(1) 00:00–06:00 (Night Shift)	415	12.28	0.726	1.304
	(2) 06:00–12:00 (Morning Shift)	2289	67.74		
	(3) 12:00–18:00 (Afternoon Shift)	541	16.01		
	(4) 18:00–24:00 (Evening Shift)	134	3.97		
	Total	3379	100.00		
Mode Choice	Hospital car (1)		65.10%		
	Bus (0)		34.90%		

3.2. Model Fitting and Performance

Table 3 illustrates the hyperparameter optimization of machine learning techniques using Bayesian Optimization. This process aims to reduce model overfitting by comparing algorithmic loss between training and testing datasets [31,49,50]. Table 4 compares the performance of the models, revealing that CatBoost outperformed both XGBoost and the traditional Binary Logit Model in predicting the factors influencing the choice of hospital transport travel mode. The comparative analysis demonstrates the superior predictive accuracy of the CatBoost model, achieving high performance during the testing phase with the following metrics: ACC = 0.8092, SNS = 0.9060, SP = 0.6333, PRC = 0.8178, F1-score = 0.8596, and AUC = 0.7696. These metrics reflect the model's strong capability in reliably identifying the influential factors in the selection of hospital transport travel mode [51,52].

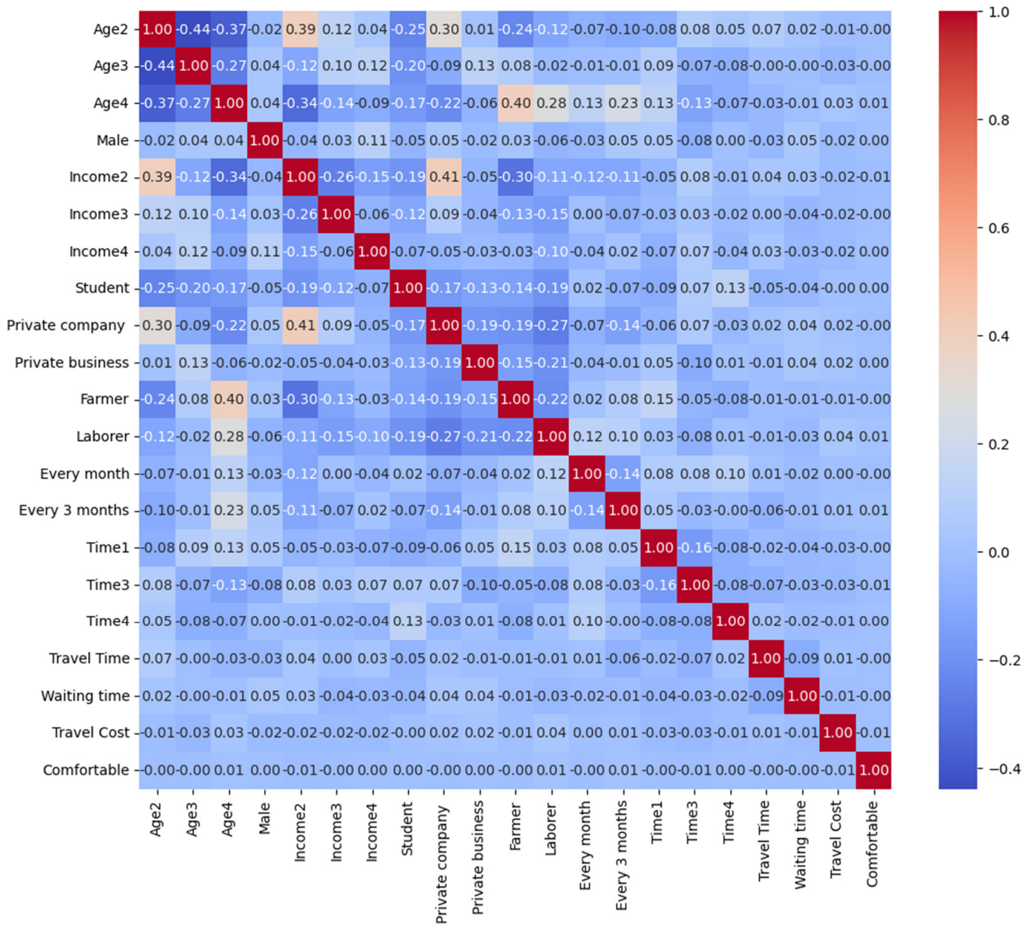

Figure 2. The correlations of all variables.

Table 3. Hyperparameter values determined using Bayesian optimization algorithm.

CatBoost		**XGBoost**	
Parameter	**Value**	**Parameter**	**Value**
depth	9.654	colsample_bytree	0.9822
iterations	230.1	learning_rate	0.2699
learning_rate	0.1309	max_depth	8.752
		n_estimators	237.9
		subsample	0.9557

Table 4. Comparison of Model Performance.

Model		**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**F1 Score**	**AUC**
Binary logit	Train	0.6674	0.6855	0.9040	0.2259	0.7797	0.6654
	Test	0.6731	0.6966	0.8818	0.2839	0.7783	0.6851
CatBoost	Train	0.9345	0.9609	0.8850	0.9401	0.9504	0.9229
	Test	0.8092	0.9060	0.6333	0.8178	0.8596	0.7696
XGBoost	Train	0.9478	0.9666	0.9127	0.9541	0.9603	0.9396
	Test	0.7855	0.8716	0.6292	0.8102	0.8398	0.7504

When considering the confusion matrices (Figure 3), the CatBoost model demonstrates clearly superior classification performance compared with the Binary Logit and XGBoost models. Furthermore, the 5-fold Cross-Validation results, shown in Table 5, indicate clear differences in predictive performance among the models. The CatBoost model exhibited the highest overall performance with Accuracy (ACC) = 0.7937 ± 0.0020, Sensitivity = 0.8623 ± 0.0146, Specificity = 0.6658 ± 0.0222, Precision = 0.8282 ± 0.0072, F1-score = 0.8448 ± 0.0034, and AUC = 0.7640 ± 0.0040.

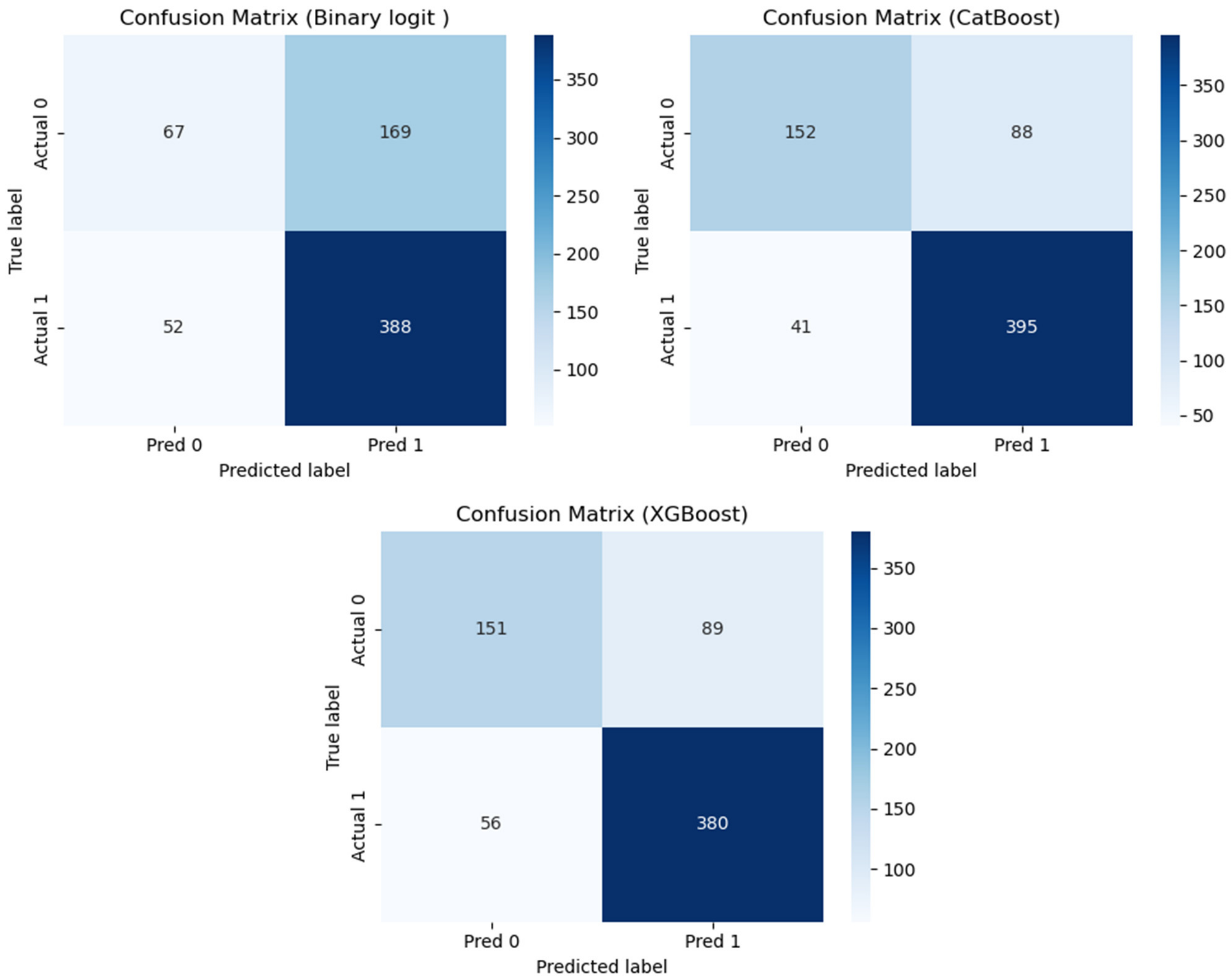

Figure 3. Confusion matrices of the three models on the test dataset.

Table 5. Cross-validation analysis of model performance.

Model	Accuracy	Sensitivity	Specificity	Precision	F1 Score	AUC
Binary logit	0.6623 ±0.0087	0.6830 ±0.0036	0.8986 ±0.0294	0.2214 ±0.0339	0.7759 ±0.0098	0.6565 ±0.0275
CatBoost	0.7937 ±0.0020	0.8623 ±0.0146	0.6658 ±0.0222	0.8282 ±0.0072	0.8448 ±0.0034	0.7640 ±0.0040
XGBoost	0.7878 ±0.0130	0.8500 ±0.0178	0.6718 ±0.0364	0.8289 ±0.0150	0.8391 ±0.0100	0.7609 ±0.0164

In addition, the class-specific performance metrics of the three models were analyzed, as presented in Table 6. The Binary Logit model exhibited a strong bias toward identifying Class 1, with a very low recall of 0.284 for Class 0. In contrast, both the XGBoost and

CatBoost models substantially improved the ability to identify Class 0 cases, achieving recall values of approximately 0.63. Among the machine learning models, CatBoost demonstrated the best overall performance, achieving the highest recall (0.906) and the highest F1-score (0.860) for Class 1. These results confirm that the machine learning models provide a more balanced classification performance for the analyzed dataset.

Table 6. Presents the per-class performance metrics.

Model	Class	Sensitivity (Recall)	Precision	F1 Score
Binary logit	Class 0	0.2839	0.5630	0.3775
	Class 1	0.8818	0.6966	0.7783
CatBoost	Class 0	0.633	0.7876	0.7021
	Class 1	0.906	0.818	0.8596
XGBoost	Class 0	0.6292	0.7295	0.7295
	Class 1	0.8716	0.8102	0.8398

Note: Class 1 = Hospital car and Class 0 = Bus.

3.3. Analysis Result of the Binary Logit Model

The results from the traditional model analysis, as presented in Table 7, indicate that passengers are increasingly likely to utilize hospital transportation services when confronted with the limitations of public transportation systems, such as extended travel times and long waiting periods. This tendency is primarily driven by the inability of public transport to provide immediate access and timely responses during emergencies [1]. Additionally, constraints related to service hours, frequency, and long walking distances to bus stops further contribute to passengers' preference for hospital transportation services. Comfort is also a key factor influencing the decision to use hospital transportation services. Passenger groups with a higher likelihood of utilizing such services include the elderly, females, high-income individuals, students, and farmers. In contrast, higher travel costs and increased travel frequency are factors that reduce the probability of choosing hospital transportation services. This may be attributed to the fact that, in developing countries such as Thailand, passengers are more cost-conscious regarding transportation. As travel frequency increases, so do the associated expenses, prompting greater financial awareness. Consequently, individuals may be deterred from using transport options that entail unnecessarily high costs [6,53]. Moreover, nighttime (12:00 a.m.–6:00 a.m.: Time 1) and late-evening (6:00 p.m.–12:00 a.m.: Time 4) travel are positively associated with the increased use of hospital transportation services, owing to their superior capacity to respond to emergency situations. In cases of severe illness or urgent medical needs, individuals are more likely to choose transportation modes that can compensate for the inadequate accessibility of regular public transport [1]. As a result, passengers may opt for hospital transportation services that are better equipped to meet their specific needs.

3.4. Factors Affecting Mode Choice in Transportation

Assessing the factors influencing travel mode choice is essential for understanding the characteristics and behaviors of passengers that affect their decision-making processes. The importance of each variable, as illustrated in Figures 4–6, is determined by calculating the mean absolute Shapley values, with features ranked according to their significance using SHAP values. Figure 4 presents the analysis results from the XGBoost model, while Figure 5 shows the results from the CatBoost model. The variable rankings from both models yielded similar results. The most influential factor in the choice of ambulance services is waiting time, followed by travel time, travel cost, and travel convenience. Other important factors include age, gender, income, travel frequency, occupation, and time of travel, respectively. Moreover, the findings derived from machine learning models align well with those obtained from the Binary Logit Model.

Table 7. Model Parameter Estimates for Travel Mode Choice.

Variable	Coefficient	Standard Error	Significance	Variance Inflation Factor (VIF)
const	0.6692	0.043	0.000 *	54.637
Travel Time	0.1677	0.043	0.000 *	1.032
Waiting time	0.1272	0.042	0.002 *	1.027
Travel Cost	−0.0886	0.042	0.037 *	1.011
Comfortable	0.1099	0.042	0.009 *	1.001
Age2	0.2777	0.072	0.000 *	3.226
Age3	0.3982	0.072	0.000 *	3.031
Age4	0.4630	0.075	0.000 *	3.277
Gen	−0.0867	0.043	0.044 *	1.044
Income2	0.1868	0.064	0.004 *	2.417
Income3	0.2698	0.058	0.000 *	1.804
Income4	0.2533	0.058	0.000 *	1.336
Student	0.1929	0.070	0.006 *	2.954
Private company	0.0496	0.056	0.372	1.731
Private business	−0.0548	0.056	0.324	1.771
Farmer	0.1495	0.068	0.029 *	2.550
Laborer	−0.0060	0.067	0.929	2.563
Every month	−0.1680	0.044	0.000 *	1.117
Every 3 months	0.0750	0.046	0.105	1.137
Time1	0.2538	0.049	0.000 *	1.093
Time3	0.0283	0.044	0.519	1.118
Time4	0.1896	0.046	0.000 *	1.067

Note: * = Statistically significant.

The comparison between models in predicting the key factors influencing mode choice in transportation indicates that the traditional Binary Logit Model has the lowest predictive performance due to several limitations, such as handling data imbalance, the complexity of the data, and managing nonlinear relationships and low data variance [6,54]. These limitations result in the traditional model's performance being inferior to machine learning models, especially when compared to CatBoost and XGBoost, which demonstrate significantly superior predictive performance. Among these, the CatBoost model outperforms all other models, offering advantages in reducing overfitting [28,29] and effectively handling imbalanced and complex data. These advantages allow for more efficient predictions without compromising the interpretability and accuracy of the overall model [6,55]. However, predicting the factors influencing mode choice is critical to understanding traveler behavior. The results from the most efficient model in this study, interpreted through the SHAP method, are shown in Figures 4–6. This method focuses on linking the responses of each data case with travel behavior patterns to provide more accurate insights [6,56].

The results from the most efficient model, CatBoost, suggest that Waiting Time has a significant influence on the choice to use hospital transport. Specifically, shorter waiting times positively impact passengers' decisions to opt for hospital vehicles. This finding aligns with previous studies that have shown that modes of transportation with shorter waiting times tend to attract more passengers [6,56]. Furthermore, previous studies have indicated that longer waiting times and extended travel durations can delay treatment and increase the risk of severe health outcomes. These factors may lead individuals to prefer faster modes of transportation [57]. This insight can contribute to improving service efficiency and providing a competitive advantage over other transportation modes, making it a key factor that drives passengers to choose the service more frequently.

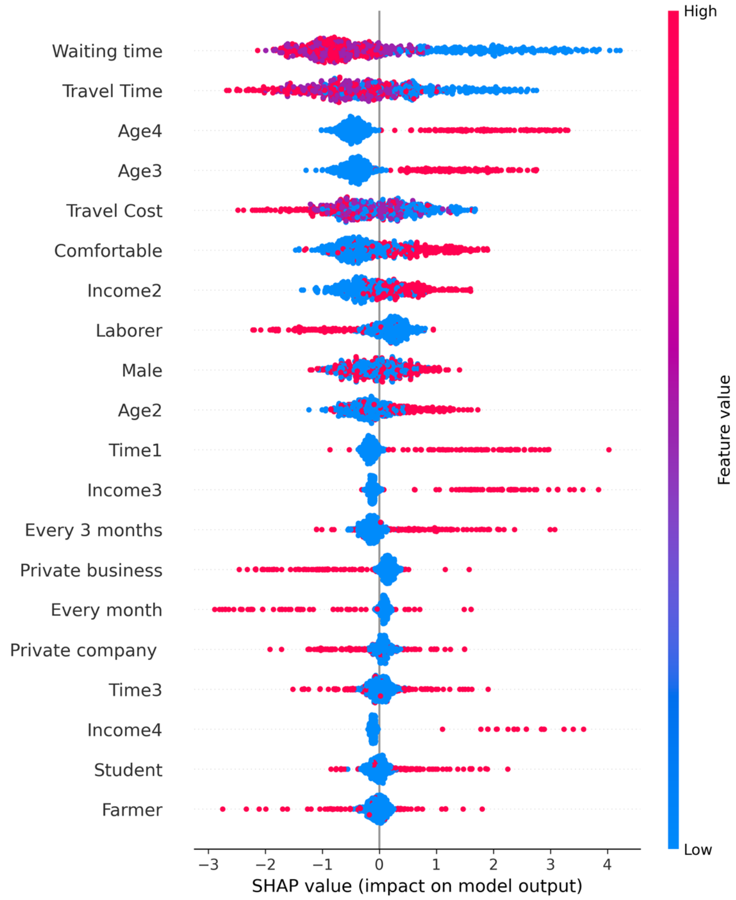

Figure 4. The factors influencing the choice to use hospital transport, utilizing SHAP values in conjunction with the XGBoost model.

Travel Time has a significant influence on the choice to use hospital transport. This means that shorter travel times increase the likelihood of passengers choosing hospital vehicles. This finding aligns with several previous studies, which indicate that passengers tend to prioritize faster modes of transportation [6,58–60]. A shorter travel time not only benefits passengers but also reduces the duration of medical care [1,9,61]. Therefore, shorter travel times represent a key competitive advantage that can motivate passengers to increasingly opt for hospital transport services.

Travel Cost has a significant influence on the choice to use hospital transport. This means that when travel costs decrease, passengers are more likely to choose hospital vehicles. This finding is consistent with previous studies, which have shown that passengers tend to prioritize travel costs; as transportation costs rise, passengers are more likely to opt for alternative, cheaper modes of transportation [62,63]. In particular, for low-income groups, transportation costs constitute a significant barrier to accessing healthcare services [64]. Given this importance, to gain a competitive advantage, setting travel costs lower than other transportation modes is a key factor in attracting more passengers to use hospital transport services.

Comfort is an important factor influencing the choice to use hospital transport, especially in countries with overcrowded public transportation systems that do not meet the demands of passengers, such as Thailand. As a result, comfort in travel becomes a key factor or competitive advantage in attracting passengers who require medical care. This finding is consistent with previous studies showing that, during major outbreaks, passengers are more likely to opt for hospital transport services rather than public transportation due to reduced comfort and increased crowding, which poses a higher risk of disease transmission [65].

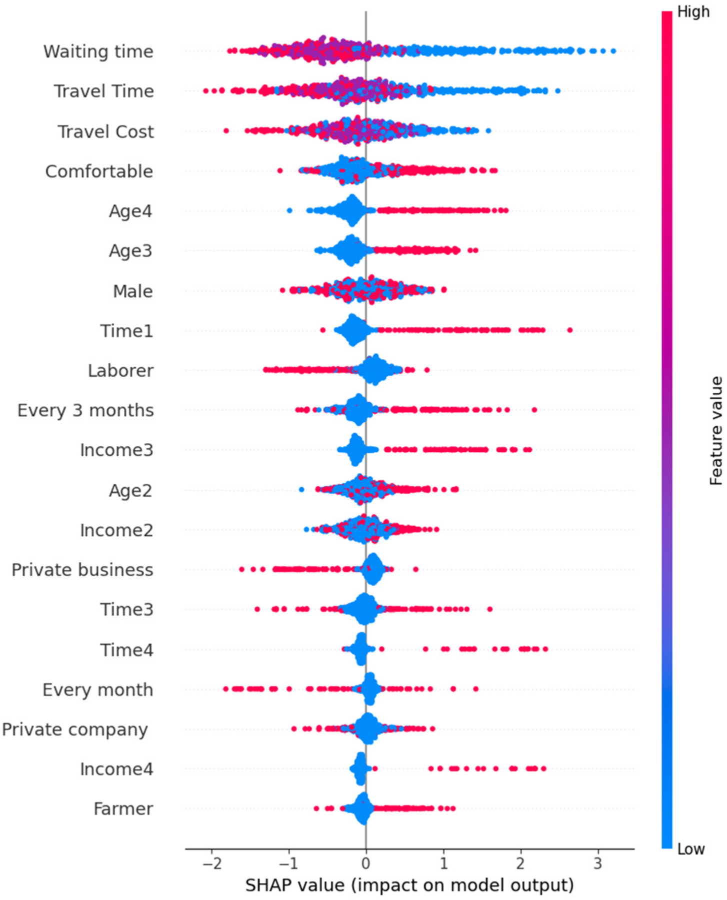

Figure 5. The factors influencing the choice to use hospital transport, utilizing SHAP values in conjunction with the CatBoost model.

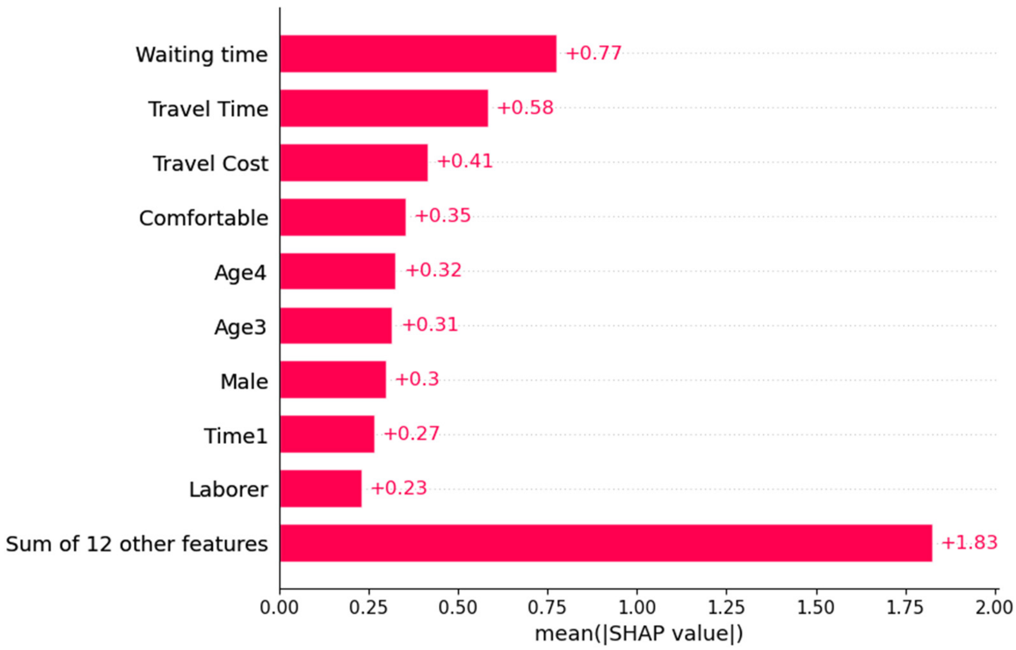

Figure 6. The importance of factors influencing the choice to use hospital transport, based on SHAP values.

Age is a factor influencing the choice to use hospital transport, with the results of the study indicating that as passengers age, they are more likely to opt for hospital transport services. This finding aligns with several previous studies that have shown that older passengers tend to choose more convenient modes of transportation rather than using buses [1,66]. It is possible that, as people age, limitations in their ability to stand and wait for long periods, as well as the difficulty of boarding and disembarking from public buses, become more challenging. These factors lead older individuals to prefer transportation modes that offer door-to-door service and minimize physical movement [67].

Gender is a factor that influences the choice to use hospital transport, with the results of the study indicating that female passengers are more likely to choose hospital transport services than male passengers. It is possible that women are more aware of health-related concerns and are more willing to seek medical assistance. Additionally, they may place greater importance on safety and comfort during travel [6]. These key factors contribute to a higher likelihood of women opting for hospital transport services.

Income has a significant impact on the choice to use hospital transport, with the results of the study indicating that passengers with lower incomes are less likely to opt for hospital transport services. This finding is consistent with previous studies that have shown that passengers with lower incomes or those who must consider travel expenses may find the higher costs of transportation that offer greater convenience and reliability prohibitive [1,6,62]. In particular, for low-income groups, transportation costs constitute a significant barrier to accessing healthcare services [64]. On the other hand, passengers with higher incomes are more likely to use hospital transport services, as supported by previous research indicating that individuals with higher incomes are generally more willing to pay for transportation options that offer greater convenience, speed, and reliability [6,68–71].

The frequency of travel has a significant influence on the choice to use hospital transport, with the results of the study showing that passengers with higher travel frequency are less likely to opt for hospital transport services. This finding is supported by previous studies indicating that the frequency of transportation service usage is significantly associated with the replacement of traditional transportation modes [65]. Passengers who travel more frequently are accustomed to traditional modes of transport and are less likely to switch to hospital transport, often choosing cheaper transportation options due to their higher tolerance for travel [6,63]. This is especially true in developing countries, where passengers are more likely to consider costs and choose lower-cost transportation modes [6,53]. In contrast, passengers who travel less frequently are more likely to use hospital transport services. This may be because their need to travel to the hospital is typically associated with emergency situations. When such events occur, they may perceive their condition as severe and in need of urgent medical attention, leading them to choose hospital transport services more frequently [72].

Occupation plays a significant role in influencing the choice to use hospital transport services. Passengers employed as farmers are more likely to choose hospital transport. This insight highlights an important demographic factor, as in rural areas, owning a private vehicle may be difficult for those with moderate to low incomes [73], and access to public transportation services is often limited, especially in countries facing public transportation system challenges like Thailand [6]. Therefore, hospital transport may serve as a more convenient option for those living in remote areas. On the other hand, passengers in other occupations, such as business owners, private sector employees, and laborers, are less likely to use hospital transport. This could be due to time constraints, as these individuals often work against the clock, and taking time off or utilizing vacation days may be difficult. Additionally, their understanding of benefits and concerns regarding travel costs may lead them to choose transportation modes that better meet their needs.

Time of travel significantly influences the choice to use hospital transport services. Passengers traveling between 12:00 a.m.–6:00 a.m. (early morning) and 6:00 p.m.–12:00 a.m. (evening) are more likely to choose hospital transport services. This finding suggests that other transportation modes, such as public transport, may have limited service hours [74]. Additionally, travel during late hours is often associated with medical emergencies. This is consistent with previous studies, which found that during emergencies or when seriously ill, passengers tend to choose transportation options that compensate for the limited access of other modes [1]. Passengers may perceive hospital transport as a more convenient, faster, and safer option compared to other modes. Conversely, passengers traveling during the daytime are less likely to choose hospital transport services. This finding confirms that, during regular hours, most passengers tend to opt for public buses, likely due to the higher frequency and availability of service during these hours, as well as the lower cost of travel.

Figure 6 illustrates the importance of each factor's attributes ranked according to the average SHAP value, where higher values indicate greater importance. The results show that the most significant factor in the choice to use hospital transport is Waiting Time, followed by Travel Time, Travel Cost, and Comfort, in that order.

4. Recommendation and Implementation

To maximize the public health impact of this study, essential policy recommendations and practical guidelines have been proposed to support relevant public health and transportation agencies. These measures are designed to promote inclusive and health-supportive transportation options, reduce transportation-related barriers, and ultimately ensure equitable, timely, and sustainable access to healthcare services, particularly for vulnerable and underserved populations. The following strategic policies and recommendations aim to guide the development of an integrated health and transportation system aligned with the principles of universal health coverage and social equity.

Travel Time Policy: Promote the development of passenger confidence in the quick, reliable, and punctual service provided for transportation. Allocate an adequate number of hospital transport in each area to reduce waiting time and promptly respond to demands. by reducing travel and waiting times, reducing travel and waiting times would encourage passengers to make greater use of hospital transportation services for healthcare.

Travel Cost Policy: Promote the establishment of fair and transparent service charges by adopting pricing based on travel distance and the type of hospital transport service provided, while ensuring standardized and affordable rates. Service fees should also be reduced for individuals who travel frequently for healthcare to encourage greater use of hospital transportation. Furthermore, shared hospital transport services for non-emergency patients should be promoted to help lower travel costs.

Comfort Policy: Promote the ergonomic design of hospital vehicles by ensuring an appropriate number of seats to facilitate passenger mobility and support the effective operation of medical personnel.

Target Group Promotion Policy

Promoting awareness and building confidence in hospital transport use among younger passengers, as this group tends to be less inclined to choose hospital transport services. Therefore, raising awareness can focus on organizing online campaigns to provide information about the advantages of using hospital transport services. Additionally, publicizing that hospital transport is available for all age groups, will help demonstrate the importance of hospital transport services to a broader audience. These approaches will encourage this group to consider using hospital transport services more frequently.

https://doi.org/10.3390/bdcc10030077

Promoting trust in hospital transport services for the elderly: Continuously promote the trustworthiness of using hospital transport services among elderly passengers, as this group tends to choose hospital transport more frequently. To enhance this trust, it is important to focus on providing special equipment tailored to the needs of the elderly. Additionally, training medical personnel to have the necessary skills in elderly care, such as proper patient transfer techniques and effective communication, is crucial. This is because the elderly often have mobility limitations, which leads them to prefer modes of transportation that are more convenient and reduce the need for physical movement [67,75].

Promoting trust in hospital transport services for women: Efforts should be made to continuously strengthen confidence in patient transport services among female users, as this group tends to utilize such services more frequently. Emphasis should be placed on enhancing service safety, particularly during nighttime operations, and promoting shared patient transport services can help address these concerns.

Promoting awareness and building trust in hospital transport services for male passengers: As this group tends to use patient transport services to hospitals less frequently, emphasizing the importance of utilizing hospital patient transportation can help shift these behaviors. Additionally, promoting positive role models [76,77] or individuals with personal experiences of using hospital transport services can serve as a powerful tool in creating a positive image and encouraging more men to use hospital transport services.

Promoting awareness and building trust in hospital transport services for low-income passengers: As this group tends to consider travel costs [1,62] and is less likely to use hospital transport services due to higher expenses. Therefore, promoting the benefits of hospital transport services should focus on government support, such as expanding benefits to cover hospital transport services for low-income individuals, making it clear that they can access hospital transport services without excessive costs. Additionally, setting reasonable and accessible distance-based service charges could serve as an effective incentive to encourage this group to increase their use of hospital transport services.

Promoting confidence in hospital transport services for high-income Passengers: A high level of service should be emphasized to meet users' expectations for convenience, speed, and reliability [6,68–71]. These approaches help build trust in and increase the attractiveness of hospital patient transport services for high-income passengers.

Promoting confidence in hospital transport services for high-frequency Travelers: Efforts to continuously build trust in hospital transport services for high-frequency travelers should focus on enhancing hygiene standards and improving safety and cleanliness within the hospital transport. Additionally, elevating the skills of personnel and staff regarding safe driving and passenger care is essential. These measures will contribute to strengthening the confidence of frequent travelers in choosing hospital transport services, ensuring their ongoing satisfaction and trust.

Promoting awareness and building trust in hospital transport services for low-frequency travelers: For low-frequency travelers, who are more accustomed to traditional modes of transportation and tend to choose cheaper alternatives, Efforts should focus on providing information about the benefits of hospital transportation services, as well as establishing appropriate pricing or discount schemes to encourage usage. Implementing these strategies can motivate infrequent travelers to consider using hospital transport services more regularly for healthcare in the future.

Promoting trust in hospital transport services for farmers: For farmers, who are more likely to use hospital transport services, ongoing efforts should focus on building trust and awareness. This can be achieved through local communication channels to highlight the benefits of hospital transport services in rural or remote areas, ensuring that these communities understand the importance of having access to such services. Government

https://doi.org/10.3390/bdcc10030077

support in covering hospital transport travel costs for those in remote or rural areas is essential. These strategies will motivate various professions, including farmers, to choose hospital transport services more consistently, while also addressing issues of equitable access to medical services in underserved areas.

Promoting trust in hospital transport services for passengers traveling between 12:00 a.m.–6:00 a.m. and 6:00 p.m.–12:00 a.m.: For passengers traveling during late-night and evening hours (12:00 a.m.–6:00 a.m. and 6:00 p.m.–12:00 a.m.), who are more likely to use hospital transport services, emphasis should be placed on building confidence in the service during these periods by ensuring that hospital transport operates with sufficient service frequency to meet demand and guarantee service readiness. In addition, offering appropriately priced promotional schemes for passengers traveling during regular hours represents another strategy to attract this group and encourage them to make greater use of hospital patient transport services.

5. Conclusions

This study conducted a comparative analysis of machine learning algorithms to investigate the factors influencing individuals' decisions regarding modes of transportation to access hospital services, an issue central to healthcare accessibility and public health equity. The analysis considered a range of influential variables, including travel time, waiting time, travel costs, comfort, and demographic characteristics. These factors are recognized as key determinants shaping individuals' healthcare-seeking behavior. Among the models compared, the CatBoost algorithm demonstrated the highest predictive performance and offered clear interpretability when applied in conjunction with the SHAP (SHapley Additive exPlanations) technique. It outperformed both the XGBoost model and the conventional Binary Logit model. The ability of CatBoost to accurately identify and explain influential factors positions it as a highly effective tool for policy planning in both the public health and transportation sectors, particularly in relation to healthcare service delivery.

The findings revealed that waiting time, travel time, travel cost, and comfort significantly influenced individuals' transportation choices for accessing hospital services. In addition, demographic factors such as age, gender, income level, travel frequency, occupation and travel time were also strongly associated with mode choice decisions.

From a public health perspective, these results highlight the existing barriers to healthcare access and underscore the need for targeted policies and interventions informed by empirical data. The ultimate goal is to promote equitable and sustainable access to healthcare, particularly for vulnerable or underserved populations. The insights derived from this study can be applied in joint planning efforts between the public health and transportation sectors to improve service quality and reduce long-term disparities in healthcare accessibility.

6. Limitations and Further Research

This study provides valuable insights; however, several limitations should be acknowledged. First, the research focuses on the Nakhon Chai Burin area, where variations in infrastructure, hospital distribution, and public transportation systems may limit the generalizability of the findings. Expanding the study area and sample size in future research would provide broader insights into travel behavior. Second, the study captures short-term behavior, which may not fully reflect long-term travel patterns. Longitudinal research is therefore recommended to better understand behavioral dynamics and changes over time. Third, the study relies on stated preference (SP) data, which may be subject to hypothetical bias, as respondents' stated choices may differ from their actual behavior. Future research should consider integrating revealed preference (RP) data to enhance external validity and better reflect real-world travel behavior. Finally, some survey responses may still contain inherent biases, as real-world behavior can differ from stated preferences, potentially leading to inconsistencies in travel mode preferences over time. This uncertainty may influence behavioral patterns and the overall conclusions of the study.

Author Contributions: Conceptualization: M.S. and R.K. Data curation: M.S., K.T., P.P. and N.D. Formal analysis: M.S. and R.K. Funding acquisition: R.K. Investigation: M.S. and R.K. Methodology: M.S. and P.W. Project administration: R.K. Supervision: R.K. Visualization: M.S. and R.K. Writing—original draft: M.S. and R.K. Writing—review and editing: M.S., T.C. and R.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by (i) Suranaree University of Technology (SUT) (ii) Thailand Science Research and Innovation (TSRI), and (iii) National Science, Research and Innovation Fund (NRSF) [Project code: 204288].

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Suranaree University of Technology (COE No.196/2567, 18 December 2024) and this study has obtained informed consent from all participants.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors express their gratitude to the Suranaree University of Technology (SUT), Thailand Science Research and Innovation (TSRI), and National Science, Research and Innovation Fund for their support in doing this research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Du, M.; Cheng, L.; Li, X.; Yang, J. Factors affecting the travel mode choice of the urban elderly in healthcare activity: Comparison between core area and suburban area. *Sustain. Cities Soc.* **2020**, *52*, 101868. [CrossRef]
2. McGrail, M.R.; Humphreys, J.S. Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. *Appl. Geogr.* **2009**, *29*, 533–541. [CrossRef]
3. Anantharaman, V.; Han, L.S. Hospital and emergency ambulance link: Using *IT* to enhance emergency pre-hospital care. *Int. J. Med. Inform.* **2001**, *61*, 147–161. [CrossRef]
4. Zhang, S.; Wang, L.; Shen, Y.; Zhang, Y.; Gao, Y.; Xu, T.; Zhang, Z. What causes Emergency Medical Services (EMS) delay? Unravel high-risk buildings using citywide ambulance trajectory data. *Habitat Int.* **2024**, *153*, 103198. [CrossRef]
5. Wiwekananda, K.; Hamukti, R.; Yogananda, K.; Egadia Calisto, K.; Utomo, P. Understanding factors of ambulance delay and crash to enhance ambulance efficiency: An integrative literature review. *J. Community Empower. Health* **2020**, *3*, 213. [CrossRef]
6. Banyong, C.; Hantanong, N.; Wisutwattanasak, P.; Champahom, T.; Theerathitichaipa, K.; Seefong, M.; Ratanavaraha, V.; Jomnonkwao, S. A machine learning comparison of transportation mode changes from high-speed railway promotion in Thailand. *Results Eng.* **2024**, *24*, 103110. [CrossRef]
7. Liu, S.; Yao, E.; Cheng, X.; Zhang, Y. Evaluating the impact of new lines on entrance/exit passenger flow of adjacent existing stations in urban rail transit system. *Transp. Res. Procedia* **2017**, *25*, 2625–2638. [CrossRef]
8. Heinen, E.; Harshfield, A.; Panter, J.; Mackett, R.; Ogilvie, D. 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* **2017**, *6*, 396–410. [CrossRef]
9. Li, X.; Zhang, Y.; Du, M. Analysis of Travel Decision-Making for Urban Elderly Healthcare Activities under Temporal and Spatial Constraints. *Sustainability* **2018**, *10*, 1560. [CrossRef]
10. Du, F.; Mao, L.; Wang, J. Determinants of travel mode choice for seeking healthcare: A comparison between elderly and non-elderly patients. *J. Transp. Geogr.* **2021**, *92*, 103023. [CrossRef]
11. Abulibdeh, A. Analysis of mode choice affects from the introduction of Doha Metro using machine learning and statistical analysis. *Transp. Res. Interdiscip. Perspect.* **2023**, *20*, 100852. [CrossRef]
12. Wang, F.; Ross, C.L. Machine Learning Travel Mode Choices: Comparing the Performance of an Extreme Gradient Boosting Model with a Multinomial Logit Model. *Transp. Res. Rec.* **2018**, *2672*, 35–45. [CrossRef]

13. Li, X.; Shi, L.; Shi, Y.; Tang, J.; Zhao, P.; Wang, Y.; Chen, J. Exploring interactive and nonlinear effects of key factors on intercity travel mode choice using XGBoost. *Appl. Geogr.* **2024**, *166*, 103264. [CrossRef]
14. Yin, G.; Huang, Z.; Fu, C.; Ren, S.; Bao, Y.; Ma, X. Examining active travel behavior through explainable machine learning: Insights from Beijing, China. *Transp. Res. Part D Transp. Environ.* **2024**, *127*, 104038. [CrossRef]
15. Banyong, C.; Hantanong, N.; Wisutwattanasak, P.; Champahom, T.; Theerathitichaipa, K.; Kasemsri, R.; Seefong, M.; Ratanavaraha, V.; Jomnonkwao, S. Analyzing High-Speed Rail's Transformative Impact on Public Transport in Thailand Using Machine Learning. *Infrastructures* **2025**, *10*, 57. [CrossRef]
16. Banyong, C.; Hantanong, N.; Nanthawong, S.; Se, C.; Wisutwattanasak, P.; Champahom, T.; Ratanavaraha, V.; Jomnonkwao, S. Machine Learning-Based Analysis of Travel Mode Preferences: Neural and Boosting Model Comparison Using Stated Preference Data from Thailand's Emerging High-Speed Rail Network. *Big Data Cogn. Comput.* **2025**, *9*, 155. [CrossRef]
17. Díaz-Ramírez, J.; Estrada-García, J.A.; Figueroa-Sayago, J. Predicting transport mode choice preferences in a university district with decision tree-based models. *City Environ. Interact.* **2023**, *20*, 100118. [CrossRef]
18. Dahmen, V.; Weikl, S.; Bogenberger, K. Interpretable Machine Learning for Mode Choice Modeling on Tracking-Based Revealed Preference Data. *Transp. Res. Rec.* **2024**, *2678*, 2075–2091. [CrossRef]
19. Hensher, D.A.; Rose, J.M.; Greene, W.H. *Applied Choice Analysis*; Cambridge University Press: Cambridge, UK, 2015.
20. Hfocus. The National Health Security Office (NHSO) has Approved Support for Local Authorities to Allocate Funds from the Local Health Security Fund (LHSF) for Providing Patient Transportation Services to Hospitals. Available online: https://www.hfocus.org/content/2024/08/31375 (accessed on 25 December 2025).
21. PORNGAROON. A Business of Genuine Care that Cannot be Easily Replaced. Available online: https://www.porngaroon.com/news/angel-care (accessed on 28 December 2025).
22. KASPAPHARMACEUTICAL. Ambulance Services for Patient Transportation. Available online: https://kaspapharma.com (accessed on 25 December 2025).
23. The Government Public Relations Department. Public transportation in Bangkok. Available online: https://www.thailand.go.th/public/index.php/visit-thailand-detail/002_011?utm_source=chatgpt.com (accessed on 23 December 2025).
24. Wisutwattanasak, P.; Jomnonkwao, S.; Se, C.; Champahom, T.; Ratanavaraha, V. Correlated random parameters model with heterogeneity in means for analysis of factors affecting the perceived value of road accidents and travel time. *Accid. Anal. Prev.* **2023**, *183*, 106992. [CrossRef] [PubMed]
25. Tanriverdi, S.C.; Shakibaei, S.; Tezcan, H.O. A Stated Preference (SP) Study on Individuals' Transportation Decisions, Focused on Marmaray Project in İstanbul. *Procedia-Soc. Behav. Sci.* **2012**, *54*, 19–26. [CrossRef]
26. Cakmakyapan, S.; Goktas, A. A comparison of binary logit and probit models with a simulation study. *J. Soc. Econ. Stat.* **2013**, *2*, 1–17.
27. Hancock, J.T.; Khoshgoftaar, T.M. CatBoost for big data: An interdisciplinary review. *J. Big Data* **2020**, *7*, 94. [CrossRef]
28. Zhai, W.; Li, C.; Fei, S.; Liu, Y.; Ding, F.; Cheng, Q.; Chen, Z. CatBoost algorithm for estimating maize above-ground biomass using unmanned aerial vehicle-based multi-source sensor data and SPAD values. *Comput. Electron. Agric.* **2023**, *214*, 108306.
29. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, QC, Canada, 2–8 December 2018; Volume 31.
30. Huang, G.; Wu, L.; Ma, X.; Zhang, W.; Fan, J.; Yu, X.; Zeng, W.; Zhou, H. Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *J. Hydrol.* **2019**, *574*, 1029–1041. [CrossRef]
31. Laphrom, W.; Se, C.; Champahom, T.; Jomnonkwao, S.; Wipulanusatd, W.; Satiennam, T.; Ratanavaraha, V. XGBoost-SHAP and unobserved heterogeneity modelling of temporal multivehicle truck-involved crash severity patterns. *Civ. Eng. J.* **2024**, *10*, 1890–1908. [CrossRef]
32. Chen, T.; Guestrin, C. XGBoost. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
33. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
34. Xu, Y.; Zhao, X.; Chen, Y.; Yang, Z. Research on a mixed gas classification algorithm based on extreme random tree. *Appl. Sci.* **2019**, *9*, 1728. [CrossRef]
35. Wang, Y.; Prato, C.G. Determinants of injury severity for truck crashes on mountain expressways in China: A case-study with a partial proportional odds model. *Saf. Sci.* **2019**, *117*, 100–107. [CrossRef]
36. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*. [CrossRef]
37. Adland, R.; Jia, H.; Lode, T.; Skontorp, J. The value of meteorological data in marine risk assessment. *Reliab. Eng. Syst. Saf.* **2021**, *209*, 107480. [CrossRef]
38. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.

39. Vega, G.M.; Aznarte José, L. Shapley additive explanations for NO2 forecasting. *Eco. Inform* **2020**, *56*, 101039. [CrossRef]
40. Shafighfard, T.; Kazemi, F.; Bagherzadeh, F.; Mieloszyk, M.; Yoo, D.Y. Chained machine learning model for predicting load capacity and ductility of steel fiber–reinforced concrete beams. *Comput.-Aided Civ. Infrastruct. Eng.* **2024**, *39*, 3573–3594. [CrossRef]
41. Seefong, M.; Wisutwattanasak, P.; Se, C.; Theerathitichaipa, K.; Jomnonkwao, S.; Champahom, T.; Ratanavaraha, V.; Kasemsri, R. Big Data Analytics with the Multivariate Adaptive Regression Splines to Analyze Key Factors Influencing Accident Severity in Industrial Zones of Thailand: A Study on Truck and Non-Truck Collisions. *Big Data Cogn. Comput.* **2023**, *7*, 156. [CrossRef]
42. Ibrahim, A.A.; Abdulaziz, R.O. Analysis of Titanic Disaster using Machine Learning Algorithms. *Eng. Lett.* **2020**, *28*, 1161–1167.
43. Provost, F. Glossary of terms special issue on applications of machine learning and the knowledge discovery process. *Mach. Learn.* **1998**, *30*, 271–274.
44. Mokhtarimousavi, S.; Anderson, J.C.; Azizinamini, A.; Hadi, M. Factors affecting injury severity in vehicle-pedestrian crashes: A day-of-week analysis using random parameter ordered response models and Artificial Neural Networks. *Int. J. Transp. Sci. Technol.* **2020**, *9*, 100–115.
45. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]
46. Seefong, M.; Wisutwattanasak, P.; Se, C.; Banyong, C.; Theerathitichaipa, K.; Jomnonkwao, S.; Champahom, T.; Ratanavaraha, V.; Kasemsri, R. Empirical Analysis of Risk Behavior in Truck Drivers Across Industrial Zones and Policy Recommendations. *Civ. Eng. J.* **2025**, *11*, 4162–4180. [CrossRef]
47. Kutner, M.H.; Nachtsheim, C.J.; Neter, J.; Li, W. *Applied Linear Statistical Models*; McGraw-hill: Columbus, OH, USA, 2005.
48. Hair, J.F., Jr.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*, 7th ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2009.
49. Ryu, S.-E.; Shin, D.-H.; Chung, K. Prediction model of dementia risk based on XGBoost using derived variable extraction and hyper parameter optimization. *IEEE Access* **2020**, *8*, 177708–177720. [CrossRef]
50. Wu, S.; Yuan, Q.; Yan, Z.; Xu, Q. Analyzing accident injury severity via an extreme gradient boosting (XGBoost) model. *J. Adv. Transp.* **2021**, *2021*, 3771640. [CrossRef]
51. Zhang, H.; Zhang, L.; Liu, Y.; Zhang, L. Understanding travel mode choice behavior: Influencing factors analysis and prediction with machine learning method. *Sustainability* **2023**, *15*, 11414. [CrossRef]
52. Ibrahim, A.A.; Ridwan, R.L.; Muhammed, M.M.; Abdulaziz, R.O.; Saheed, G.A. Comparison of the CatBoost classifier with other machine learning methods. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 738–748. [CrossRef]
53. Pei, Y.L.; Zhou, K.; Peng, T. Prediction model of passenger waiting time in high-speed rail hub based on BP neural network. In Proceedings of the Applied Mechanics and Materials, Hong Kong, China, 17–18 August 2013; pp. 1903–1906.
54. Tang, L.; Tang, C.; Fu, Q.; Ma, C. Predicting travel mode choice with a robust neural network and Shapley additive explanations analysis. *IET Intell. Transp. Syst.* **2024**, *18*, 1339–1354. [CrossRef]
55. Chen, H.; Cheng, Y. Travel Mode Choice Prediction Using Imbalanced Machine Learning. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 3795–3808. [CrossRef]
56. Zaidan, E.; Abulibdeh, A.; Alban, A.; Jabbar, R. Motivation, preference, socioeconomic, and building features: New paradigm of analyzing electricity consumption in residential buildings. *Build. Environ.* **2022**, *219*, 109177. [CrossRef]
57. Purwito, D.; Linggardini, K.; Jaitieng, A. Barriers To Healthcare Access: Examining Travel Time, Waiting Times, And Service Costs in Indonesia Primary Health Care. *J. Ber. Ilmu Keperawatan* **2025**, *18*, 236–245. [CrossRef]
58. Deng, Y.; Bai, Y.; Cui, L.; He, R. Travel Mode Choice Behavior for High-Speed Railway Stations Based on Multi-Source Data. *Transp. Res. Rec.* **2023**, *2677*, 525–540. [CrossRef]
59. Liu, L.; Zhang, M. High-speed rail impacts on travel times, accessibility, and economic productivity: A benchmarking analysis in city-cluster regions of China. *J. Transp. Geogr.* **2018**, *73*, 25–40. [CrossRef]
60. Wang, J.; Zhao, W.; Liu, C.; Huang, Z. A System Optimization Approach for Trains' Operation Plan with a Time Flexible Pricing Strategy for High-Speed Rail Corridors. *Sustainability* **2023**, *15*, 9556. [CrossRef]
61. Du, M.; Cheng, L. Better Understanding the Characteristics and Influential Factors of Different Travel Patterns in Free-Floating Bike Sharing: Evidence from Nanjing, China. *Sustainability* **2018**, *10*, 1244. [CrossRef]
62. Witchayaphong, P.; Pravinvongvuth, S.; Kanitpong, K.; Sano, K.; Horpibulsuk, S. Influential factors affecting travelers' mode choice behavior on mass transit in Bangkok, Thailand. *Sustainability* **2020**, *12*, 9522. [CrossRef]
63. Le, J.; Teng, J. Understanding Influencing Factors of Travel Mode Choice in Urban-Suburban Travel: A Case Study in Shanghai. *Urban Rail Transit* **2023**, *9*, 127–146. [CrossRef]
64. Syed, S.T.; Gerber, B.S.; Sharp, L.K. Traveling Towards Disease: Transportation Barriers to Health Care Access. *J. Community Health* **2013**, *38*, 976–993. [CrossRef]
65. Shi, K.; Shao, R.; De Vos, J.; Cheng, L.; Witlox, F. The influence of ride-hailing on travel frequency and mode choice. *Transp. Res. Part D Transp. Environ.* **2021**, *101*, 103125. [CrossRef]

66. Szeto, W.Y.; Yang, L.; Wong, R.C.P.; Li, Y.C.; Wong, S.C. Spatio-temporal travel characteristics of the elderly in an ageing society. *Travel Behav. Soc.* **2017**, *9*, 10–20. [CrossRef]
67. Schwanen, T.; Dijst, M.; Dieleman, F.M. Leisure trips of senior citizens: Determinants of modal choice. *Tijdschr. Voor Econ. En Soc. Geogr.* **2001**, *92*, 347–360. [CrossRef]
68. Karmarkar, O.D.; Jana, A.; Velaga, N.R. Willingness to pay for reduced travel time: Case study of Mumbai—Ahmedabad High Speed Railway. *Case Stud. Transp. Policy* **2023**, *12*, 100974. [CrossRef]
69. Cartenì, A.; Pariota, L.; Henke, I. Hedonic value of high-speed rail services: Quantitative analysis of the students' domestic tourist attractiveness of the main Italian cities. *Transp. Res. Part A Policy Pract.* **2017**, *100*, 348–365. [CrossRef]
70. Kim, S.; Ulfarsson, G.F. Travel Mode Choice of the Elderly: Effects of Personal, Household, Neighborhood, and Trip Characteristics. *Transp. Res. Rec.* **2004**, *1894*, 117–126. [CrossRef]
71. Ashiabor, S.; Trani, A.; Baik, H.; Hinze, N. Development of a Intercity Mode Choice Models for New Aviation Technologies. In *Aviation: A World of Growth*; American Society of Civil Engineers: Reston, VA, USA, 2007; pp. 61–77.
72. Toloo, G.; FitzGerald, G.J.; Aitken, P.J.; Ting, J.Y.S.; McKenzie, K.; Rego, J.; Enraght-Moony, E. Ambulance Use Is Associated With Higher Self-rated Illness Seriousness: User Attitudes and Perceptions. *Acad. Emerg. Med.* **2013**, *20*, 576–583. [CrossRef] [PubMed]
73. Klein, N.J.; Basu, R.; Smart, M.J. In the driver's seat: Pathways to automobile ownership for lower-income households in the United States. *Transp. Res. Interdiscip. Perspect.* **2023**, *18*, 100787. [CrossRef]
74. Rasca, S.; Saeed, N. Exploring the factors influencing the use of public transport by commuters living in networks of small cities and towns. *Travel Behav. Soc.* **2022**, *28*, 249–263. [CrossRef]
75. Zhang, W.; Cao, K.; Liu, S.; Huang, B. A multi-objective optimization approach for health-care facility location-allocation problems in highly developed cities such as Hong Kong. *Comput. Environ. Urban Syst.* **2016**, *59*, 220–230. [CrossRef]
76. Seefong, M.; Wisutwattanasak, P.; Se, C.; Theerathitichaipa, K.; Jomnonkwao, S.; Champahom, T.; Ratanavaraha, V.; Kasemsri, R. A study of motorcycle riders related to speeding behavior in Thailand's Industrial zones. *Sci. Rep.* **2024**, *14*, 29889. [CrossRef] [PubMed]
77. Ajzen, I. *Attitudes, Personality and Behaviour*; McGraw-Hill Education (UK): Maidenhead, UK, 2005.