International Journal of Multiphysics
Volume 18, No. 3, 2024
ISSN: 1750-9548

Behavioral Analysis of Urban Travel Mode Selection Based on Random Forest Algorithm

Hai Zhang[1*], Na Liu[1]
1 School of Traffic and Transportation, Xi'an Traffic Engineering Institute, Xi'an, Shaanxi, 710300, China
*Corresponding author e-mail: 1371102993@qq.com
Na Liu: 1002626540@qq.com

Abstract- Assessing the travel demand involves a knowledge of people make choices with regard to mode of transportation. It was stated that machine learning (ML) techniques are beneficial for predicting achievement and additionally suggested for modeling mode choice patterns. Nevertheless, establishing an effective rationale for the association between inputs and outputs is challenging because of the black-box structure of ML. This research examines the predictability and interpretability of the mathematical framework by analyzing travel mode selections using an innovative adaptive travel modeoptimized random forest (AWPO-RF) approach. The prediction performance of the RF is enhanced by implementing the AWPO strategy. The predictive efficacy of the suggested technique is investigated using Python and trip journal information which has been gathered. The experimental outcomes show that the suggested strategy outperformed other existing strategies for more accurate travel mode selection estimate. Moreover, the AWPO-RF technique that has been proposed to computes the relative importance of explanatory variables and their association with mode selections. This was necessary for understanding and realistic modelling of travelling behaviors.

Keywords-*Travel mode, behavior, machine learning (ML), adaptive travel modeoptimized-random forest (AWPO-RF)*

1. Introduction

The selection of urban transport was one of the important ingredients in today's life which had a significant effect on social integration, sustainability and the way people interacted. The world is affected by people's choices to move around cities [1]. The significance and intricacy that accompany making a decision means of transport in urban settings underlines diversity and an urgent call for understanding and innovative responses[2]. Residents who live in cities assess a number of factors, including simplicity, cost, the time it takes to get somewhere, and their own preferences, when deciding to move around the city on a routine basis. Urban transit mode has implications bigger than personal commuting. These affect traffic patterns, levels of pollution, and quality of life while reverberating through the fabric or urban structure for social justice [3]. The prevalence of private automobiles in many cities makes pollution in the air, congestion and emissions of carbon worse constituting serious risks to community health and sustainable development. The significant expenditures on bicycle systems, pedestrian-friendly facilities and transportation systems can reduce these unfavorable externalities and promote safer, equitable and resilient urban settings [4].

The dynamics of choosing urban transport modes are complicated by the emergence of innovative technology and changing cultural attitudes. The environment of urban transportation constitutes the emergence of cab services, the spread of electric bicycles and the impending advent of autonomous vehicles [5]. Securing modern innovations constitutes environmentally friendly and efficient urban transportation networks that demand quick policy interventions and creative urban planning techniques. The factors that impact the choice of mode are complex and context-specific they include personal preferences, economic status, habits of land usage, infrastructure for transportation, governmental initiatives and norms of culture [6].An integrated and multidisciplinary strategy addresses the numerous issues related to choosing a mode of transportation for metropolitan areas. To create and carry out successful remedies, developers, transport designers, legislators, ecologists, health care specialists, economics and community groups work together harmoniously. Furthermore, interaction between communities and citizens was necessary to promote environmentally friendly urban transportation [7]. Urban planning places more emphasis on human-focused design concepts that favor walking, bicycle facilities and accessible transportation possess unrestricted flow of automobile traffic. Improving quality of life, lively, pedestrian-friendly urban areas promote relationships, financial growth and resilience in communities [8].

The study aim is to develop a novel adaptive travel modeoptimized random forest (AWPO-RF) technique to analyze trip mode selections, this research investigates the predictability and interpretability of the mathematical framework.

2. Related works

Research [9] examined urban transportation networks using LPR and CL. The customized machine learning technique were divided into two partssuch as a unique multi-stage zero-shot classifier and an operational multi-grained inspecting collective learning system. By combining the distinct advantages of LPR information, retrieved spatiotemporal transport features, ranging and filtering CL information, the former seeks to predict the volume of traffic constituted in a single link.The outcome of the experiment exhibits the quantity of traffic estimated by using information collected from many sources. The GPS information was utilized to determine travel modes using machine-learning categorization approach. The infer trip phases of the GPS information utilizing an approach that consists of two phases. The initial phase to identify transit types. The additional modes of transportation are determined in the second phase through Gaussian procedure classification [10]. The experimental outcome demonstrated that the suggested strategy was designed for assigning modes of transportation by GPS.

Study [11] examined the streamlined operation of green way compatibility assessment was achieved by applying machine learning methods and GIS resources in combination with a range of freshly collected urban areas information, such as street-level visuals, PoIs and LBS geolocation information. The experimental outcome demonstrated the feasible and rehabilitate green paths by metropolitan systems. Article [12] examined the modes of transport used by travelers from their GPS itineraries. A substantial number of annotated GPS itineraries were unutilized while tagging work was carried out by simulations in a supervised manner. Consequently, a deep SECA design was suggested to autonomously retrieve relevant characteristics from GPS intervals. The result shows that compared with alternative approaches, the suggested strategy was superior.

Research [13] examined to forecast the travel durations on metropolitan systems were partially detected through mobile sensors. The machine learning algorithms used for estimating journey durations on urban areas are partially captured by mobile instruments: such as MFFN and RF systems. The experimental findings demonstrated that the suggested RF and MFFN approaches provide superior prediction efficiency. Article [14] examined a unique unconstrained additive method of learning for road traffic jam identification and targeting, interactively across time, addressing two major issues in transport assessment. Such as hyper-dimensional processing and the IKASL method gradually discover a long time for transport assessment. The anticipated time for travel of subsequent location forecasting issues constitute the travel path of a single vehicle throughout the metropolitan area can be predicted by the vehicle's subsequent destination and its arrival period [15]. The LSTM neural systems were the foundation of deep learning algorithms utilized for long travel paths. The result compared with contrasting approaches, the suggested LSTM approach was superior.

Historically, models have used statistical regression frameworks, such as nested logit models, multinomial logit models, linear regression models, and Poisson regression models, to estimate the choice of travel mode. These models, however, are predicated on certain underlying relationships between the explanatory and dependent variables, and therefore have their own set of assumptions. The multinomial logit model, for instance, makes the assumption that each pair of alternatives' choice probabilities are independent of the existence or attributes of every other alternative. When these presumptions are broken, biassed predictions and inconsistent parameter estimates result. The inability to assess the relative effects of explanatory factors on travel mode choices is another serious flaw in statistical regression models. Travel mode choice prediction might be greatly aided by knowing the relative relevance of the explanatory variables, which would enhance travel demand forecasting. For traditional statistical regression models, the significance test or sensitivity analysis can be performed, but only one variable is assessed at a time with the presumption that other variables stay constant. Consequently, it is possible to overlook the significant interactions between factors [16].

Machine learning techniques offer a viable substitute for statistical models when modelling the selection of transport modes. Machine learning techniques learn to describe complicated relationships in a data-driven manner, as opposed to making tight assumptions. Transportation research has shown the value of machine learning techniques such as support vector machines, decision trees, and neural networks for forecasting travel mode preferences. These machine-learning techniques often include selecting the top model and using its estimated parameters to forecast results in various scenarios. It is debatable, though, whether creating and utilising a single model is always the optimal course of action given the several sources of error and uncertainty in the study of transport mode choices. Errors may exist in the input data, the sample may be biassed, the model may be stochastic, and the prediction scenarios may not match the real evolution of transportation systems.

This study investigates the possibilities of a so-called ensemble method to predict travel mode choices in order to address this issue. When using many learning algorithms, ensemble approaches in machine learning yield greater prediction performance than using only one of the individual learning algorithms. The random forest (RF) approach, created by Breiman, is the most well-liked ensemble method and has excellent prediction and classification performance. To maximise prediction performance, the RF deliberately combines several straightforward decision trees rather than fitting a single "best" tree model. The study accommodate for variations in travel decision heuristics by applying the random forest approach as a plurality of decision trees when it comes to travel mode selections [17]. Various decision trees within the ensemble might identify distinct sources of variability and uncertainty in the data. Therefore, it would be predicted that the accuracy of model estimate and prediction would improve from a purely technical standpoint. The random forest method allows for the identification and interpretation of pertinent factors and interactions by utilising methods and insights from both statistical and machine learning approaches. We can better comprehend model outcomes thanks to this method's interpretability, which is also crucial for analysing the connections between mode choice and its contributing components. The random forest approach has been widely used and successfully applied to many different research topics. This study uses the RF approach to address classification and prediction problems related to transportation. Travel choice behaviour, traffic incident prediction, traffic time/flow prediction, and pattern identification are the four general categories into which they are divided.

3. Methodology

Figure(1) depicts the proposed methodology. We propose a novel adaptive travel modeoptimized random forest (AWPO-RF) approach examines the predictability and interpretability of the mathematical framework by analyzing travel mode selections.

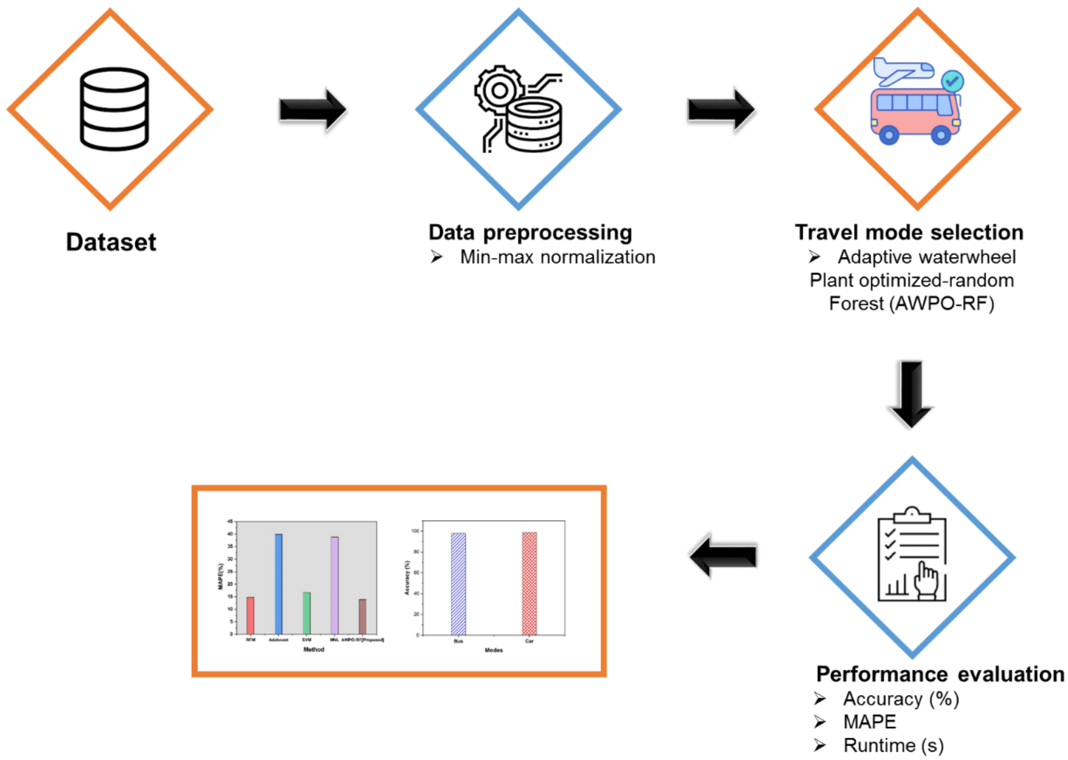

Figure (1): Proposed Methodology

3.1 Dataset

Initially, we obtained a dataset from github, (https://github.com/nekketsu2010/sussex-huawei-locomotion-challenge2023?tab=readme-ov-file). The SHL dataset (SHL) are used to assess the AWPO-RF method. Three individuals gathered the SHL dataset. There are four different kinds of transport modes are consisted. Every sample has data from a magnetometer, gyroscopic and speedometer. SHL datasets constitute an average sample rate of 100 Hz.

3.2 Data Pre-Processing

3.2.1 Min-max normalization

Min Max Normalization offers consistent analysis across several data sources, which enhances data normalization. Operations may improve their cybersecurity architecture with this strategy, minimizing risks and protecting networks from erratic cyberattacks. In order to mitigate the significant disparities in data values resulting from dimension differences, we propose the Min-Max Normalization technique, which may be expressed as follows.

$$Y = \frac{y-min}{max-min} \qquad (1)$$

The inscriptions Min and Max, respectively, stand for the maximum and minimum values of each dimension. The precision and speed of the model's convergence can be increased by using the Min-Max Normalization to map data between 0 and 1 without affecting the linear relationship between the original data.

3.3 Adaptive Wheel Plant Optimization (AWPO)

AWPO promotes environmentally friendly urban transportation options through smart grid integration and renewable energy production. The suggested AWPO is a population-centered strategy that, depending on people's capacity to navigate through the universe of potential issue solutions, iteratively provides an acceptable response. Each travel mode make up the AWPO community has a different value for each issue variable depending in the search region. Consequently, every waterwheel symbolizes a travel mode choice that can be expressed mathematically as a matrix.

All of the travel modes in the WWPA community are expressed in equation (2). At the beginning of the AWPO execution, the locations of the travel choice in the process of searching space are created randomly by using equation (3).

$$O=\begin{bmatrix} O_1 \\ \vdots \\ O_j \\ \vdots \\ O_M \end{bmatrix}=\begin{bmatrix} o_{1,1} & \cdots & o_{1,i} & \cdots & o_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ o_{j,1} & \cdots & o_{j,i} & \cdots & o_{j,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ o_{M,1} & \cdots & o_{M,i} & \cdots & o_{M,n} \end{bmatrix} \quad (2)$$

$$o_{j,i}=ka_i+q_{j,i}\cdot(va_i-ka_i), \quad j=1,2,\dots,M, \quad i=1,2,\dots,n \quad (3)$$

Where M and n stand for the number of travel model factors, respectively; o_j is the jth travel mode position, $q_{j,i}$represent an arbitrary value in the interval [0,1]; $ka_i$and va_i represent the upper and lower limits of the ith issue factor;O is the population-based vector of positions; and $o_{j,i}$constitute $j\ th$ dimensions ranging.

The target function may be determined for every travel choice as they symbolize a possible fix for the issue illustrated in equation (4)

$$E=\begin{bmatrix} E_1 \\ \vdots \\ E_J \\ \vdots \\ E_M \end{bmatrix}=\begin{bmatrix} E(W_1) \\ \vdots \\ E(W_J) \\ \vdots \\ E(W_M) \end{bmatrix} \quad (4)$$

Here E_J is the estimated value for jth and$E$is an array containing all parameters. The primary measurements used to choose the most suitable options are the desired functional assessments. Consequently, the greatest value of the desired function indicates the most suitable potential solution, while the lowest value indicates the least favorable potential solution. The optimal solution must change over time as a result of the travel mode of choice' varied throughout the search area during each iteration.

By simulating the travel mode's proximity to the bug can ascertain the travel behaviour shift in position.

$$X\rightarrow=q_1\rightarrow\cdot(\omega(e)\cdot O(s)+2L \quad (5)$$

$$O(s+1)=O(s)+X\rightarrow.(2L+q_2) \quad (6)$$

The desired travel mode of choice's location can be adjusted with the help of the subsequent equation:

$$O(s+1)=Gaussian\ (\mu o,\sigma)+q_1(O(s)+2LX\rightarrow) \quad (7)$$

Here the independent variables q_1 and q_2 possess the intervals [0, 2] and [0,1] respectively. Additionally, the travel modeuses X→an array that specifies the circumference of wrap for potential places and L represents a quadratic integer with quantities from [0,1].

The predicted travel mode choice behavior serves as the basis for the community upgrade in AWPO. The simulation's capability of moving insect to the proper tube, which causes slight adjustments to the location of the travel choice in the area of search. AWPO abuse authority was boosted during the local seek are converged to replicate the travel mode' natural behavior, AWPO designers first choose a random site for every travel mode in the community appropriate of location for consuming bugs. The waterwheel was shifted to the new location.

$$X\rightarrow=q_3\rightarrow\cdot(\omega(e)\cdot LO_{best}(s)+q_3O(s)) \quad (8)$$

$$O(s+1)=S(s)+LX\rightarrow \quad (9)$$

The variables $q_3$and $O(s)$represent the O_{best}solutions, respectively, at iteration $s$and [0,2], respectively.

The subsequent modification was performed to ensure that local minima are avoided if the approach fails to improve the iterations:

$$O(s+1)=(q_1+L)\sin(ED\theta) \quad (10)$$

The random factors E and Dconstitute the interval among [-5,5]. Furthermore, the following formula shows the quantity of L falls significantly:

International Journal of Multiphysics
Volume 18, No. 3, 2024
ISSN: 1750-9548

$$L = \left(1 + \frac{2s^2}{S_{max}} + E\right) \tag{11}$$

We employ an evolutionary variable$e$that dynamically alters depending on the pursuit state, to incorporate the inertia weight ω:

$$\omega(e) = \frac{1}{1+1,5f^{-2.6e}} \quad with \quad \omega(e) \in [0.4,0.9] \ \forall e \in [0,1] \tag{12}$$

Exploitation and extraction are dynamically balanced by using the ineffective weight in the subsequent equations (5) and (8).

The AWPO algorithm's ability to locate optimum solutions was enhanced by the adaptive inertial weight mechanism, which allows the system to adapt its hunting setting and strike an improved equilibrium between exploration and extraction.

3.4 Random Forest (RF)

The Random Forest method can forecast and evaluate variations in urban areas travel, including bus, car, railway and subway. The approach may enhance the development of infrastructure, transport networks and urban development for greater efficiency and environmental sustainability. The machine learning method termed as Random Forest was utilized to sort vast volumes of information into categories. To achieve a high degree of precision, Random Forest combines several trees of data used for training. By using a randomized choice of features techniques and dynamic pooling, random Forest was utilized for the improvement of the CART approach. The following Breiman and Cutler's approach for the Random Forest method:

Select a size-n randomly selected sample from information clusters was recovered. They termed the phase as bootstrapping. The tree was developed by using a bootstrapping instance until it reaches its largest dimension before being pruned. In order to create a tree, a randomized choice of features was used. Specifically, m explicating factors are randomly picked where $n << p$, and \$ n \$informative variables are used to determine the sorter.Where k trees are found in the canopy and procedures 1 and 2 are repeated. The victor of randomly selected forest classification was determined by tallying the votes cast in each tree; the tree with the highest number of points wins.

As per Yin, Random Forest creation employs a specific method to ascertain the division that will function as a single node based on the index of Gini significance:

$$Gini(T) = 1 \cdot \sum_{j=1}^{l} oj^2 \tag{13}$$

The possibility of T belongs to group j was represented by oj.

The Gini value has been determined, by using the following formula to determine the Gini Gain values:

$$GiniGain\ (T) = Gini(T) - Gini(B,T) = Gini(T) \cdot \sum_{j=1}^{m} \frac{|T_j|}{|T|} Gini(T_j) \tag{14}$$

Here $T_j$represents T division carried through feature B.

3.5 Adaptive Waterwheel Plan Optimized-Random Forest (AWPO-RF)

The innovative hybrid strategy to improve choice-making in complicated urban contexts integrates Random Forest (RF) with Adaptive Travel modeOptimization (AWPO) for urban transport mode selection. AWPO constitutes the kinetics of a travel mode of choice system and uses reactive learning and incremental upgrades to optimize resource and attribute allocation, effectively capturing the fluctuating and nonlinear aspects of urban traffic behavior. The hybrid model combines the predictive ability of RF with the AWPO by integrating the Random Forest approach, possess a strong collective method for learning. Urban transportation systems become more resilient and adaptable to fluctuating customer needs and situations. Therefore, maximizing transportation effectiveness, lowering traffic and encouraging environmentally friendly travel alternatives, the hybrid method can greatly aid in the development of urban travel.

4. Experimental Results

From January 20, 2024, to December 11, 2024, researchers used an offline questionnaire to survey residents of three Shanghai urban clusters along their main roadways.

Using Tensorflow 1.12.0 was used to complete the recommended task where Python software was installed for the procedure to be completed. We assess the proposed strategy and calculate its effectiveness using the following indicators: Runtime(s) and MAPE, Accuracy (%). We also present an efficacy comparison between our proposed strategy and other current approaches. The existing methods include RFM, AdaBoost, SVM and MNL [16], DNN, LSTM and CL-TRANSMODE, Bus, Car, railway and subway [17].

Every traveler is limited by their own logic. Few passengers take random selection variables, represented by mutation factors in the model, into account while deciding on a travel mode; the majority, however, adhere to the idea of greatest utility. So, tourists will likely switch up their mode of transportation regardless of the maximization of utility concept. Figure (2a) shows the comparison of accuracy between Bus and Car. Figure (2b) shows the comparison of accuracy between Railway and Subway. The suggested AWPO-RF method compared with various modes such as bus, car, railway and subway. The AWPO-RF method exhibit the accuracy of modes.

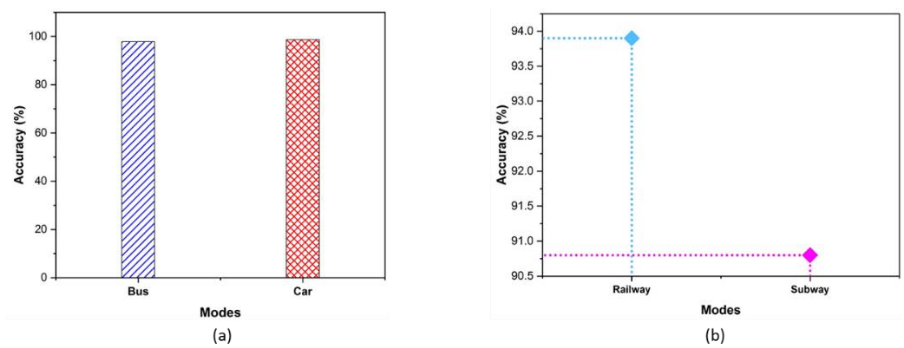

Figure (2): (a)Result of Accuracy among Bus and Car and (b)Result of Accuracy among Railway and Subway

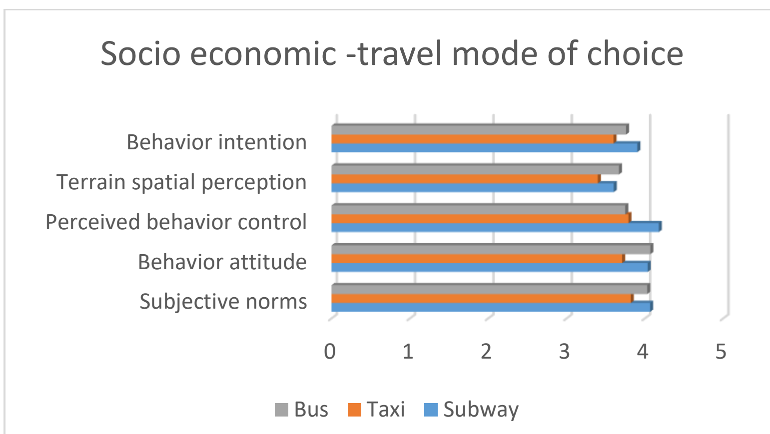

Figure 3- socio economic-travel mode of choice

Subway riders tend to fall into one of five categories: women, people of middle age and younger, those living in the area, people with moderate to low incomes, and those who do not own a car. The greater physical difficulties of driving in the complicated urban topography of a river valley type metropolis may explain why women are less likely to go by automobile than males. The subway is not as popular among commuters over the age of 50 as cars and buses. This might be because taking the subway is more difficult and less accessible than the other two choices. Subway riders of

International Journal of Multiphysics
Volume 18, No. 3, 2024
ISSN: 1750-9548

a certain age may, therefore, face challenging terrain on their way out of the station in this river valley metropolis. People who have more disposable income and a car tend to take cabs more often, which is consistent with research on commuting behavior in plain cities. This suggests that, because of the relatively flat topographical impression of river valley-type cities, higher-income commuters with more private cars are not easily convinced to take the subway (refer Figure 3). Commuters in the area tend to favor taking the metro. Possible explanations include a higher rate of private vehicle ownership and familiarity with the intricate topography of the valley-type city among locals.

The MAPE between the Projected Outcomes and the factual results exhibits how MAPE evaluates a model's accuracy. The comparative evaluation of MAPE was shown in Figure (4). When compared to presently existing methodologies, the suggested AWPO-RFhas a MAPE value of 13.95. The proposed methodology demonstrates superiority over the existing methods for travel mode selections.

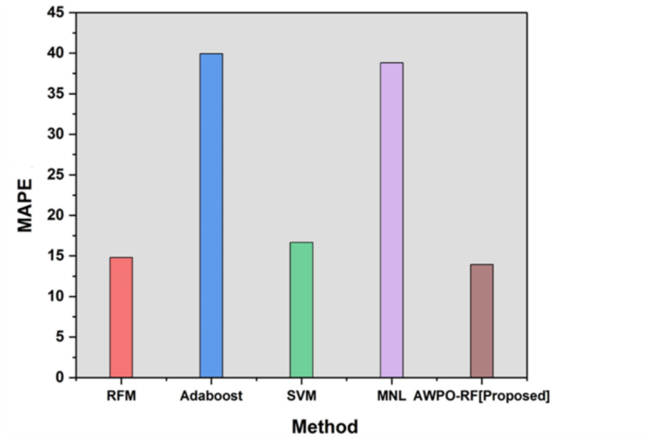

Figure (4): Result of MAPE

Runtime usually expressed in seconds, the amount of time that an application or process executes. It shows how long it takes for a certain action or activity to be completed inside an application or system initiative. The comparative evaluation of Runtime was shown in Figure (5). When compared to presently existing the methodologies, the suggested AWPO-RF has a runtime value of 9.86s. Our proposed method provided superior results for travel mode selections.

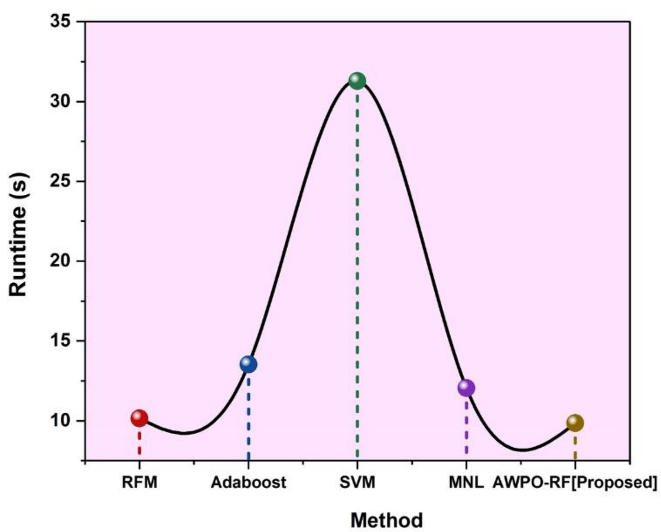

Figure (5): Result of Runtime

The comparison of real or anticipated value and accuracy describes how precise or accurate a measurement, computation forecast. Figure (6) shows the comparative evaluation of accuracy between the proposed and traditional methods. In contrast to various methods the suggested AWPO-RF achieves an accuracy of 98.4%. Our suggested approach optimizes travel mode selection. Table 1 shows that the numerical outcomes of parameters.

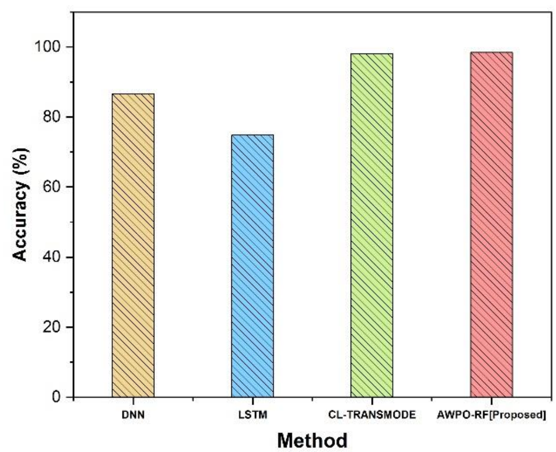

Figure (6): Result of Accuracy

Table (1): Result parameters

MAPE					
Methods	**RFM**	**AdaBoost**	**SVM**	**MNL**	**AWPO-RF [proposed]**
MAPE (%)	14.81	39.93	16.66	98.82	13.95
Runtime (s)					
Methods	**RFM**	**AdaBoost**	**SVM**	**MNL**	**AWPO-RF [proposed]**
Runtime (s)	10.15	13.52	31.28	12.05	9.86
Accuracy (%)					
Methods	**DNN**	**LSTM**	**CL-TRANSMODE**	**AWPO-RF [proposed]**	
Accuracy (%)	86.6	74.9	98.1	98.4	

5. Conclusion

In this work, we proposed a revolutionary strategy called the adaptive waterwheel plant optimized random forest (AWPO-RF) approach for travel mode selections. Initially, we obtained a dataset from github, to train our suggested model. The SHL dataset (SHL) is used to assess the AWPO-RF method. The prediction performance of the RF is further enhanced by implementing the AWPO strategy. The experimental results showed MAPE (13.95), runtime (9.86s) and accuracy(98.4%).When the assessment results are compared to the previously used approaches, the suggested AWPO-RF approach calculates the relative significance of explanatory factors and their correlation with mode selections. The disparity in accessibility of transportation among urban populations gives rise to issues of equality, particularly for underprivileged communities where access to specific modes of transport can be restricted.

The choice of travel mode was impacted by intricate behavioral dynamics that can be difficult to predict. The dynamics include economic status, cultural conventions and individual preferences. In future research, ensuring equitable access to secure, cheap and dependable transportation alternatives for all residents requires addressing equity concerns in transport design.

Acknowledgement
2024 Key research project of Xi'an Traffic Engineering Institute, project number:2024KY-31

References

1. Essien, A., Petrounias, I., Sampaio, P., &Sampaio, S. (2019, February). Improving urban traffic speed prediction using data source fusion and deep learning. In *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 1-8). IEEE.10.1109/BIGCOMP.2019.8679231
2. Chen, W., Wu, A. N., &Biljecki, F. (2021). Classification of urban morphology with deep learning: Application on urban vitality. *Computers, Environment and Urban Systems*, *90*, 101706. 10.1016/j.compenvurbsys.2021.101706
3. Bachir, D., Khodabandelou, G., Gauthier, V., El Yacoubi, M., &Puchinger, J. (2019). Inferring dynamic origin-destination flows by transport mode using mobile phone data. *Transportation Research Part C: Emerging Technologies*, *101*, 254-275. 10.1016/j.trc.2019.02.013
4. Zhou, X., Wang, M., & Li, D. (2019). Bike-sharing or taxi? Modeling the choices of travel mode in Chicago using machine learning. *Journal of transport geography*, *79*, 102479.
5. Fani, A., &Naseri, H. (2021). Travel Time Modelling of Urban Roads By Application of Coyote Optimization-based Machine Learning Method. *Amirkabir Journal of Civil Engineering*, *53*(9), 3649-3664.
6. Novak, H., Bronić, F., Kolak, A., &Lešić, V. (2023). Data-driven modeling of urban traffic travel times for short-and long-term forecasting. *Ieee transactions on intelligent transportation systems*.10.1109/TITS.2023.3287980
7. Cheng, Z., Wang, W., Lu, J., & Xing, X. (2020). Classifying the traffic state of urban expressways: A machine-learning approach. *Transportation Research Part A: Policy and Practice*, *137*, 411-428. 10.1016/j.tra.2018.10.035
8. Le, J., &Teng, J. (2023). Understanding influencing factors of travel mode choice in urban-suburban travel: a case study in Shanghai. *Urban Rail Transit*, *9*(2), 127-146.10.1007/s40864-023-00190-5
9. Liu, Z., Liu, Y., Meng, Q., & Cheng, Q. (2019). A tailored machine learning approach for urban transport network flow estimation. *Transportation Research Part C: Emerging Technologies*, *108*, 130-150.10.1016/j.trc.2019.09.006
10. Xiao, G., Cheng, Q., & Zhang, C. (2019). Detecting travel modes using rule-based classification system and Gaussian process classifier. *IEEE Access*, *7*, 116741-116752.10.1109/ACCESS.2019.2936443
11. Tang, Z., Ye, Y., Jiang, Z., Fu, C., Huang, R., & Yao, D. (2020). A data-informed analytical approach to human-scale greenway planning: Integrating multi-sourced urban data with machine learning algorithms. *Urban Forestry & Urban Greening*, *56*, 126871.10.1016/j.ufug.2020.126871
12. Dabiri, S., Lu, C. T., Heaslip, K., & Reddy, C. K. (2019). Semi-supervised deep learning approach for transportation mode identification using GPS trajectory data. *IEEE Transactions on Knowledge and Data Engineering*, *32*(5), 1010-1023.10.1109/TKDE.2019.2896985
13. Alrukaibi, F., Alsaleh, R., & Sayed, T. (2019). Applying machine learning and statistical approaches for travel time estimation in partial network coverage. *Sustainability*, *11*(14), 3822. 10.3390/su11143822
14. Bandaragoda, T., De Silva, D., Kleyko, D., Osipov, E., Wiklund, U., &Alahakoon, D. (2019, October). Trajectory clustering of road traffic in urban environments using incremental machine learning in combination with hyperdimensional computing. In *2019 IEEE intelligent transportation systems conference (ITSC)* (pp. 1664-1670). IEEE.10.1109/ITSC.2019.8917320

15. Sun, J., & Kim, J. (2021). Joint prediction of next location and travel time from urban vehicle trajectories using long short-term memory neural networks. *Transportation Research Part C: Emerging Technologies*, *128*, 103114.10.1016/j.trc.2021.103114
16. Cheng, L., Chen, X., De Vos, J., Lai, X., &Witlox, F. (2019). Applying a random forest method approach to model travel mode choice behavior. *Travel behaviour and society*, *14*, 1-10.10.1016/j.tbs.2018.09.002
17. Qin, Y., Luo, H., Zhao, F., Wang, C., Wang, J., & Zhang, Y. (2019). Toward transportation mode recognition using deep convolutional and long short-term memory recurrent neural networks. *IEEE Access*, *7*, 142353-142367.10.1109/ACCESS.2019.2944686

List of Abbreviations

Abbreviation	Definition
LPR	license Plate Recognition
CL	Cellphone Location
GPS	Global Positioning System
GIS	Geographical Information System
PoIs	Points of Interest
LBS	Location-Based Service
SECA	SEmi-Supervised Convolutional Autoencoder
MFFN	Multi-layer Feed Forward Neural Network
RF	Random Forest
IKASL	Incremental Knowledge Acquiring Self-Learning
LSTM	Long short-term memory
CART	Classification and Regression Tree
MAPE	Mean Absolute Percentage Error
RFM	Random Forest Method
AdaBoost	Adaptive Boosting
SVM	Support Vector Machine
MNL	Multinomial Logit
CL-TRANSMODE	CNN-LSTM-Transportation Mode
DNN	Deep neural networks
CNN	Convolutional neural networks
SHL	Sussex-Huawei Locomotion