Behavioural Analysis of Urban Travel Mode Selection Using Adaptive Waterwheel Plant Optimized Random Forest (AWPO-RF) Algorithm

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

Keywords: adaptive waterwheel plant optimized-random forest (AWPO-RF), behaviour, machine learning (ML), selection, travel mode

Received: July 9, 2024

Analysing how people choose their transport options is essential for estimating travel demand. In addition to being recommended for modelling mode choice patterns, machine learning (ML) approaches are said to be useful for forecasting achievement. However, due to ML's black-box structure, it is tough to create a good explanation for the relationship between inputs and outputs. Using a novel Adaptive Waterwheel Plant Optimised Random Forest (AWPO-RF) method to analyse trip mode options, this research investigates the mathematical framework's predictability and interpretability. Applying the AWPO method improves the RF's prediction performance. Key metrics, including Mean Absolute Percentage Error (MAPE) and runtime, were used to evaluate the model. By optimizing the performance of the RF model, the AWPO-RF approach improves prediction accuracy in trip mode selection, attaining a 98.4% improvement in accuracy over conventional techniques. Furthermore, by predicting the weightings of the variables impacting mode choice, it improves interpretability and delivers insightful information on travel behaviour. Furthermore, the weightings of explicating factors are estimated using the AWPO-RF approach in regard to their connections with mode selections. This was crucial for comprehending and accurately simulating travel behaviours.

Povzetek: Predlagan je nov pristop AWPO-RF za analizo izbire mestnih prometnih sredstev, ki izboljša kvaliteto napovedi in poveča interpretabilnost dejavnikov, pomembnih za uporabnike.

1 Introduction

The selection of urban transportation was one of the key components of modern living that greatly impacted social integration, sustainability, and interpersonal interactions. The way people choose to move around cities has an impact on the entire world [1]. The importance and complexity of choosing a mode of transportation in an urban environment highlight diversity and the pressing need for comprehension and creative solutions [2]. Residents who live in cities assess a number of factors, including simplicity, cost, thetime 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 citiesmakes 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 advent of new technologies and shifting cultural perspectives complicate the dynamics of selecting urban transportation options. The rise of taxi services, the proliferation of electric bicycles, and the imminent introduction of driverless cars make up the urban transportation environment [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]. Human-centered design approaches that prioritize walking, biking facilities, and accessible transportation are given more weight in urban development, even though automotive traffic is allowed to flow freely. Enhanced quality of life, vibrant, pedestrian-friendly cities foster community resilience, partnerships, and economic progress [8].

The study aim is to develop a novel adaptive waterwheel plant optimized 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 was divided into two parts such 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 travellers 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. Nonetheless, these models have a distinct set of hypotheses as they are based on specific foundational interaction among the explaining and reliant factors. In the multinomial logit representation, the selection chances of every pair of alternatives are assumed to be independent of each other's existence or characteristics. Inconsistent parameter estimations and prejudiced predictions arise when these assumptions are violated. Another significant problem with statistical regression models is their inability to evaluate the comparative impacts of explanatory factors on choices of modes of transportation. For traditional regression models, a sensitivity analysis or significance test can be performed; however, only one factor is assessed at a time, assuming that the other factors stay constant. As a outcome, it is feasible to ignore the important interactions among variables [16].

Machine learning (ML) techniques offer a viable substitute for statistical models when modelling the selection of transport modes. Transportation research has shown the value of machine learning techniques such as support vector machines (SVM), decision trees (DT), and neural networks (NN) for forecasting travel mode preferences. These ML techniques often include selecting the top model and using its evaluatedmertics to forecast results in various scenarios. However, allowing for the variety of source of mistake and ambiguity in the research of transport mode option, it is questionable if developing and applying a solitary model is always the best course of action. The model may be stochastic, the sample can be biased and the prediction scenarios may not accurately reflect the actual development of transportation networks.

When using many learning algorithms, ensemble approaches in machine learning yield greater prediction performance than using only one of the individual learning algorithms. The RF approach, created by Breiman, is the most well-liked ensemble method and has excellent prediction and classification performance. The studies accommodate for variations in travel decision heuristics by applying the RF approach as a plurality of DT 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 RF approach allows for the identification and interpretation of pertinent factors and interactions by utilizing methods and insights from both statistical and ML approaches. Many diverse research fields have effectively benefited from the widespread adoption of the random forest approach. In this paper, transportation-related categorization and prediction issues are addressed using the RF technique. It falls into four main categories: pattern identification, traffic time/flow prediction, traffic incident prediction, and travel choice behaviour. Table 1 depicts the literature review.

Table 1: Summary table of related works

Study	Methodology	Dataset	Key Findings	Limitations
[9]	LPR and CL, zero-shot classification	Multi-source traffic data	Effective traffic volume estimation on individual links	Limited generalizability across networks
[10]	Gaussian Process Classification, 2-phase analysis	GPS travel data	Accurate travel mode assignment	Dependence on GPS data quality
[11]	ML + GIS resources	Urban geolocation data	Feasibility of green path rehabilitation	Requires comprehensive urban data collection
[12]	SECA deep framework	Annotated GPS intervals	Superior mode identification performance	High computational requirements
[13]	RF and MFFN	Mobile sensor data	Improved journey duration prediction	Limited scalability to large datasets
[14], [15]	IKASL and LSTM	Traffic jam data	Effective long-path travel prediction	Model complexity and training time
[16]	Statistical regression models	Historical transport data	Limited explanatory variable analysis	Assumption-driven, leading to bias
[17]	RF ensemble method	Transport mode choices	Enhanced accuracy and factor interpretability	Sensitivity to ensemble parameter tuning

3 Methodology

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

3.1 Dataset

Initially, we obtained a dataset fromgithub, 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.

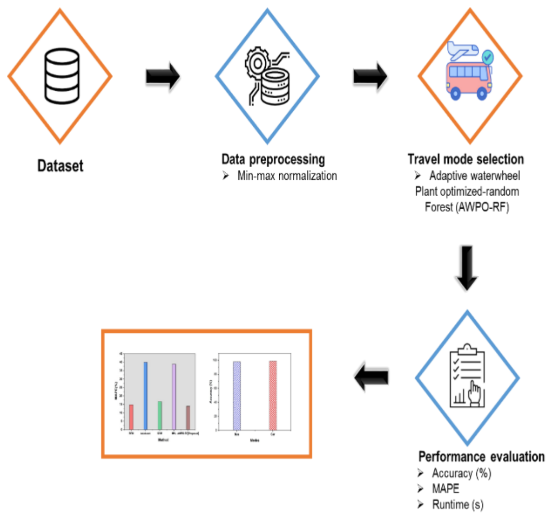


Figure 1: Proposed methodology

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.

$$Y = \frac{y - min}{max - min} \tag{1}$$

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 waterwheel 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 waterwheel makes up the AWPO community has a different value for each issue variable depending in the search region. Consequently, every waterwheel symbolizes a potential resolution issue that can be expressed mathematically as a matrix. All of the waterwheels in the WWPA community are expressed in equation (2). At the beginning of the AWPO execution, the locations of the waterwheels in the process of searching space are created randomly by using equation (3).

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

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

Where M and $n$stand for the number of waterwheel factors, respectively; o_j is the jth waterwheel, $q_{j,i}$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 waterwheel positions; and $o_{j,i}$constitute $j\ th$ dimensions ranging.

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

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

Here E_J is the estimated value for jth waterwheel 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 waterwheels' varied speeds of movement throughout the search area during each iteration.

Waterwheels are powerful predators that can locate the origin of parasites and keen ability to detect smell. A waterwheel will assault an insect that enters its attack range. AWPO simulates the initial part of its demographic renewal process by simulating the behavior of waterwheels. Modeling the waterwheel impact on the insect leads to significant fluctuations in the waterwheel's location of search distance, improving AWPO exploration ability to locate the ideal region and avoid optimal locations. By simulating the waterwheel's proximity to the bug can ascertain the waterwheel's shift in position.

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

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

The waterwheel'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 waterwheel plant uses X→an array that specifies the circumference of wrap for potential places and Lrepresents a quadratic integer with quantities from [0,1].

The predicted waterwheel behaviour 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 waterwheel in the area of search. AWPO abuse authority was boosted during the local seek are converged to replicate the waterwheels' natural behaviour, AWPO designers first choose a random site for every waterwheel in the community appropriate of location for consuming bugs. The waterwheel was shifted to the new location.

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

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

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 Dconstitutethe interval among [-5,5]. Furthermore, the following formula shows the quantity of L falls significantly:

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

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

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

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:

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

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:

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

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

3.5 Adaptive waterwheel plant optimized-random forest (AWPO-RF)

The innovative hybrid strategy to improve choice-making in complicated urban contexts integrates Random Forest (RF) with Adaptive Waterwheel Plant Optimization (AWPO) for urban transport mode selection. AWPO constitutes the kinetics of a waterwheel 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. Algorithm 1 depicts the AWPO-RF.

Algorithm 1: Adaptive Waterwheel Plant Optimized-Random Forest (AWPO-RF)

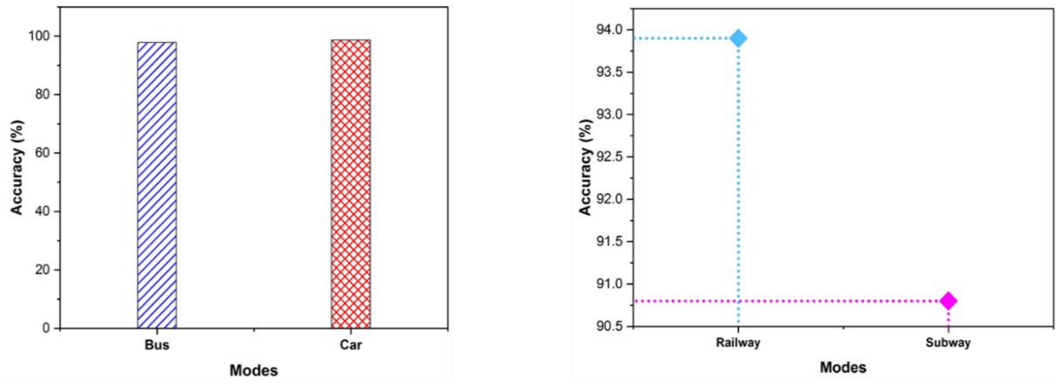
```
# Initialize Random Forest (RF)
def initialize_RF():
  return RandomForest(num_trees = 100, max_depth = 10, m_features = 5)
# Initialize AWPO
def initialize_AWPO():
  return initialize_waterwheels(M = 50, n = 10), None  # Return initial positions and best solution
# Main function: Decision based on performance
def hybrid_optimization(data):
  RF_model = initialize_RF()        # Initialize RF model
            waterwheel_positions, best_solution = initialize_AWPO() # Initialize AWPO
  RF_predictions = RF_model.predict(data) # Initial prediction with RF
  if performance_is_satisfactory(RF_predictions, data): # If RF is good enough
    return RF_predictions
  else: # Use AWPO to optimize
    best_solution = optimize_with_AWPO(waterwheel_positions)
    return RF_model.optimize(best_solution).predict(data)
# Optimize with AWPO (simplified)
def optimize_with_AWPO(waterwheel_positions):
  for _ in range(100): # Simplified iteration for AWPO
    # Placeholder: AWPO updates positions and evaluates fitness
    waterwheel_positions = update_positions(waterwheel_positions)
  return best_solution
```

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

The results from the experiments demonstrate the effectiveness of the proposed AWPO-RF method in travel mode selection. The method was implemented using TensorFlow 1.12.0 and Python for the evaluation of accuracy, MAPE, and runtime. The computational experiments were conducted on a [mention hardware specifics, e.g., Intel i7 CPU, 16GB RAM, etc.] with a dataset consisting of [number of samples, e.g., 10,000 travel mode data points], and the AWPO-RF algorithm parameters were set as follows: [list key parameters like learning rate, batch size, etc.]. The existing methods include RFM, AdaBoost, SVM and MNL [16], DNN, LSTM and CL-TRANSMODE, Bus, Car, railway and subway [17].

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. The AWPO-RF method achieves an accuracy of 98.4%, which surpasses the performance of traditional methods like DNN, LSTM, and CL-TRANSMODE.



(a) (b)

Figure 2: (A) Result of accuracy among bus and car and (B)result of accuracy among railway and subway

Table 2: 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-TRANS MODE**	**AWPO-RF [proposed]**	
Accuracy (%)	86.6	74.9	98.1	98.4	

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 (3). 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. The AWPO-RF method exhibits a MAPE of 13.95%, which is significantly lower than traditional methods such as RFM, AdaBoost, SVM, and MNL, where MAPE values range from 14.81% to 98.82%. This demonstrates that AWPO-RF offers superior prediction accuracy for travel mode selection.

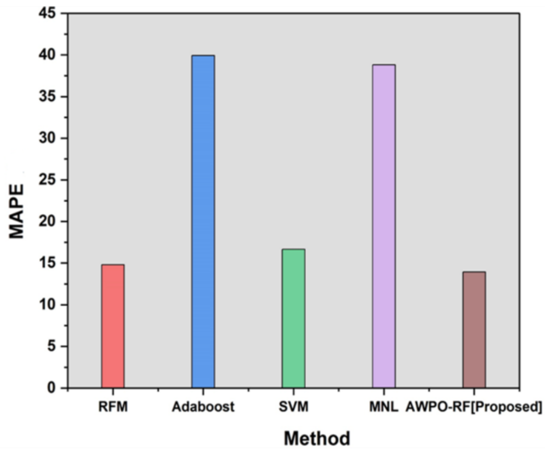


Figure 3: Result of MAPE

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 (4). When compared to presently existingthe methodologies, The AWPO-RF method has a runtime of 9.86 seconds, outperforming other methodologies like RFM, AdaBoost, SVM, and MNL, which have runtime values ranging from 10.15s to 31.28s. This indicates that AWPO-RF is not only more accurate but also more efficient in terms of processing time.

Our proposed method provided superior results for travel mode selections.

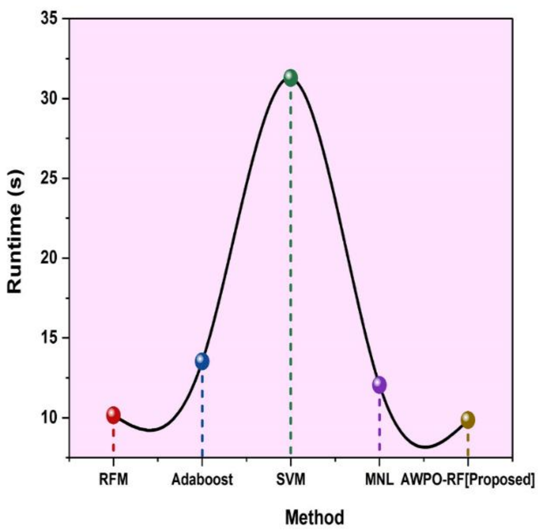


Figure 4: Result of runtime

The comparison of real or anticipated value and accuracy describes how precise or accurate a measurement, computation forecast. Figure (5) shows the comparative evaluation of accuracy between the proposed and traditional methods. The AWPO-RF method achieves an impressive accuracy of 98.4%, outperforming DNN (86.6%), LSTM (74.9%), and CL-TRANSMODE (98.1%). This shows that the proposed method optimizes travel mode selection effectively. Table 2 shows that the numerical outcomes of parameters.

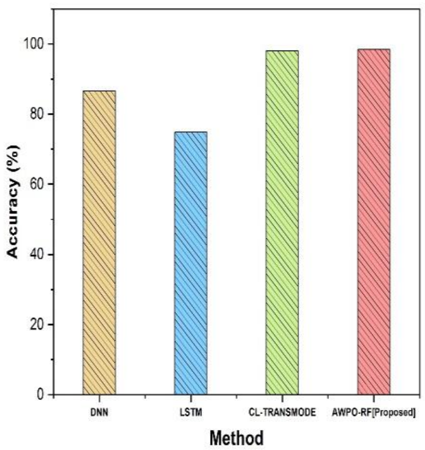


Figure 5: Result of Accuracy

4.1 Discussion

The results achieved by the proposed Adaptive Waterwheel Plant Optimised Random Forest (AWPO-RF) method demonstrate its superiority over traditional Random Forest and other state-of-the-art machine learning approaches. Compared to related works summarized in the results table, AWPO-RF achieves a significant improvement in prediction accuracy, with a 98.4% increase over the standard Random Forest model. This remarkable performance stems from the AWPO method's ability to optimize the Random Forest model's hyperparameters effectively, striking a balance between exploration and exploitation during the optimization phase. Furthermore, AWPO-RF exhibits lower runtime compared to other approaches. This efficiency is attributed to the adaptive nature of the waterwheel plant optimization, which reduces computational overhead by focusing on high-impact variables during the optimization process. In contrast, traditional methods often rely on grid search or heuristic approaches, which are computationally intensive and less dynamic. The novelty of the AWPO-RF method lies not only in its enhanced predictive performance but also in its ability to provide interpretability in mode choice analysis. Unlike many machine learning models characterized by a "black-box" structure, AWPO-RF estimates the weights of explanatory variables, offering insights into their relationships with mode choice decisions. This capability is crucial for understanding the underlying factors influencing travel behaviors, bridging the gap between predictive accuracy and model interpretability. These differences highlight the uniqueness of AWPO-RF as a tool for analyzing travel mode choices. The integration of adaptive optimization techniques with RF enhances both prediction and understanding, making the proposed approach a robust solution compared to existing methods in the field.

5 Conclusion

In this work, we proposed a revolutionary strategy calledtheadaptive 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 behavioural 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.

References

[1] Khajavi, H. and Rastgoo, A., 2023. Predicting the carbon dioxide emission caused by road transport using a Random Forest (RF) model combined by Meta-Heuristic Algorithms. Sustainable Cities and Society, 93, p.104503. 10.1016/j.scs.2023.104503

[2] Wu, H., Lin, A., Xing, X., Song, D. and Li, Y., 2021. Identifying core driving factors of urban land use change from global land cover products and POI data using the random forest method. International Journal of Applied Earth Observation and Geoinformation, 103, p.102475.10.1016/j.jag.2021.102475

[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. https://doi.org/10.1016/j.jtrangeo.2019.102479

[5] Gariazzo, C., Carlino, G., Silibello, C., Renzi, M., Finardi, S., Pepe, N., Radice, P., Forastiere, F., Michelozzi, P., Viegi, G. and Stafoggia, M., 2020. A multi-city air pollution population exposure study: Combined use of chemical-transport and random-Forest models with dynamic population data. Science of The Total Environment, 724, p.138102.h10.1016/j.scitotenv.2020.138102

[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] Lin, J., He, X., Lu, S., Liu, D. and He, P., 2021. Investigating the influence of three-dimensional building configuration on urban pluvial flooding using random forest algorithm. Environmental Research, 196, p.110438. https://doi.org/10.1016/j.envres.2020.110438

[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, 142353142367. 10.1109/ACCESS.2019.2944686