

Investigating the Use of Active Transportation Modes Among University Employees Through an Advanced Decision Tree Algorithm

Mahdi Aghaabbasi^{1*}, Muhammad Zaly Shah², Rosilawati Zainol¹

¹ Centre for Sustainable Urban Planning and Real Estate (SUPRE), Department of Urban and Regional Planning, Faculty of Built Environment, University of Malaya, Malaysia

² Center for Innovative Planning and Development (CIPD), Department of Urban and Regional Planning, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Malaysia

* Correspondence: mahdi@um.edu.my

SUBMITTED: 15 September 2021; REVISED: 27 October 2021; ACCEPTED: 28 October 2021

ABSTRACT: Now more than ever, the health and economic benefits of active transportation (AT) are evident and several planning efforts and programs are particularly targeted at improving active transportation options for different populations, such as students and seniors. Administrative employees at universities received less attention in the literature than other population groups. This population spends a lot of time doing sedentary activities and behaviors during their working time. Thus, the present study used a C5 decision tree to examine the usage of university employees' AT modes when they are out of campus to get to work, shopping, and leisure. The effects of the sociodemographic and living environment of employees on their AT mode choice were also examined. According to the results, walking was the most frequently used mode to get to work and leisure and public transport was the most frequently used mode to get to shopping. Transit station conditions (25), sidewalk availability and coverage (36), and bike path availability and coverage (30) were the most important factors in the use of AT modes by employees to get to work, shop, and leisure, respectively. Furthermore, several decision rules were extracted from the C5 tree, which included combinations of multiple factors.

KEYWORDS: Active transportation, mode choice, university employees, trip purposes, C5

1. Introduction

Active transportation (AT) refers to all forms of human-powered transportation, such as walking and cycling [1]. In addition, the [2] considers rolling and public transportation as active modes of getting to essential destinations such as school and work. AT is also considered as a recreational activity for all people. Several studies have been conducted to identify the factors that influence the use of AT modes by general and special group populations such as children, schoolchildren, women, and seniors [3-11]. This wide range of benefits of AT for societies motivates policymakers and urban planners to identify factors and attributes that contribute to a higher level of AT modes' usage [12]. To this end, several studies were conducted to identify the influencing factors on the usage of AT modes by general and special group populations

such as children, school children, women, and seniors [13-21]. However, some special groups, such as employees, attracted less attention from academia in relation to the use of AT modes.

Millions of employees across the world are more likely to drive a car or motorcycle to work, leisure, or shopping each day than they are to use any other modes. This dependence on motor vehicles might contribute to negative health outcomes caused to some extent by inactivity. This dependency can also be worse in developing countries where the rate of vehicle ownership is extremely high. Malaysia, for example, has the highest vehicle ownership rate in the ASEAN region [22] and the third highest level of vehicle ownership worldwide (93%) [23]. Thus, encouraging AT modes provides an opportunity to address improving employees' health through physical activity. However, several areas in developing countries lack characteristics that make them well-suited for AT. For example, an ineffective public transport system and sprawl are significant barriers to reducing vehicle commuting in developing countries.

Staff of universities can benefit from AT promotion like other employees. Unlike the academic staff of universities who have a more fruitful schedule and activities, the administrative staff have a monotonous schedule, which might cause them to lack physical activity options in their offices at the universities. Administrative university staff might find it difficult to be physically active while at work and are less likely to engage in moderate physical activity and are more likely to engage in sedentary behaviors, such as sitting. Thus, it is important to identify the ways to promote the use of AT modes among this group. While a number of studies have examined the workplace factors that influence the use of AT modes by university employees [24], no study is available to assess the factors that affect the use of AT by this group when they are out of their workplaces. Identifying such factors would help in developing better AT programs, where university employees would be incentivized to start using the use of AT modes in their routine life.

Several studies examined factors that affect individuals' travel mode choice. Based on previous studies, factors related to sociodemographic, trip characteristics and purpose, and built environment are among the most frequently used factors to examine the travel mode choice of the general population [25, 26]. Individuals' personal characteristics influence their travel mode choice both directly and indirectly [27-31]. Factors including age, race, education, gender, vehicle ownership, occupation, and income are regarded as directly influential on the mode of choice (Figure 1). Besides, a number of the abovementioned factors influence the destination choice [27]. For example, in China, women were more interested in choosing a workplace near their houses than men, because they had to take care of family members and they devoted less time to work compared to men. In fact, the destination itself might define the characteristics of the trip and the activity. These characteristics, include travel distance, travel time, and travel cost. Then, a certain mode is chosen based on these characteristics. While sociodemographic factors define the trip characteristics, built environment attributes also influence the trip characteristics as well [27].

Built-environment attributes such as density and land use mixture are regarded as influential factors on travel distance and time [32]. A proper implication of density and a mixture of land use might contribute to a higher number of short trips by individuals. Indeed, these characteristics provide housing close to other activities and access to facilities within a certain radius.. Several studies confirmed that these short distances can assist individuals in making these trips using active modes such as walking and biking [33-37]. The mixture of land

use and service density near the workplace can also reduce the use of motorized transport and encourage employees to use active modes such as walking and biking [38].

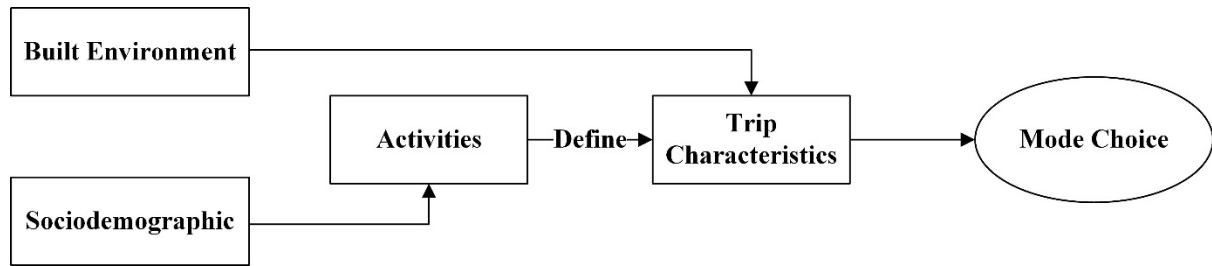


Figure 1. The possible impact of sociodemographic, built environment, and trip characteristics on mode choice.

In terms of AT research, several studies examined the impacts of sociodemographic and built environment attributes on AT mode choice. Previous study showed that the impact of sociodemographic factors (in this order) with AT to/from school and to other destinations in Hong Kong adolescents [39]. They found that age had a positive association with the frequency of AT; education had a positive association with the total weekly minutes of AT; household income, vehicle ownership, and the number of motor vehicles in the household had a positive and negative association, respectively, with the weekly frequency of non-school AT. In New Zealand, a number of studies found that younger ages and fewer vehicles were associated with AT to school among adolescents and children [40, 41]. In the Netherlands, it was found that those aged 25-44 were more likely to use AT modes for the purpose of sport. In addition, the likelihood of AT usage among individuals with university and college degrees was higher than among individuals with lower education levels [42].

Concerning the impacts of built environment attributes on AT, Yang et al. found that built environment attributes such as street connectivity and distance to certain facilities (e.g. parks) influenced the use of AT by older adults [13]. The presence of a sidewalk and a bike lane was also identified as important built-environmental factors in promoting AT modes [43-49]. Aziz et al. discovered that appropriate sidewalk width, length of the bike lane, and proportion of protected bike lane will increase the likelihood of more people using AT modes for their home-to-work commute [50]. Land use mixture were found to have a positive relationship with AT usage among different people groups in several studies [51-56]. Access to destinations, which are regarded as the presence or number of a range of destinations, including shops, banks, worship places, and so on, is shown by numerous studies to have a positive relationship with AT usage [43, 57-65].

Numerous studies investigated the travel behavior of students and staffs in university setting [26, 66-80]. These studies typically assess the impacts of several factors, including sociodemographics and the built environment, on the travel behavior patterns of the university population.. Sociodemographic attributes such as age, gender, household income, race, vehicle ownership, academic status (undergraduate and postgraduate), residential location (on-campus and off-campus) and enrolment (part-time and full-time) have been shown to affect the mode choice of university students [25, 66, 71, 75, 81-83]. In addition to the above, several studies showed that built environment attributes such as neighbourhood type, mixture of land use, density, bicycle and pedestrian network connectivity, proximity to a bus stop, influence the mode choice of university students [67, 75, 80, 84].

Among the university population, students receive more attention than staff. Zhou studied the employees' carsharing in University of California, Los Angeles (UCLA) [85]. The author identified that employees' gender and income influence their carsharing mode choice. Wang and Liu identified factors influencing public transport use by the staff in University of Queensland (UQ), and these factors were travel time and distance [86]. Travel time also were mentioned by Shannon et al. [87], who attempted to identify influential factors on the use of AT in the University of Western Australia. Uttley and Lovelace examined commuting behaviour and long-term behavioural shifts towards cycling in response to outside intervention at the organisational level in University of Sheffield [88]. They also presented information regarding the staffs' mode of travel, as well as commute distance and time.

Most AT-relevant attributes in the general population, and especially in children and seniors, can be thought of as encompassing both built environment and sociodemographic attributes. Overwhelmingly, most research examines AT modes' usage among children, seniors, and the general population through the aforementioned factors. AT mode usage by employees while they are out of their workplaces is less studied. It is important to examine the AT usage among different employees since the individual's occupation can affect the travel mode choice and AT usage [25, 27, 28].

In light of the above, this study was carried out through the analysis of the active travel data of university employees in Malaysia. Precisely, the aim of the study was to: (1) investigate the AT mode choice of university employees for different trip purposes; (2) identify the most important factors influencing the use of AT modes by university employees for different trip purposes; (3) identify behavioral patterns that occur within the active travel data by rule extraction technique; and (4) provide insights for the development of AT improvement strategies focused on university employees. A data mining technique was performed to analyze the AT data, which aims to extract knowledge from observed data previously unknown. It is a free assumption and does not require a priori probabilistic knowledge about the phenomena of interest. The main reason for using data mining techniques was also related to the complex character of the mode choice phenomenon. A mode choice can be viewed as a multi-factor event always preceded by a situation in which one or more built-environment and sociodemographic factors affect the mode choice decision. In fact, each mode choice is the result of a chain of factors which have some common actors in several mode choice decisions. Using data mining techniques to identify these factors and their interdependence can provide useful insights for the development of effective land use and transportation strategies. Precisely, the data mining method that we used was decision tree (DT), which is also called hierarchical tree-based regression (HTBR). Among several algorithms of HTBR, C5 was chosen to build the model to predict the AT mode choice of university employees. Concerning AT mode choice, identifying important factors is the key to developing an effective model to characterize future transport planning initiatives that promote AT modes among different groups of people..

The remainder of the paper is structured as follows: section 2 provides details of the statistical methodologies used in the analysis and describes the case study; section 3 explains the results of the C5 trees and of their decision rules; and finally, a discussion places the results in the context of the practice of urban and transport planning, and conclusions are drawn.

also attempted to identify the factors associated with more usage of AT modes by university employees. Eventually, we used the outcomes of the questionnaire to construct models that can predict employees' usage of AT modes. The questionnaire included 16 items related to sociodemographic information of respondents, their living environment, and their daily travel distance, time, and cost based on different trip purposes. Table 1 presents the questionnaire items, their types and values.

Table 1. Input and target variables used in this study.

Variable	Description	Type	Value
<i>Sociodemographic</i>			
A1	Age	Ordinal	(1) 18-29; (2) 30-49; (3) 50-64; (4) 65 years or more
G1	Gender	Nominal	(1) male; (2) female
E1	Education	Ordinal	(1) primary; (2) secondary; (3) diploma; (4) bachelor's degree; (5) master's degree; (6) doctorate degree
EM	Employment mode	Nominal	(1) part-time; (2) full-time
I1	Income	Interval	(1) less than RM 1000; (2) between RM 1000 and RM 2000; (3) between RM 2000 and RM 3000; (4) between RM 3000 and RM 6000; (5) between RM 6000 and RM 13000; (6) more than RM 13000
R1	Race	Nominal	(1) Malay; (2) Chinese; (3) Indian; (4) others
VO	Vehicle ownership	Nominal	(1) car; (2) motorcycle; (3) both car and motorcycle; (4) no
<i>Built environment</i>			
N1	Neighborhood	Nominal	(1) residential only; (2) residential with some commercial buildings; (3) residential with some industrial facilities; (4) a commercial area with some residential; (5) an industrial area with some residential; (6) mixed residential and commercial
NF	Neighborhood facilities in a walkable distance	Nominal	(1) schools; (2) childcare facilities; (3) public transport; (4) taxi station; (5) shops; (6) banks; (7) healthcare facilities; (8) leisure facilities; (9) parks and other open spaces; (10) place of worship; (11) two of facilities; (12) three or more facilities
SAC	Sidewalk availability and coverage	Nominal	(1); No sidewalks; (2) Discontinuous, narrow sidewalks; (3) narrow sidewalks along all major streets; (4) adequate sidewalks along all major streets
BAC	Bike path availability and coverage	Nominal	(1) Little or none; (2) some bike paths or routes; (3) many cycle paths, lanes, or routes forming network
TSC	Transit stops conditions	Nominal	(1) No shelters; (2) some bus stop shelters; (3) widely available bus stop shelters
<i>Trip characteristics</i>			
DTD	Daily travel distance	Ordinal	(1) ≤ 1 km; (2) $1 \text{ km} < D \leq 5 \text{ km}$; (3) $5 \text{ km} < D \leq 10 \text{ km}$; (4) $D > 10 \text{ km}$
DTT	Daily travel time		(1) $\leq 10 \text{ min}$; (2) $10 \text{ min} < T \leq 15 \text{ min}$; (3) $15 \text{ min} < T \leq 30 \text{ min}$; (4) $T > 30 \text{ min}$
DTC	Daily travel cost	Ordinal	(1) $C \leq \text{RM } 5$; (2) $\text{RM } 5 < C \leq \text{RM } 15$; (3) $\text{RM } 15 < C \leq \text{RM } 25$; (4) $D > \text{RM } 25$
ATM*	Active travel mode	Nominal	(1) public transport; (2) walking; (3) biking; (4) private vehicle (car and/or motorcycle+ public transport; (5) private vehicle (car and/or motorcycle+ walking and/or biking; (6) public transport+ walking and/or biking; (7) public transport+ ridesourcing; (8) walking and/or biking+ ridesourcing

*Target variable

2.2 Hierarchical tree-based regression (HTBR)

In the present study, the AT mode of choice is the target variable, classified into eight values. Several traditional techniques, such as multinomial logistic regression and ordinal logistic regression, are available to model the categorical variables. However, the power and accuracy of these techniques are prone to be negatively affected by the violation of their pre-defined assumptions and functions [89]. To address the aforementioned flaws, hierarchical tree-based regression (HTBR) can be used as an assumption-free option suitable for forward stepwise and non-parametric variable selection [90]. Furthermore, the HTBR is extremely efficient for problems' classification and prediction [91]. This method identifies both exclusively and exhaustive groups/subgroups of the target variable that share common characteristics and impact the dependent variable of interest.

Several algorithms of HTBR are available, including CHAID, QUEST, CART, and C5. The present study used the C5 algorithm, which was developed by Quinaln [92] and is one of the most commonly used algorithms in HTBR in transport research [93, 94]. The C5 algorithm uses both continuous and categorical data as input, and categorical data as an output variable. C5 creates non-binary trees using the information gain ratio. To prepare data for analysis, we used Automated Data Preparation (ADP) in SPSS Modeler software. The ADP allows for making data ready for quick and easy model building, without requiring prior knowledge of the statistical concepts involved. The cross-validation method (10-folds) was also used to assess the trees' structure and their accuracy.

The results of this paper are reported in two forms: the decision tree and the "if-then". The latter is a presentation form of decision trees. In this format, the "if" represents an attribute set of one or more variable and "then" represents the dependent variable state. Thus, in this study, the dependent variable (consequent) is the usage of AT modes of university employees based on their trip purpose, and the independent variables (antecedent) are factors that might affect the use of these modes. Typically, the number of rules achieved is equal to the number of tree child nodes. However, we used three parameters in every possible rule to extract significant rules which can provide useful information for implementing planning policies to increase the level of AT. In order to show how often any combination of antecedent and consequent occurs in a database, a measure of support (S) is used. To measure the cases' percentage in which a consequent appears given that the antecedent has occurred, a measure of probability (P) is defined. The third parameter is population (Po), which measures the statistical dependence of the rule by relating the observed frequency of co-occurrence to the expected frequency of co-occurrence under the assumption of conditional independence ($P = S/P$) [95]. In this study, to obtain high quality rules, those that have $P\% \geq 75$, $S\% \geq 0.9$, $Po\% \geq 1.2$ were chosen as significant.

3. Result

The present study conducted a questionnaire survey among 511 UTM administration staff. After a careful review of the completed questionnaires, the respondents that reported to use the active modes for purposes of work, shopping, and leisure were selected to participate in this study. Thus, 45 participants that reported using AT modes to travel to work, 38 of those that reported using AT modes to travel to shop, and 84 of those that reported using AT modes to travel to leisure destinations were included in the study. The sociodemographic of the

participants and built environment attributes reported by them are presented in Appendixes A and B, respectively. The frequency of daily travel distance and time by the trip purposes of participants are shown in Table 2. In work, shopping, and leisure categories, DTC of below RM 5 and DTT of 10 min and below have the greatest frequencies. While the highest DTD frequency related to travel to work and leisure is between 1 km and 5 km, the highest DTD frequency related to travel to shopping is between 5 km and 10 km.

The frequencies of the usage of AT modes are presented in Table 3. In the category of "work", walking has the greatest usage frequency, while no participant used the combination of public transport and ride sourcing. Public transport had the highest usage frequency in the "shopping" category, while no UTM employees used a combination of private vehicles and active modes or a combination of public transport and ride sourcing. In the "leisure" category, the greatest usage frequency belongs to walking, and the frequency of the combination of private vehicles and active modes is zero. Between the categories, walking for leisure has the greatest usage frequency (50%).

Table 2. Average daily travel distance, time, and cost by trip purpose.

	Work		Shopping		Leisure	
	Frequency	%	Frequency	%	Frequency	%
Daily travel distance (km)						
≤1 km	5	11.1	3	7.9	24	28.6
1 km < D ≤ 5 km	24	53.3	12	31.6	44	52.4
5 km < D ≤ 10 km	10	22.2	13	34.2	6	7.1
D > 10 km	6	13.3	10	26.3	10	11.9
Daily travel cost (RM)*						
C ≤ RM 5	32	71.1	19	50.0	65	77.4
RM 5 < C ≤ RM 15	9	20.0	13	34.2	8	9.5
RM 15 < C ≤ RM 25	2	4.4	4	10.5	8	9.5
D > RM 25	2	4.4	2	5.3	3	3.6
Daily travel time (min)						
≤ 10 min	31	68.9	19	50.0	65	77.4
10 min < T ≤ 15 min	7	15.6	6	15.8	8	9.5
15min < T ≤ 30 min	6	13.3	11	28.9	8	9.5
T > 30 min	1	2.2	2	5.3	3	3.6

*RM 1.00 = 0.24 USD

Table 3. Frequency of AT modes usage by trip purpose.

	Work		Shopping		Leisure	
	Frequency	%	Frequency	%	Frequency	%
Public transport	11	24.4	15	39.5	10	11.9
Walking	19	42.2	12	31.6	42	50.0
Biking	2	4.4	4	10.5	15	17.9
Private vehicle (car and/or motorcycle+ public transport	3	6.7	0	0	0	0
Private vehicle (car and/or motorcycle+ walking and/or biking	2	4.4	0	0	0	0
Public transport+ walking and/or biking	2	4.4	0	0	9	10.7
Public transport+ ride sourcing	0	0	4	10.5	2	2.4
Walking and/or biking+ ride sourcing	6	13.3	3	7.9	6	7.1
TOTAL TRIPS	45		38		84	

3.1. Relative importance of influential factors

Figure 3 shows the relative importance of input variables for work, shopping, and leisure trips of university employees through the AT modes. These influential factors and their importance were identified through the C5 decision trees. One C5 model was developed for each trip's purpose. For work purposes, sidewalk availability and coverage (SAC) and daily travel time (TDT) were identified as the most important factors, with a relative importance of 16%. SAC is also the most important factor for shopping purposes (relative importance = 36%). For leisure trips, biking path availability and coverage (BAC) has the highest relative importance (30%). The least important factor for work is neighborhood type (N1), with a relative importance of 12%. For both shopping and leisure trips, the least important factor is neighborhood facilities within a walkable distance (NF), with a relative importance of 6% and 20%, respectively.

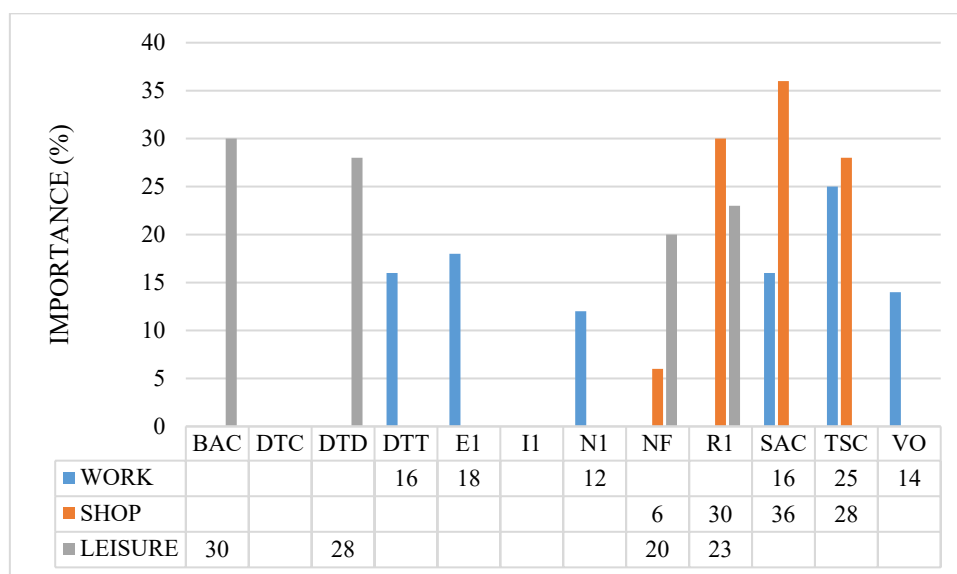


Figure 3. Importance of input variables with C5.

3.2 Work purpose

Figure 4 shows the C5 model that predicts the use of active transport modes for travelling to the workplace. The overall accuracy of this model is 71.11%. The tree identified eight associate rules with the use of AT modes (Table 4). Among these rules, five rules were identified as significant ($P\% = 75$, $S\% = 0.9$, and $Po\% = 1.2$). The model shows that the node 0 is split based on the variable of transit stop conditions (TSC) and this shows that the TSC is the single best variable to predict the use of AT modes by UTM employees to travel to work. C5 directs the involved TSCs of "some bus stop shelters" and "widely available bus stop shelters" to the left, forms node 1; directs the TSC of "no shelters" to the right, forms node 14. As shown by node 14, if an employee perceives the transit stop conditions in his/her living area as "no shelters", the tree predicts that the employee's AT mode choice to work is most likely to be walking (Rule 8: $P\% = 68.75$, $S\% = 24.44$, $Po\% = 35.56$ /non-significant).

Concerning the division of node 1 and based on TSC and an employee's daily travel time (DTT) to work, this node is split based on DTT. The C5 sends the DTT value of 10 min and below to the node 2; the node 13 sends the DTT value of greater than 10 min to the nodes 13. As shown by node 13, if an employee's daily travel time to work is greater than 10 min and

he perceives the transit stop conditions in his/her living area as "some bus stop shelters" or "widely available bus stop shelters", the tree predicts that the employee's AT mode choice to work is most likely to be public transport (Rule 7: $P\% = 50$, $S\%=8.89$, $Po\%=17.78$ /non-significant).

Node 2 is further split in relation to sidewalk availability and coverage (SAC) and C5 sends the SAC values of "no sidewalks" and "discontinuous, narrow sidewalks" to node 3; sends the SACs of "narrow sidewalks along all major streets" and "adequate sidewalks along all major streets" to node 10. Nodes 3 and 10 are further split in relation to SAC and vehicle ownership (VO). At node 3, the C5 sends the SACs of "no sidewalks" and "discontinuous, narrow sidewalks" to nodes 4 and 5, respectively. As shown by node 4 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters" and DTT of "10 min and below", the tree predicts that the employee's AT mode choice to work is most likely to be a combination of private vehicle and public transport if an employee perceives the sidewalk conditions in his/her living area as "no sidewalks" or "discontinuous, narrow sidewalks" (Rule 1: $P\% = 100$, $S\%=4.44$, $Po\%=4.44$). At node 10, the C5 sends the VO of "no" and "yes" to nodes 11 and 12, respectively. As shown by node 11 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters", DTT of "10 min and below", and SACs of "narrow sidewalks along all major streets" or "adequate sidewalks along all major streets", the tree predicts that the employee's AT mode choice to work is most likely to be walking if an employee does not own a private vehicle (Rule 5: $P\% = 75$, $S\%=6.67$, $Po\%=8.89$). As shown by node 12 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters", DTT of "10 min and below", and SACs of "narrow sidewalks along all major streets" or "adequate sidewalks along all major streets", the tree predicts that the employee's AT mode choice to work is most likely to be a combination of walking, biking, and ridesourcing if an employee owns a private vehicle (Rule 6: $P\% = 60$, $S\%=6.67$)

Turning to division of node 5 and conditioned on SAC, DTT, TSC, and employees' education (E1), the C5 sends the E1 values of "primary", "secondary", and "diploma" to node 6; sends the E1 values of "bachelor's degree", "master's degree", "doctorate degree" to node 7. As shown by node 6 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters", DTT of "10 min and below", and SACs of "no sidewalks" or "discontinuous, narrow sidewalks", the tree predicts that the employee's AT mode of choice to work is most likely to be walking if an employee's education is primary, secondary, or diploma (Rule 2: $P\% = 100$, $S\%=4.44$, $Po\%=4.44$). The C5 further splits node 7 in relation to the neighborhood type (N1) and sends the N1 values of "commercial area with some residential" and "mixed residential and commercial" to node 8; sends the N1 values of "residential only", "residential with some commercial buildings", and "residential with some industrial facilities" to node 9. As shown by node 8 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters", DTT of "10 min and below", and SACs of "no sidewalks" or "discontinuous, narrow sidewalks", E1 values of "primary", "secondary", or diploma, the tree predicts that the employee's AT mode choice to work is most likely to be walking if the type of employee's neighborhood is "commercial area with some residential" or "mixed residential and commercial" (Rule 3: $P\% = 100$, $S\%=4.44$, $Po\%=4.44$). At node 9 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters", DTT of "10 min and below", and SACs of "no sidewalks" or "discontinuous, narrow sidewalks", E1 values of "primary", "secondary", or diploma, the tree predicts that the employee's AT mode choice to work is most

likely to be public transport if the type of employee's neighborhood is "residential only", "residential with some commercial buildings", or "residential with some industrial facilities" (Rule 4: P% = 83.33, S%=11.11, Po%=13.33).

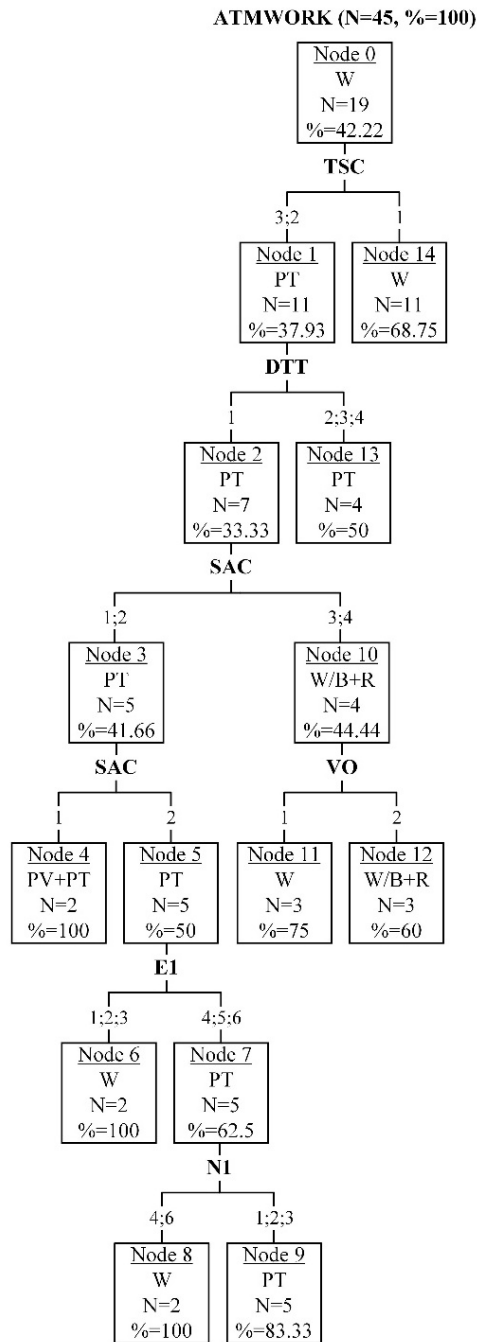


Figure 4. C5 model-predicting employee's AT mode choice for work purpose.

Table 4. Description of the rules according to the C5 for work purpose.

Rule ID	Node	Rule description	Then	P%	S%	Po%
1	4	(SAC = 1;2) AND (DTT = 1) AND (TSC = 3;2)	PV+PT	100	4.44	4.44
2	6	(E1 = 1;2;3) AND (SAC = 1;2) AND (DTT = 1) AND (TSC = 3;2)	W	100	4.44	4.44
3	8	(N1 = 4;6) AND (E1 = 1;2;3) AND (SAC = 1;2) AND (DTT = 1) AND (TSC = 3;2)	W	100	4.44	4.44
4	9	(N1 = 1;2;3) AND (E1 = 1;2;3) AND (SAC = 1;2) AND (DTT = 1) AND (TSC = 3;2)	PT	83.33	11.11	13.33
5	11	(VO = 1) AND (SAC = 3;4) AND (DTT = 1) AND (TSC = 3;2)	W	75	6.67	8.89
6*	12	(VO = 2) AND (SAC = 3;4) AND (DTT = 1) AND (TSC = 3;2)	W/B+R	60	6.67	11.11
7*	13	(DTT = 2;3;4) AND (TSC = 3;2)	PT	50	8.89	17.78
8*	14	(TSC = 1)	W	68.75	24.44	35.56

*Rule excluded
PT=Public Transport; W=Walking; B=Biking; PV=Private Vehicle; R=Ridesourcing

3.3 Shopping purpose

Figure 5 shows the C5 model that predicts the use of active transport modes for travelling to shopping. The overall accuracy of this model is 92.11%. The tree identified seven associate rules with the use of AT modes (Table 5). Among these rules, five rules were identified as significant (P% = 75, S% = 0.9, and Po%=1.2). The model shows that the node 0 is split based on the variable of transit stop conditions (TSC) and this shows that the TSC is the single best variable to predict the use of AT modes by UTM employees to travel to shopping. C5 directs the involved TSCs of "some bus stop shelters" and "widely available bus stop shelters" to the left, forms node 1; directs the TSC of "no shelters" to the right, forms node 7. The nodes 1 and 7 are further split in relation to SAC and employees' race (R1), respectively. Conditioned on TSC and SAC, the C5 splits node 1 based on SAC and sends the SACs of "narrow sidewalks along all major streets" and "adequate sidewalks along all major streets" to node 2; and sends the SACs of "no sidewalks" and "discontinuous, narrow sidewalks" to node 3. As shown by node 2, if an employee perceives the transit stop conditions in his/her living area as "some bus stop shelters" or "widely available bus stop shelters", and perceives the sidewalk conditions as "no sidewalks" or "discontinuous, narrow sidewalks", the tree predicts that the employee's AT mode of choice to work is most likely to be public transport (Rule 1: P% = 73.33, S%=28.95, Po%=39.48/non-significant). Node 3 is further split based on neighborhood facilities within a walkable distance (NF). The C5 sends the NF value of "parks and other open spaces" to node 4; sends the NF values of "childcare facilities; "public transport", and "three or more facilities" to node 5; sends the NF values of "schools", "shops", "banks", "place of worship", and "two facilities" to node 6. At node 4 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters", and SACs of "narrow sidewalks along all major streets" and "adequate sidewalks along all major streets", the tree predicts that the employee's AT mode choice to work is most likely to be a combination of walking, biking, and ride sourcing if the employee's neighborhood has parks and other open spaces within a walkable distance (Rule 2: P% = 100, S%=26.63). As shown by node 5 and conditioned on TSCs of "some bus stop

shelters" or "widely available bus stop shelters", and SACs of "narrow sidewalks along all major streets" and "adequate sidewalks along all major streets", the tree predicts that the employee's AT mode choice to work is most likely to be walking if the employee's neighborhood has childcare facilities, public transport, or three or more facilities in a walkable distance from his/her house (Rule 3: P% = 100, S%=7.89, Po%=7.89). At node 6 and conditioned on TSCs of "some bus stop shelters" or "widely available bus stop shelters", and SACs of "narrow sidewalks along all major streets" and "adequate sidewalks along all major streets," the tree predicts that the employee's AT mode of choice to work is most likely to be public transport if the employee's neighborhood has schools, shops, banks, places of worship, and "two facilities" within a walkable distance from his/her house (Rule 4: P% = 80, S% = 10.53, Po% = 80).

Turning to division of node 7, conditioned on the TSC of "no shelters" and different values of employees' race (R1), C5 divides node 7 based on R1 and directs the races of "Indian" and "others" to node 8; directs the race of "Chinese" to node 9; directs the race of "Malay" to node 10. The tree predicts that the employee's AT mode choice to shopping is most likely to be biking if the employee's race is Indian or others (Rule 5: P% = 66.66, S%=5.26, Po%=7.90). The tree continues to predict that the employee's AT mode choice for shopping is most likely to be a combination of walking, biking, and ride sourcing if the employee's race is Chinese (Rule 6: P% = 100, S%=2.63, Po%=2.63). The tree then predicts that the employee's AT mode choice to shopping is most likely to be walking if the employee's race is Malay (Rule 7: P% = 80, S%=21.05, Po%=26.32).

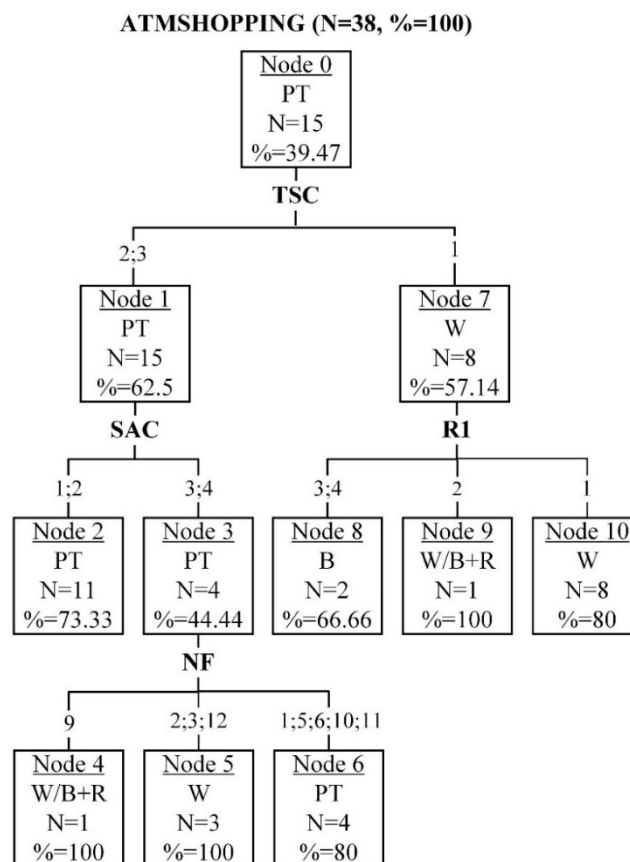


Figure 5. C5 model-predicting employee's AT mode choice for shopping purpose.

Table 5. Description of the rules according to the C5 for shopping purpose.

Rule ID	Node	Rule	Then	P%	S%	Po%
1*	2	(SAC = 1;2) AND (TSC = 2;3)	PT	73.33	28.95	39.48
2	4	(SAC = 3;4) AND (TSC = 2;3) AND (NF = 9)	W/B+R	100	2.63	2.63
3	5	(SAC = 3;4) AND (TSC = 2;3) AND (NF = 2;3;12)	W	100	7.89	7.89
4	6	(SAC = 3;4) AND (TSC = 2;3) AND (NF = 1;5;6;10;11)	PT	80	10.53	13.16
5*	8	(R1 = 3;4) AND (TSC = 1)	B	66.66	5.26	7.90
6	9	(R1 = 2) AND (TSC = 1)	W/B+R	100	2.63	2.63
7	10	(R1 = 1) AND (TSC = 1)	W	80	21.05	26.32

*rule excluded

PT=Public Transport; W=Walking; B=Biking; PV=Private Vehicle; R=Ridesourcing

3.4 Leisure purpose

Figure 6 shows the C5 model that predicts the use of active transport modes for travelling to the leisure destinations. The overall accuracy of this model is 84.52%. The tree identified eight associate rules with the use of AT modes (Table 6). Among these rules, four rules were identified as significant ($P\% = 75$, $S\% = 0.9$, and $Po\% = 1.2$). The model shows that the node 0 is split based on the variable of biking path availability and coverage (BAC) and this shows that the BAC is the single best variable to predict the use of AT modes by UTM employees to travel to leisure. C5 directs the involved BAC of "little or none" to the left, forming node 1; directs the BACs of "some bike paths or routes" and "many cycle paths, lanes, or routes forming a network" to the right, forming node 2. As shown by node 1, if an employee perceives the biking facilities in his/her living area as "little or none", the tree predicts that the employee's AT mode of choice for leisure is most likely to be walking (Rule 1: $P\% = 61.66$, $S\% = 44.05$, $Po\% = 71.44$ /non-significant).

Turning to the division of node 2, conditioned on the BACs of some bike paths or routes "and" many cycle paths, lanes, or routes forming a network "and" daily travel distance (DTD) to leisure, the C5 splits node 2 based on DTD and sends the DTD of 5 km and below to node 3; the DTD of greater than 5 km to node 9. As shown by node 9, if an employee's daily travel time to leisure is greater than 5 km, the tree predicts that the employee's AT mode choice to leisure is most likely to be a combination of walking and biking (Rule 6: $P\% = 75$, $S\% = 3.57$, $Po\% = 4.76$). Node 3 is further split in relation to neighborhood facilities within a walkable distance (NF). C5 sends the NFs of schools, childcare facilities, public transport shops, banks, leisure facilities, parks and other open spaces, and places of worship to node 4; and sends the NF of three or more facilities to node 5. At node 4, if an employee lives in a neighborhood with the abovementioned facilities in a walkable distance from his/her house, the tree predicts that the employee's AT mode choice to leisure is most likely to be biking (Rule 2: $P\% = 100$, $S\% = 11.90$, $Po\% = 11.90$). The node 5 is further split in relation to employees' race (R1). C5 sends the R1 of "Indian" to node 6; sends the R1 of "Chinese" to node 7; sends the R1 of "Malay" and "others" to node 8. As shown by node 6, if the race of the employee is Indian, the tree predicts that the employee's AT mode of choice for leisure is most likely to be public transport (Rule 3: $P\% = 100$, $S\% = 1.19$, $Po\% = 1.19$). At node 7, if the race of the employee is Chinese, the tree predicts that the employee's AT mode of choice for leisure is most likely to be biking (Rule 4: $P\% = 80$, $S\% = 4.76$, $Po\% = 5.95$). As shown by node 8, if the race of the employee is Malay or others, the tree predicts that the employee's AT mode of choice for leisure is most likely to be biking (Rule 5: $P\% = 100$, $S\% = 4.76$, $Po\% = 4.76$).

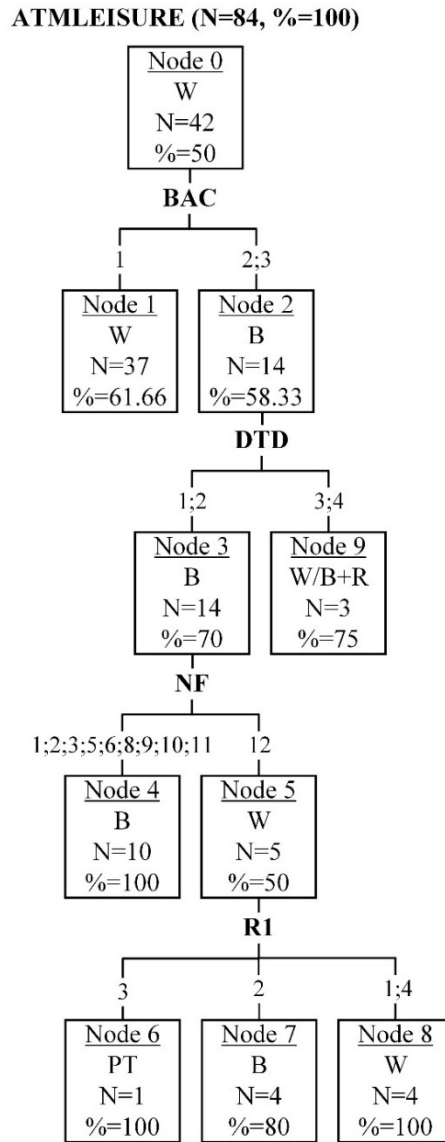


Figure 6. C5 model-predicting employee's AT mode choice for leisure purpose.

Table 6. Description of the rules according to the C5 for leisure purpose.

Rule ID	Node	Rule	Then	P%	S%	Po%
1*	1	BAC = 1	W	61.66	44.05	71.44
2	4	(NF = 1;2;3;5;6;8;9;10;11) AND (DTD = 1;2) AND (BAC = 1)	B	100	11.90	11.90
3*	6	(R1 = 3) AND (NF = 12) AND (DTD = 1;2) AND (BAC = 2;3)	PT	100	1.19	1.19
4	7	(R1 = 2) AND (NF = 12) AND (DTD = 1;2) AND (BAC = 2;3)	B	80	4.76	5.95
5	8	(R1 = 1;4) AND (NF = 12) AND (DTD = 1;2) AND (BAC = 2;3)	W	100		4.76
6	9	(DTD = 3;4) AND (BAC = 2;3)	W/B+R	75	3.57	4.76

*Rule excluded

PT=Public Transport; W=Walking; B=Biking; PV=Private Vehicle; R=Ridesourcing

4. Discussions

Analysis results showed that the usage of AT modes is strongly sensitive to several combinations of the living environment, sociodemographic, and trip attributes. The C5 models identified transit station conditions (TSC), sidewalk availability and coverage (SAC), and bike path availability and coverage (BAC) as the most important factors affecting the use of AT

modes for work, shopping, and leisure trips, respectively. These findings are in line with those of several studies that have shown the positive influence of the availability of well-conditioned transit stations on the use of public transport, particularly buses [96-98]. Several studies also noted the importance of sidewalk and bike path conditions in increasing people's willingness to use them [99-101].

The present study created several rules associated with the use of AT modes for work, shopping, and leisure trips. Each rule was a combination of factors from the living environment, sociodemographics, and trip characteristics. Concerning the work trips, the tree created five significant decision rules. According to the rules, two built environment factors, including SAC and TSC, were presented in all five rules, while another built environment factor, neighborhood type (N1), was only available in two rules. The frequent presence of SAC and TSC in all significant rules is in line with the findings of Heinen [102], who showed that the availability of comfortable walking and biking paths, as well as proximity to bus stops and busways, increase the share of commute trips involving any active travel to work. Concerning the factors of trip characteristics, the DTT was also available in all five significant rules, and its value was 10 min and below (value = 1). In general, therefore, it seems that the use of active modes is suitable for those employees that daily spend 10 min or less to get to work. This fixed value of DTT resulted in various outputs of AT. The association of other factors, including TSC, SAC, E1, N1, and VO with these decision rules, can possibly explain these differences. The association of other factors, including TSC, SAC, E1, N1, and VO with these decision rules, can possibly explain these differences. This association and the frequent presence of built environment factors, especially those that support walking and public transport (e.g., SAC and TSC), implies that these could be major factors, influencing the AT mode choice of employees to get to work. Taken together, these results suggest that the combination of SAC, DTT, and TSC can be a reliable predictor and perform better than other combinations (or a single factor) to predict the AT mode choice of employees to work. An important practical implication is that any improvement in sidewalk and transit station conditions might increase the usability of AT modes among employees to work. However, this action might not be sufficient for those who spend more than 10 minutes getting to work each day.

With regards to sociodemographic factors influencing the AT mode choice of employees to work, vehicle ownership (VO) and education level (E1) of employees were available in significant decision rules. The VO appeared in only one significant rule (rule 5), while the E1 appeared in three significant rules (rules 2-4). While a number of studies have identified vehicle ownership as a significant factor in AT participation [39, 103], it is surprising that vehicle ownership is not included in more significant rules to predict AT mode choice for work trips. However, this finding does not imply that vehicle ownership is not important for this purpose. It can therefore be assumed that when an employee decides to use an AT mode for work purposes, owning a motor vehicle at home does not impact his or her AT mode choice. Concerning the employees' education level, the values of E1 were primary, secondary, and diploma (values = 1-3). It is somewhat surprising that the lower levels of education are associated with the use of AT modes. These results differ from some published studies that note that people with higher education are much more inclined to use the AT modes [39, 104]. The association of other factors, including TSC, SAC, DTT, and N1 with these decision rules, can possibly explain this result. These findings suggest that higher AT levels are associated with lower employees' education level.

The C5 created five significant decision rules in regard to employees' AT mode choice to shopping. Concerning the built environment factors, TSC appeared in all five significant rules. The majority of employees used buses to get to shopping (39.5%) and possibly the design, location, and condition of the bus stops was of concern for the employees. This attention of employees to the bus stops can explain the frequent presence of TSC in the rules. TSC appeared in all the rules, but it does not mean that this factor is the only predictor of AT mode choice to get to shopping. Besides, SAC and NF appeared in the three rules. As the combination of SAC, TSC, and NF frequently appeared in the rules (three out of five), this combination could be an efficient predictor for AT mode choice of employees to shop. In the rules 2-4 of shopping trips, the values of SAC and TSC were "narrow sidewalks along all major streets" or "adequate sidewalks along all major streets" and "some bus stop shelters" or "widely available bus stop shelters", respectively. Combined with NF, the same values of SAC and TSC resulted in different AT modes, including a combination of walking/biking with ridesourcing, walking, and public transport. The variability in NF values can explain the variability in AT modes caused by these rules. It is therefore likely that such connections exist between the availability of sidewalks and transit stations in neighborhoods and higher rates of using walking and public transport to get to shopping by the employees.

R1, as a sociodemographic factor, appeared in two rules (6 and 7). Combined with R1, the same value of TSC (no shelters) resulted in different AT modes, including a combination of walking/biking with ride sourcing and walking. It can thus be suggested that when the value of a built environment factor such as TSC does not change, the likelihood of a combination of walking/biking and ride sourcing is higher for Chinese employees to get to shopping compared to other races; the likelihood of walking is higher for Malay employees to get to shopping compared to other races.

Concerning the purpose of leisure, the tree created four significant rules. According to the rules, a built environment factor, BAC, and a trip characteristics factor, DTD, appeared in all rules. The frequent presence of BAC in the decision rules was not surprising since biking was the second major mode of leisure activity. Additionally, the bike is regarded as a means of recreation in Malaysia. Thus, the importance of BAC for leisure trips makes sense. Another built environment factor, NF, was presented in three rules. Furthermore, the race was influential on rules 4 and 5. While the values of NF, DTD, and BAC do not change in these rules, the different values of R1 resulted in various AT outputs, including biking and walking to leisure activities. It can thus be suggested that when the built environment and trip characteristic factors such as NF, BAC, and DTD are fixed, the likelihood of biking is higher for Chinese employees compared to other races, and the likelihood of walking is higher for Malay and others (e.g., foreigners and Indians). Because this combination appeared in all of the rules, the combination of DTD and BAC can be an effective predictor of employees' AT mode choice for leisure. This frequent presence also shows that the AT mode choice of employees to leisure is mostly influenced by DTD and BAC.

5. Limitations

Within the scope of this research, although the present study produced notable insights on the use of AT modes among university employees, there are some notable limitations. First, the study population was purposefully kept limited to those employees that used active modes for the purposes of work, shopping, and leisure three times per week. This ensures that the target

population includes frequent users of AT, and these users are developing a consistent pattern of AT. Second, the present study was conducted among university employees in a developing country where the overall condition of active transport infrastructure, including sidewalks, bike lanes, and bus stations, is undesirable. Therefore, caution must be applied as the results of proposed models may not be transferable to developed countries. Third, the sample size used in this study might have reduced the generalizability of the results. Therefore, the findings of this study may not be generalizable to employees in universities of varying sizes with diverse demographics, and participation bias may have influenced the results. Finally, the self-report data of AT modes' usage was used in this study. Further studies can complement self-report with trip observations.

6. Conclusions

This study set out to investigate the AT mode choice of employees for different trip purposes, including work, shopping, and leisure. To this end, a sample of administrative university employees was selected to be studied. This study then analyzed the data using the C5 algorithm. The data analyzed was presented in the form of decision trees and decision rules. The most obvious finding to emerge from this study is that the TSC, SAC, and BAC are the most important factors affecting the use of AT modes by employees for work, shopping, and leisure trips, respectively. The second major finding was that combinations of (a) SAC, DTT, and TSC to get to work, (b) SAC, TSC, and NF to get to shopping, and (c) DTD, and BAC to get to leisure are the best predictors for AT mode choice of university employees. The results of this research support the idea that it is necessary to simultaneously use strategies that improve the built environment conditions, reduce the travel time and distance with regards to the trip purposes, and consider the sociodemographic attributes of employees in order to achieve a high level of active transport among the employees. The current findings also enhance our understanding of the AT modes' characteristics and their influencing factors. This research will serve as a base for future studies and further research to understand the underlying cause of complex employees' mode choice behavior with the living environment and sociodemographic change.

Conflicts of Interest

The authors declare no conflict of interest.

References

Uncategorized References

- [1]. Public Health Agency of Canada. What is active transportation? 2014 [cited 2019 4th May]; Available from: <http://www.phac-aspc.gc.ca/hp-ps/hl-mvs/pa-ap/at-ta-eng.php>.
- [2]. Health in All Policies Task Force, 2014-2018 Active Transportation Action Report. 2018, California Health in All Policies (HiAP) Task Force: California. pp. 1-6.
- [3]. Voss, C. (2018). 1 - Public Health Benefits of Active Transportation. in Children's Active Transportation, Eds, Larouche, R., Elsevier. pp. 1-20.
- [4]. Gordon, C. (2018). 3 - Economic Benefits of Active Transportation. in Children's Active Transportation, Eds, Larouche, R., Elsevier. pp. 39-52.
- [5]. Hong, A. (2018). 2 - Environmental Benefits of Active Transportation. in Children's Active Transportation, Eds, Larouche, R., Elsevier. pp. 21-38.

- [6]. Pérez, K.; Olabarria, M.; Rojas-Rueda, D.; Santamariña-Rubio, E.; Borrell, C. Nieuwenhuijsen, M. (2017). The health and economic benefits of active transport policies in Barcelona. *Journal of Transport & Health*, 4, 316-324. <https://doi.org/10.1016/j.jth.2017.01.001>.
- [7]. Brown, V.; Diamedi, B.Z.; Moodie, M.; Veerman, J.L. Carter, R. (2016). A systematic review of economic analyses of active transport interventions that include physical activity benefits. *Transport Policy*, 45, 190-208. <https://doi.org/10.1016/j.tranpol.2015.10.003>.
- [8]. Kornas, K.; Bornbaum, C.; Bushey, C. Rosella, L. (2017). Exploring active transportation investments and associated benefits for municipal budgets: a scoping review. *Transport Reviews*, 37, 465-487. 10.1080/01441647.2016.1252446.
- [9]. Arsenio, E. Ribeiro, P. (2015). The Economic Assessment of Health Benefits of Active Transport. in Sustainable Urban Transport Emerald Group Publishing Limited. pp. 1-22.
- [10]. Kent, J.L. (2014). Carsharing as active transport: What are the potential health benefits? *Journal of Transport & Health*, 1, 54-62. <https://doi.org/10.1016/j.jth.2013.07.003>.
- [11]. Tainio, M.; Goodman, A. Woodcock, J. (2017). 1977 - Developing New Method to Estimate Active Mode (walking, cycling) Related Health Benefits for Department for Transport (DfT) Transport Analysis Guidance. *Journal of Transport & Health*, 5, S29-S30. <https://doi.org/10.1016/j.jth.2017.05.313>.
- [12]. Lu, W.; Scott, D.M. Dalumpines, R. (2018). Understanding bike share cyclist route choice using GPS data: Comparing dominant routes and shortest paths. *Journal of Transport Geography*, 71, 172-181. <https://doi.org/10.1016/j.jtrangeo.2018.07.012>.
- [13]. Yang, Y.; Xu, Y.; Rodriguez, D.A.; Michael, Y. Zhang, H. (2018). Active travel, public transportation use, and daily transport among older adults: the association of built environment. *Journal of Transport & Health*, 9, 288-298.
- [14]. Duncan, S.; White, K.; Mavoa, S.; Stewart, T.; Hinckson, E. Schofield, G. (2016). Active transport, physical activity, and distance between home and school in children and adolescents. *Journal of Physical Activity and Health*, 13, 447-453.
- [15]. Mackey, D.C.; Perkins, A.D.; Tai, K.H.; Sims-Gould, J. McKay, H.A. (2018). Men on the Move: A Randomized Controlled Feasibility Trial of a Scalable, Choice-Based Physical Activity and Active Transportation Intervention for Older Men. *Journal of aging and physical activity*, 1-45.
- [16]. Rothman, L.; Macpherson, A.K.; Howard, A.; Parkin, P.C.; Richmond, S.A. Birken, C.S. (2016). Direct observations of active school transportation and stroller use in kindergarten children. *Preventive medicine reports*, 4, 558-562.
- [17]. Roberts, J.D.; Rodkey, L.; Ray, R. Saelens, B.E. (2018). Do not forget about public transportation: analysis of the association of active transportation to school among Washington, DC area children with parental perceived built environment measures. *Journal of physical activity and health*, 15, 474-482.
- [18]. Cubbin, C.; Kim, Y. Lee, R.E. (2017). California Women with Young Children Use Active Transportation Most Often in Very Unsafe Neighborhoods Regardless of SES. *Journal of Transport & Health*, 5, S8.
- [19]. Smith, L.; Norgate, S.H.; Cherrett, T.; Davies, N.; Winstanley, C. Harding, M. (2015). Walking school buses as a form of active transportation for children—a review of the evidence. *Journal of school health*, 85, 197-210.
- [20]. Bhosale, J.; Duncan, S. Schofield, G. (2017). Intergenerational change in children's independent mobility and active transport in New Zealand children and parents. *Journal of Transport & Health*, 7, 247-255.
- [21]. Jones, R.A.; Blackburn, N.E.; Woods, C.; Byrne, M.; van Nassau, F. Tully, M.A. (2019). Interventions promoting active transport to school in children: A systematic review and meta-analysis. *Preventive medicine*.
- [22]. MIDA, Malaysian Automotive Industry, in Automotive Industry. 2018, Malaysian Investment Development Authority: Kuala Lumpur. pp. 1-12.
- [23]. Nielsen. Rising middle class will drive global automotive demand in the coming two years. 2014 [cited 2019 4th May]; Available from: <https://www.nielsen.com/my/en/press-room/2014/rising-middle-class-will-drive-global-automotive-demand.html>.
- [24]. Bopp, M.; Sims, D.; Colgan, J.; Rovniak, L.; Matthews, S.A. Poole, E. (2016). An examination of workplace influences on active commuting in a sample of university employees. *Journal of Public Health Management and Practice*, 22, 387-391.

- [25]. Zhou, J. (2012). Sustainable commute in a car-dominant city: Factors affecting alternative mode choices among university students. *Transportation Research Part A: Policy and Practice*, 46, 1013-1029. 10.1016/j.tra.2012.04.001.
- [26]. Zhou, J. (2014). From better understandings to proactive actions: Housing location and commuting mode choices among university students. *Transport Policy*, 33, 166-175. 10.1016/j.tranpol.2014.03.004.
- [27]. Yang, M.; Wang, W.; Chen, X.; Wang, W.; Xu, R.Gu, T. (2010). Modeling destination choice behavior incorporating spatial factors, individual sociodemographics, and travel mode. *Journal of Transportation Engineering*, 136, 800-810. 10.1061/ASCE/TE.1943-5436.0000140.
- [28]. Bernetti, G.; Longo, G.; Tomasella, L. Violin, A. (2008). Sociodemographic Groups and Mode Choice in a Middle-Sized European City. *Transportation Research Record: Journal of the Transportation Research Board*, 2067, 17-25. 10.3141/2067-03.
- [29]. Cheng, L.; Chen, X.; Yang, S.; Wang, H.Wu, J. (2016). Modeling Mode Choice of Low-Income Commuters with Sociodemographics, Activity Attributes, and Latent Attitudinal Variables: Case Study in Fushun, China. *Transportation Research Record: Journal of the Transportation Research Board*, 2581, 27-36. 10.3141/2581-04.
- [30]. Lind, H.B.; Nordfjærn, T.; Jørgensen, S.H.Rundmo, T. (2015). The value-belief-norm theory, personal norms and sustainable travel mode choice in urban areas. *Journal of Environmental Psychology*, 44, 119-125. 10.1016/j.jenvp.2015.06.001.
- [31]. Kim, S.Ulfarsson, G.F. (2004). Travel mode choice of the elderly effects of personal, household, neighborhood, and trip characteristics. *Transportation Research Record: Journal of the Transportation Research Board*, 1894, 117-126.
- [32]. Scheiner, J. (2010). Interrelations between travel mode choice and trip distance: trends in Germany 1976–2002. *Journal of Transport Geography*, 18, 75-84. 10.1016/j.jtrangeo.2009.01.001.
- [33]. Manaugh, K.El-Geneidy, A.M. (2013). Does distance matter? Exploring the links among values, motivations, home location, and satisfaction in walking trips. *Transportation Research Part A: Policy and Practice*, 50, 198-208. <https://doi.org/10.1016/j.tra.2013.01.044>.
- [34]. Ermagun, A.Levinson, D. (2017). “Transit makes you short”: On health impact assessment of transportation and the built environment. *Journal of Transport & Health*, 4, 373-387. <https://doi.org/10.1016/j.jth.2016.08.004>.
- [35]. Du, Y.; Deng, F.Liao, F. (2019). A model framework for discovering the spatio-temporal usage patterns of public free-floating bike-sharing system. *Transportation Research Part C: Emerging Technologies*, 103, 39-55. <https://doi.org/10.1016/j.trc.2019.04.006>.
- [36]. Loo, B.P.Y.Siiba, A. (2019). Active transport in Africa and beyond: towards a strategic framework. *Transport Reviews*, 39, 181-203. 10.1080/01441647.2018.1442889.
- [37]. Kim, S.Ulfarsson, G.F. (2008). Curbing automobile use for sustainable transportation: analysis of mode choice on short home-based trips. *Transportation*, 35, 723-737.
- [38]. Chatman, D.G. (2003). How density and mixed uses at the workplace affect personal commercial travel and commute mode choice. *Transportation Research Record*, 1831, 193-201.
- [39]. Barnett, A.; Sit, C.H.; Mellecker, R.R.Cerin, E. (2018). Associations of socio-demographic, perceived environmental, social and psychological factors with active travel in Hong Kong adolescents: The iHealt (H) cross-sectional study. *Journal of Transport & Health*.
- [40]. Oliver, M.; Badland, H.; Mavoa, S.; Witten, K.; Kearns, R.; Ellaway, A.; Hinckson, E.; Mackay, L.Schluter, P.J. (2014). Environmental and socio-demographic associates of children’s active transport to school: a cross-sectional investigation from the URBAN study. *International Journal of Behavioral Nutrition and Physical Activity*, 11, 70.
- [41]. Mandic, S.; de la Barra, S.L.; Bengoechea, E.G.; Stevens, E.; Flaherty, C.; Moore, A.; Middlemiss, M.; Williams, J.Skidmore, P. (2015). Personal, social and environmental correlates of active transport to school among adolescents in Otago, New Zealand. *Journal of science and medicine in sport*, 18, 432-437.
- [42]. Scheepers, E.; Wendel-Vos, W.; van Kempen, E.; Panis, L.I.; Maas, J.; Stipdonk, H.; Moerman, M.; den Hertog, F.; Staatsen, B.van Wesemael, P. (2013). Personal and environmental characteristics associated with choice of active transport modes versus car use for different trip purposes of trips up to 7.5 kilometers in The Netherlands. *PloS one*, 8, e73105.

- [43]. McCormack, G.R.; Shiell, A.; Giles-Corti, B.; Begg, S.; Veerman, J.L.; Geelhoed, E.; Amarasinghe, A. Emery, J.H. (2012). The association between sidewalk length and walking for different purposes in established neighborhoods. *International Journal of Behavioral Nutrition and Physical Activity*, 9, 92. [10.1186/1479-5868-9-92](https://doi.org/10.1186/1479-5868-9-92).
- [44]. Carlson, C.; Aytur, S.; Gardner, K. Rogers, S. (2012). Complexity in built environment, health, and destination walking: a neighborhood-scale analysis. *Journal of Urban Health*, 89, 270-284.
- [45]. Pikora, T.J.; Giles-Corti, B.; Knuiman, M.W.; Bull, F.C.; Jamrozik, K. Donovan, R.J. (2006). Neighborhood environmental factors correlated with walking near home: using SPACES. *Medicine and science in sports and exercise*, 38, 708-714.
- [46]. Ma, L. Dill, J. (2015). Associations between the objective and perceived built environment and bicycling for transportation. *Journal of Transport & Health*, 2, 248-255. <https://doi.org/10.1016/j.jth.2015.03.002>.
- [47]. Zahabi, S.A.H.; Chang, A.; Miranda-Moreno, L.F. Patterson, Z. (2016). Exploring the link between the neighborhood typologies, bicycle infrastructure and commuting cycling over time and the potential impact on commuter GHG emissions. *Transportation Research Part D: Transport and Environment*, 47, 89-103. <https://doi.org/10.1016/j.trd.2016.05.008>.
- [48]. Braun, L.M.; Rodriguez, D.A.; Cole-Hunter, T.; Ambros, A.; Donaire-Gonzalez, D.; Jerrett, M.; Mendez, M.A.; Nieuwenhuijsen, M.J. de Nazelle, A. (2016). Short-term planning and policy interventions to promote cycling in urban centers: Findings from a commute mode choice analysis in Barcelona, Spain. *Transportation Research Part A: Policy and Practice*, 89, 164-183. <https://doi.org/10.1016/j.tra.2016.05.007>.
- [49]. Mertens, L.; Compennolle, S.; Deforche, B.; Mackenbach, J.D.; Lakerveld, J.; Brug, J.; Roda, C.; Feuillet, T.; Oppert, J.-M.; Glonti, K.; Rutter, H.; Bardos, H.; De Bourdeaudhuij, I. Van Dyck, D. (2017). Built environmental correlates of cycling for transport across Europe. *Health & Place*, 44, 35-42. <https://doi.org/10.1016/j.healthplace.2017.01.007>.
- [50]. Aziz, H.M.A.; Nagle, N.N.; Morton, A.M.; Hilliard, M.R.; White, D.A. Stewart, R.N. (2017). Exploring the impact of walk-bike infrastructure, safety perception, and built-environment on active transportation mode choice: a random parameter model using New York City commuter data. *Transportation*, 45, 1207-1229. [10.1007/s11116-017-9760-8](https://doi.org/10.1007/s11116-017-9760-8).
- [51]. Christian, H.; Knuiman, M.; Divitini, M.; Foster, S.; Hooper, P.; Boruff, B.; Bull, F. Giles-Corti, B. (2017). A longitudinal analysis of the influence of the neighborhood environment on recreational walking within the neighborhood: results from RESIDE. *Environmental health perspectives*, 125, 077009.
- [52]. Sung, H.; Lee, S. Jung, S. (2014). Identifying the Relationship between the Objectively Measured Built Environment and Walking Activity in the High-Density and Transit-Oriented City, Seoul, Korea. *Environment and Planning B: Planning and Design*, 41, 637-660. [10.1068/b39116](https://doi.org/10.1068/b39116).
- [53]. Turrell, G.; Haynes, M.; Wilson, L.-A. Giles-Corti, B. (2013). Can the built environment reduce health inequalities? A study of neighbourhood socioeconomic disadvantage and walking for transport. *Health & Place*, 19, 89-98. <https://doi.org/10.1016/j.healthplace.2012.10.008>.
- [54]. Munshi, T. (2016). Built environment and mode choice relationship for commute travel in the city of Rajkot, India. *Transportation Research Part D: Transport and Environment*, 44, 239-253. <https://doi.org/10.1016/j.trd.2015.12.005>.
- [55]. Christiansen, L.B.; Cerin, E.; Badland, H.; Kerr, J.; Davey, R.; Troelsen, J.; van Dyck, D.; Mitáš, J.; Schofield, G.; Sugiyama, T.; Salvo, D.; Sarmiento, O.L.; Reis, R.; Adams, M.; Frank, L. Sallis, J.F. (2016). International comparisons of the associations between objective measures of the built environment and transport-related walking and cycling: IPEN adult study. *Journal of Transport & Health*, 3, 467-478. <https://doi.org/10.1016/j.jth.2016.02.010>.
- [56]. Oliver, L.; Schuurman, N.; Hall, A. Hayes, M. (2011). Assessing the influence of the built environment on physical activity for utility and recreation in suburban metro Vancouver. *BMC Public Health*, 11, 959. [10.1186/1471-2458-11-959](https://doi.org/10.1186/1471-2458-11-959).
- [57]. Wasfi, R.A.; Dasgupta, K.; Eluru, N. Ross, N.A. (2016). Exposure to walkable neighbourhoods in urban areas increases utilitarian walking: Longitudinal study of Canadians. *Journal of Transport & Health*, 3, 440-447. <https://doi.org/10.1016/j.jth.2015.08.001>.
- [58]. Sundquist, K.; Eriksson, U.; Kawakami, N.; Skog, L.; Ohlsson, H. Arvidsson, D. (2011). Neighborhood walkability, physical activity, and walking behavior: The Swedish

- Neighborhood and Physical Activity (SNAP) study. *Social Science & Medicine*, 72, 1266-1273. <https://doi.org/10.1016/j.socscimed.2011.03.004>.
- [59]. Van Dyck, D.; Cardon, G.; Deforche, B.; Sallis, J.F.; Owen, N. De Bourdeaudhuij, I. (2010). Neighborhood SES and walkability are related to physical activity behavior in Belgian adults. *Preventive Medicine*, 50, S74-S79. <https://doi.org/10.1016/j.ypmed.2009.07.027>.
- [60]. Van Dyck, D.; Deforche, B.; Cardon, G. De Bourdeaudhuij, I. (2009). Neighbourhood walkability and its particular importance for adults with a preference for passive transport. *Health & Place*, 15, 496-504. <https://doi.org/10.1016/j.healthplace.2008.08.010>.
- [61]. Koohsari, M.J.; Owen, N.; Cole, R.; Mavoa, S.; Oka, K.; Hanibuchi, T. Sugiyama, T. (2017). Built environmental factors and adults' travel behaviors: Role of street layout and local destinations. *Preventive Medicine*, 96, 124-128. <https://doi.org/10.1016/j.ypmed.2016.12.021>.
- [62]. Riva, M.; Gauvin, L.; Apparicio, P. Brodeur, J.-M. (2009). Disentangling the relative influence of built and socioeconomic environments on walking: The contribution of areas homogenous along exposures of interest. *Social Science & Medicine*, 69, 1296-1305. <https://doi.org/10.1016/j.socscimed.2009.07.019>.
- [63]. Thielman, J.; Rosella, L.; Copes, R.; Lebenbaum, M. Manson, H. (2015). Neighborhood walkability: Differential associations with self-reported transport walking and leisure-time physical activity in Canadian towns and cities of all sizes. *Preventive Medicine*, 77, 174-180. <https://doi.org/10.1016/j.ypmed.2015.05.011>.
- [64]. Kelley, E.A.; Kandula, N.R.; Kanaya, A.M. Yen, I.H. (2016). Neighborhood Walkability and Walking for Transport Among South Asians in the MASALA Study. *Journal of Physical Activity and Health*, 13, 514-519. 10.1123/jpah.2015-0266.
- [65]. Jack, E. McCormack, G.R. (2014). The associations between objectively-determined and self-reported urban form characteristics and neighborhood-based walking in adults. *International Journal of Behavioral Nutrition and Physical Activity*, 11, 71. 10.1186/1479-5868-11-71.
- [66]. Davison, L.; Ahern, A. Hine, J. (2015). Travel, transport and energy implications of university-related student travel: a case study approach. *Transportation Research Part D: Transport and Environment*, 38, 27-40. 10.1016/j.trd.2015.04.028.
- [67]. Namgung, M. Akar, G. (2015). Influences of neighborhood characteristics and personal attitudes on university commuters' public transit use. *Journal of the Transportation Research Board*, 2500, 93-101. 10.3141/2500-11.
- [68]. Jomnonkwao, S.; Sangphong, O.; Khampirat, B.; Siridhara, S. Ratanavaraha, V. (2016). Public transport promotion policy on campus: evidence from Suranaree University in Thailand. *Public Transport*, 8, 185-203. 10.1007/s12469-016-0122-2.
- [69]. Tezcan, H.O. (2016). Potential of carpooling among unfamiliar users: case of undergraduate students at Istanbul Technical University. *Journal of Urban Planning and Development*, 142, 1-11. 10.1061/(asce)up.1943-5444.0000283.
- [70]. Zhan, G.; Yan, X.; Zhu, S. Wang, Y. (2016). Using hierarchical tree-based regression model to examine university student travel frequency and mode choice patterns in China. *Transport Policy*, 45, 55-65. 10.1016/j.tranpol.2015.09.006.
- [71]. Zhou, J. (2016). Proactive sustainable university transportation: marginal effects, intrinsic values, and university students' mode choice. *International Journal of Sustainable Transportation*, 10, 815-824. 10.1080/15568318.2016.1159357.
- [72]. Swiers, R.; Pritchard, C. Gee, I. (2017). A cross sectional survey of attitudes, behaviours, barriers and motivators to cycling in University students. *Journal of Transport & Health*, 6, 379-385. 10.1016/j.jth.2017.07.005.
- [73]. Daisy, N.S.; Hafezi, M.H.; Liu, L. Millward, H. (2018). Understanding and modeling the activity-travel behavior of university commuters at a large Canadian university. *Journal of Urban Planning and Development*, 144, 1-10. 10.1061/(asce)up.1943-5444.0000442.
- [74]. Hafezi, M.H.; Sultana Daisy, N.; Liu, L. Millward, H. (2018). Daily activity and travel sequences of students, faculty and staff at a large Canadian university. *Transportation Planning and Technology*, 41, 536-556. 10.1080/03081060.2018.1469286.
- [75]. Mitra, R. Nash, S. (2018). Can the built environment explain gender gap in cycling? An exploration of university students' travel behavior in Toronto, Canada. *International Journal of Sustainable Transportation*, .doi.org/10.1080/15568318.2018.1449919, 1-10. 10.1080/15568318.2018.1449919.

- [76]. Nguyen-Phuoc, D.Q.; Amoh-Gyimah, R.; Tran, A.T.P.Phan, C.T. (2018). Mode choice among university students to school in Danang, Vietnam. *Travel Behaviour and Society*, 13, 1-10. 10.1016/j.tbs.2018.05.003.
- [77]. Sims, D.; Bopp, M.Wilson, O.W.A. (2018). Examining influences on active travel by sex among college students. *Journal of Transport & Health*, 9, 73-82. 10.1016/j.jth.2018.05.009.
- [78]. Tan, S.B.; Zegras, P.C.; Wilhelm, E.Arcaya, M.C. (2018). Evaluating the effects of active morning commutes on students' overall daily walking activity in Singapore: Do walkers walk more? *Journal of Transport & Health*, 8, 220-243. 10.1016/j.jth.2018.01.002.
- [79]. Peer, S. (2019). To bike or not to bike? – Evidence from a university relocation. *Transportation Research Part D: Transport and Environment*, 70, 49-69. 10.1016/j.trd.2019.03.003.
- [80]. Nash, S.Mitra, R. (2019). University students' transportation patterns, and the role of neighbourhood types and attitudes. *Journal of Transport Geography*, 76, 200-211. 10.1016/j.jtrangeo.2019.03.013.
- [81]. Wang, X.; Khattak, A.J.Son, S. (2013). What can be learned from analyzing university student travel demand? *Transportation Research Record: Journal of the Transportation Research Board*, 2322, 129-137. 10.3141/2322-14.
- [82]. Delmelle, E.M.Delmelle, E.C. (2012). Exploring spatio-temporal commuting patterns in a university environment. *Transport Policy*, 21, 1-9. 10.1016/j.tranpol.2011.12.007.
- [83]. Chen, X. (2012). Statistical and activity-based modeling of university student travel behavior. *Transportation Planning and Technology*, 35, 591-610. 10.1080/03081060.2012.701818.
- [84]. Lundberg, B.Weber, J. (2014). Non-motorized transport and university populations: an analysis of connectivity and network perceptions. *Journal of Transport Geography*, 39, 165-178. 10.1016/j.jtrangeo.2014.07.002.
- [85]. Zhou, J. (2013). Study of Employee Carsharing on the University Campus. *Journal of Urban Planning and Development*, 139, 301-310. 10.1061/(asce)up.1943-5444.0000153.
- [86]. Wang, D.Liu, Y., Factors influencing public transport use: a study of university commuters' travel and mode choice behaviours, in State of Australian Cities Conference. 2015.
- [87]. Shannon, T.; Giles-Corti, B.; Pikora, T.; Bulsara, M.; Shilton, T.Bull, F. (2006). Active commuting in a university setting: Assessing commuting habits and potential for modal change. *Transport Policy*, 13, 240-253.
- [88]. Uttley, J.Lovelace, R. (2016). Cycling promotion schemes and long-term behavioural change: A case study from the University of Sheffield. *Case Studies on Transport Policy*, 4, 133-142. 10.1016/j.cstp.2016.01.001.
- [89]. Griselda, L.Joaquín, A. (2012). Using decision trees to extract decision rules from police reports on road accidents. *Procedia-Social and Behavioral Sciences*, 53, 106-114.
- [90]. Washington, S.; Jean, W.Guensler, R. (1997). Binary recursive partitioning method for modeling hot-stabilized emissions from motor vehicles. *Journal of the Transportation Research Board*, 96-105.
- [91]. Lemon, S.C.; Roy, J.; Clark, M.A.; Friedmann, P.D.Rakowski, W. (2003). Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Annals of behavioral medicine*, 26, 172-181.
- [92]. Quinaln, J.R. (1993). C4. 5: programs for machine learning. Kuafmann, Morgan San Mateo, California.
- [93]. O'Connor, A. (2015). An Analysis of the Predictive Capability of C5.0 and Chaid Decision Trees and Bayes Net in the Classification of fatal Traffic Accidents in the UK.
- [94]. Zhou, G.; Wang, L.; Wang, D.Reichle, S. (2010). Integration of GIS and Data Mining Technology to Enhance the Pavement Management Decision Making. *Journal of Transportation Engineering*, 136, 332-341. 10.1061/共ASCE天TE.1943-5436.0000092.
- [95]. Pande, A.Abdel-Aty, M. (2009). Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool. *Safety Science*, 47, 145-154. 10.1016/j.ssci.2007.12.001.
- [96]. Wong, R.C.P.; Szeto, W.Y.; Yang, L.; Li, Y.C.Wong, S.C. (2017). Elderly users' level of satisfaction with public transport services in a high-density and transit-oriented city. *Journal of Transport & Health*, 7, 209-217. 10.1016/j.jth.2017.10.004.

- [97]. Miao, Q.; Welch, E.W.Sriraj, P.S. (2019). Extreme weather, public transport ridership and moderating effect of bus stop shelters. *Journal of Transport Geography*, 74, 125-133. 10.1016/j.jtrangeo.2018.11.007.
- [98]. Chakour, V.Eluru, N. (2016). Examining the influence of stop level infrastructure and built environment on bus ridership in Montreal. *Journal of Transport Geography*, 51, 205-217. 10.1016/j.jtrangeo.2016.01.007.
- [99]. Mateo-Babiano, I. (2016). Pedestrian's needs matter: Examining Manila's walking environment. *Transport Policy*, 45, 107-115. 10.1016/j.tranpol.2015.09.008.
- [100]. Nasar, J.L.; Holloman, C.Abdulkarim, D. (2015). Street characteristics to encourage children to walk. *Transportation Research Part A: Policy and Practice*, 72, 62-70. 10.1016/j.tra.2014.12.004.
- [101]. El-Assi, W.; Salah Mahmoud, M.Nurul Habib, K. (2017). Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. *Transportation*, 44, 589-613. 10.1007/s11116-015-9669-z.
- [102]. Heinen, E.; Panter, J.; Mackett, R.Ogilvie, D. (2015). Changes in mode of travel to work: a natural experimental study of new transport infrastructure. *Int J Behav Nutr Phys Act*, 12, 81. 10.1186/s12966-015-0239-8.
- [103]. Guliani, A.; Mitra, R.; Buliung, R.N.; Larsen, K.Faulkner, G.E.J. (2015). Gender-based differences in school travel mode choice behaviour: Examining the relationship between the neighbourhood environment and perceived traffic safety. *Journal of Transport & Health*, 2, 502-511. 10.1016/j.jth.2015.08.008.
- [104]. Winters, M.; Brauer, M.; Setton, E.M.Teschke, K. (2010). Built environment influences on healthy transportation choices: bicycling versus driving. *Journal of urban health*, 87, 969-993.



© 2021 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).