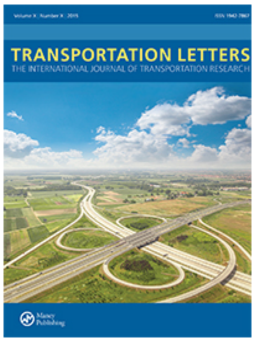

Taylor & Francis
Taylor & Francis Group

Transportation Letters
The International Journal of Transportation Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ytrl20

Application of machine learning with a surrogate model to explore seniors' daily activity patterns

Yiling Deng

TRANSPORTATION LETTERS
https://doi.org/10.1080/19427867.2021.1969169

Application of machine learning with a surrogate model to explore seniors' daily activity patterns

Yiling Deng

School of Design and Architecture, Zhejiang University of Technology, Hangzhou, Zhejiang Province, China

ABSTRACT

Investigating seniors' daily activity patterns (DAPs) is essential to understand their activity-travel needs. Although some studies have applied machine learning (ML) to derive DAPs, few of them have sought to improve the interpretability of ML. This study aims to predict and interpret seniors' DAPs in the Chinese context by using ML with a surrogate model. First, a boosted C5.0 algorithm was employed to model seniors' DAPs, which provided more accurate predictions than the multinomial logit (MNL) model. Second, a rule-based C5.0 algorithm was used as a surrogate model to approximate the prediction function of the boosted C5.0 algorithm and to provide insight into the underlying decision processes in the boosted C5.0 algorithm. The results show that retired men are most likely to lack out-of-home activities. A good residential built environment, especially good walkability and public transit accessibility, increases seniors' out-of-home activities. This study provides recommendations for increasing seniors' mobility.

KEYWORDS

Daily activity pattern; seniors; machine learning; surrogate model; C5.0 algorithm; multinomial logit model

Introduction

Population aging has become a notable and common demographic phenomenon in most countries. According to World Population Prospects 2019 released by the Population Division of the United Nations, the number of people aged 60 and above will double by 2050, rising from 1.0 billion in 2020 to 1.4 billion in 2030 and 2.1 billion in 2050 globally (World Population Prospect 2019). As the most populous country, China has been witnessing prominent demographic aging because of increased longevity and declining birth rates. The share of the aging population in China is 17.9% in 2018 and is projected to reach 25% in 2030 (China's Population Aging Trend Forecast Report 2018). The rapid increase of the aging population is supposed to dramatically influence urban and transportation systems. Considering the special activity-travel needs of seniors is important for urban planners and transportation operators to create a better travel environment for seniors.

Since trips are derivatives of out-of-home activities, examining seniors' daily activity patterns (DAPs) rather than trip characteristics can provide a full understanding of the underlying decision processes of their travel behavior. However, previous studies have focused more on seniors' trip characteristics, such as trip frequency (Cheng et al. 2019; Szeto et al. 2017; Böcker, Van Amen, and Helbich 2017; Hahn et al. 2016; Choo, Sohn, and Park 2016; Feng 2017), trip mode (Szeto et al. 2017; Böcker, Van Amen, and Helbich 2017; Zhang et al. 2019; Soltani et al. 2018), and distance traveled (Feng 2017), and less on seniors' DAPs. Methodologically, most studies used logit models to investigate individuals' DAPs (Dianat, Habib, and Miller 2020; Kristoffersson, Berglund, and Algers 2020). Machine learning (ML), such as the boosted C5.0 algorithm, has received increasing attention owing to its predictive power. However, the lack of interpretability is a serious drawback of the boosted C5.0 algorithm as well as other ML algorithms. It is worth investigating how to make boosted C5.0 algorithm as well as other ML algorithms as interpretable as econometric or rule-based techniques, which would facilitate broader applications of ML in modeling DAPs and other activity-travel behaviors.

Seniors' DAPs

Researchers and practitioners have developed different activity-based models (e.g., constraint-based models, econometric models, and computational process models) to provide the realistic behavioral representations of the underlying decision processes of travel behavior (Rasouli and Timmermans 2014). The DAP choice or scheduling module is the cornerstone of the activity-based model and the key feature that distinguishes the activity-based model from the trip-based model (Auld and Mohammadian 2009). A growing body of literature has investigated individuals' DAPs (Dianat, Habib, and Miller 2020; Kristoffersson, Berglund, and Algers 2020; Millward, Hafezi, and Daisy 2019; Hafezi, Liu, and Millward 2018; Allahviranloo and Recker 2013). However, few studies have assessed seniors' DAPs. Two exceptions are the studies of Habib and Hui (2017) and Lai et al. (2019). They both used MDCEV models.

Seniors' DAPs have not yet been fully explored in the Chinese context. The special household structure, early retirement age, and social and cultural contexts may differentiate Chinese seniors' DAPs and their influences from those in Western countries. Although Lai et al.'s study was based in Hong Kong, they only analyzed older couples in two-member households. In addition, the sociodemographics and built environment of mainland Chinese cities are very different from those of Hong Kong. For example, Hong Kong has a higher population density and lower car ownership. More research is needed to understand Chinese seniors' DAPs.

CONTACT Yiling Deng yiling@zjut.edu.cn School of Design and Architecture, Zhejiang University of Technology, Hangzhou, Zhejiang Province, China

Methods to estimate DAPs

Modeling DAPs is a discrete choice problem that can be solved with logit models or ML. Logit models have been most commonly used, including multinomial logit (MNL) models (Dianat, Habib, and Miller 2020), nested logit (NL) models (Kristoffersson, Berglund, and Algers 2020), and multiple discrete continuous extreme value (MDCEV) models, which are a marriage of a discrete choice model to determine participation or not and a linear regression model to determine activity duration (Habib and Hui 2017; Lai et al. 2019; Shamshiripour and Samimi 2019). A few studies applied ML to derive DAPs, such as random forests (Hafezi, Liu, and Millward 2018) and support vector machines (Allahviranloo and Recker 2013). MNL, NL, and MDCEV models all presuppose a linear model structure, while ML allows for more flexible model structures, which reduces model incompatibilities with empirical data. Thus, ML can perform better than logit models in prediction (Zhao et al. 2020; Martín-Baos, García-Ródenas, and Rodriguez-Benitez 2021).

ML also has limitations and its application in modeling DAPs should be improved in at least two ways. First and foremost, ML has often been criticized for being 'black-box', which is a major obstacle to its use in many practical applications. Few studies have sought to improve the interpretability of ML (Koushik, Manoj, and Nezamuddin 2020). Training a surrogate model is an effective way to interpret decision processes in ML (Molnar 2020). To the best of our knowledge, the surrogate model has not been used to explain ML in activity-travel behavior studies. Second, decision trees are the most widely used algorithms in ML and are the base classifiers for ensemble learning, such as boosting and bagging. Among decision tree algorithms, CHAID (Chi-squared Automatic Interaction Detection) (Rashidi and Mohammadian 2011; Marquet and Miralles-Guasch 2014; Zhan et al. 2016; Fasihozaman Langerudi, Rashidi, and Mohammadian 2016) and CART (Classification and Regression Trees) (Hafezi, Liu, and Millward 2019; Pitombo, de Souza, and Lindner 2017; Pitombo, Kawamoto, and Sousa 2011; Pitombo, Sousa, and Filipe 2009) have been often used in activity-travel behavior studies. C5.0, a thorough decision tree algorithm developed by Quinlan (2014), has been rarely used, although it works well for most types of problems directly out of the box. C5.0 differs from CART and CHAID in the split mechanism, pruning process, and definition of purity index. Compared with CART, C5.0 is more accurate because it is not restricted to making binary trees. Compared with CHAID, the trees of C5.0 are usually smaller, which makes the results more interpretable and generalizable. C5.0 has a boosted version, which is faster than AdaBoost and random forest because it uses many fewer trees. Fernandez-Delgado et al. assessed 179 classifiers from 17 families using 121 data sets and found that C5.0 was among the top 5 algorithms that produced the best results (Fernandez-Delgado et al. 2014).

Objectives of this research

This study aims to model seniors' DAPs to better understand their activity-travel needs. We followed the framework of discrete choice activity-based models, such as DaySim (Bowman and Ben-Akiva 2001) and CT-RAMP (Davidson et al. 2010). These widely used prototype activity-based models usually generate DAPs first and then generate tours as well as trips for individuals. Rather than using a set of frequency choice models, we formulated a DAP choice model making integrated decisions on activity agenda, tour purposes, and tour complexity for seniors. We used an ML technique – boosted C5.0 algorithm to achieve a more accurate prediction of DAPs and an interpretable surrogate model – rule-based C5.0 algorithm to reveal the inner processes in the ML technique.

The contributions of this paper lie on the followings:

1. to apply the boosted C5.0 algorithm to predict seniors' DAPs;
2. to apply the rule-based C5.0 algorithm as a surrogate model to approximate and explain the boosted C5.0 algorithm; and
3. to reveal the factors influencing seniors' DAPs in the Chinese context.

The remainder of this paper is organized as follows. Section 2 presents data and variables. Section 3 introduces the methodology. Section 4 analyzes the results. Section 5 discusses the conclusions.

Data and variables

Data

This study was based on the Nanjing Household Travel Survey carried out in 2012. The data included one-day travel diaries and household and personal characteristics of 5,974 individuals from 2,007 households. In this study, seniors were defined as those aged 50 and over. The definition of seniors is controversial. It is usually associated with changes in the life course, such as retirement, or with the age at which seniors receive pensions or medical benefits. Most studies in Western countries defined seniors as those aged over 60 or 65, which reflects the legal retirement age in these societies. The reason for choosing the age of 50 as the selection criterion for seniors in this study was that the statutory retirement age in China is 60 for men and 55 or 50 for women. Seniors were further divided into two main age cohorts: age 50–60 as the younger senior cohort (transitional cohort) and age 60 or older as the senior cohort. The inclusion of the younger senior cohort can reflect the impacts of life transitions on DAPs as well as provide information on the activity and travel

Table 1. List of variables

Variable	Coding	Type	Choice
Daily activity pattern	DAP	Categorical	SN, SM, SN-SN, H, CN, CM, SM-SN, SM-SM
Work status	Job	Binary	Working, Retired
Age	Age	Binary	Younger senior (50–59), Senior (≥ 60)
Gender	Gen	Binary	Female, Male
Education level	Edu	Binary	Low, High
Bus pass ownership	BP	Binary	Yes, No
Driver's license ownership	DL	Binary	Yes, No
Household car ownership	Car	Ordered	0, 1, ≥2
Household bicycle ownership	Bike	Ordered	0, 1, 2, ≥3
Household income	Inc	Ordered	Low, Mid, High
Family composition	FC	Categorical	Elderly, Adult, Extended
Residential built environment	BE	Categorical	1, 2, 3, 4

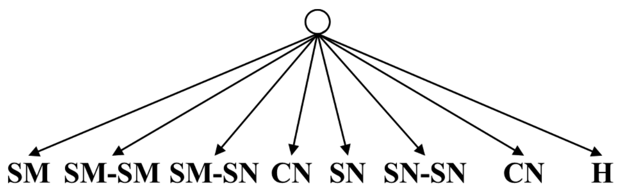

Figure 1. Choice structure of DAPs.

needs of the future senior cohort. The same definition of seniors has been used in several studies in China (Feng 2017; Feng et al. 2013) and a similar classification and naming of senior cohorts have been used in studies in the U.S. (Moniruzzaman et al. 2013), Canada (Moniruzzaman and Páez 2016), and Singapore (Hou 2019).

After data cleaning, 1589 seniors from 968 households were screened out. All variables used in the model are presented in Table 1. The dependent variable is seniors' DAPs. The independent variables include sociodemographics and built environment factors.

DAPs

DAPs were identified by the presence of tours. In practice, different activity-based models use different DAP coding schemes, which have implications for overall model system design. In this study, we proposed a realistic but simplified DAP coding scheme, which can be predicted in a single model.

Eight precoded categories of trip purpose were reported by older adults: work, personal business, shopping, recreation/entertainment, social, medical appointment, return, and other. These activities were grouped into mandatory activities, including work, and nonmandatory activities, including maintenance activities, i.e., personal business, shopping, and medical appointment, and discretionary activities, i.e., recreation/entertainment, social, and other. We only distinguished mandatory and nonmandatory purposes because of the small amount of data.

Tours were defined as home-to-home loops. The tour purpose was named based on the activity purpose. If a tour included two or more activities with different purposes, the tour purpose was assigned based on the primary activity. A tour with only one stop was called a simple tour. A tour with more than one stop was called a complex tour. Considering the tour purpose and number of stops, tours were classified into four categories: simple mandatory tour (SM), simple nonmandatory tour (SN), complex mandatory tour (CM), and complex nonmandatory tour (CN). A complex mandatory tour can also include nonmandatory activities, although not necessarily. However, a complex nonmandatory tour can only involve nonmandatory activities.

DAPs were simply the combinations of tours without considering sequences. By analyzing the data, thirteen categories of DAPs were detected. Five categories of DAPs were rather small, together accounting for 3% of all DAPs. These DAPs were grouped into their proximate DAPs as follows: twenty-one CN-SN were grouped into CN, nineteen SN-SN-SN were grouped into SN-SN, seven CM-SM were grouped into CM, three SM-SM-SO were grouped into SM-SO, and one CN-CN were grouped into CN. After processing, eight categories of DAPs were remained. The codes and total numbers of these DAPs are as follows: SN (649), SM (412), SN-SN (207), H (119), CN (70), CM (54), SM-SN (35), and SM-SM (43). Note that H (stay home all day) was also a category of DAP, which is quite common among seniors. The choice structure of DAPs is illustrated in Figure 1.

Sociodemographics

Sociodemographics included gender, age, work status, education level, driver's license ownership, bus pass ownership, household income, household car ownership, household bicycle ownership, and family composition. Age, gender, driver's license ownership, and bus pass ownership all had two options and were treated as nominal variables. Other variables with multiple choices were processed as follows and then treated as nominal variables: the options for household income were merged from seven to three (i.e., low, middle, or high), for work status from nine to two (i.e., working or retired), and for education level from four to two (i.e., low or high). Household car ownership was limited to three choices (i.e., 0, 1, 2 +). Household bicycle ownership was limited to four choices (i.e., 0, 1, 2, 3+).

Chinese household structure is unique. For economic and cultural reasons, seniors usually live with their married or unmarried adult children. Three categories of family composition were discerned as follows and then treated as a categorical variable: elderly family (single or couple seniors), adult family (seniors living with adult children), and extended family (seniors living with children and grandchildren). The proportions of seniors who lived in elderly families, adult families, and extended families were approximately 16%, 63%, and 21%, respectively.

Built environment

The '5 Ds' introduced by Ewing and Cervero (i.e., density, diversity, design, destination accessibility, and distance to transit) are the classic variables for characterizing the residential built environment (Ewing and Cervero2010). Residential density (Res), employment density (Emp), Walk Score (WS), intersection density (Int), distance to Central Business District (CBD), distance to metro (MRT), and bus stop density (Bus) were adopted to characterize residential built environment by following the '5 Ds'. Residential density and employment density were measured within a 1 km radius around the household location to reflect the density index. The Walk Score was obtained through the application programming interface (API) of the walkscore.com website, which is closely related to the diversity index and the design index (Walk Score Methodology 2011). In addition to the Walk Score, intersection density was used to access the design index. The destination accessibility index was measured using distance to CBD. The distance to transit index was measured by distance to metro and bus stop density. Distance to metro took the form of an exponential distance decay function as $1 - e^{-d}$, in which d was the distance to the nearest metro stop (km). By using this function, distance to metro was scaled to range from 0 to 1 and was exponentially rather than linearly increased. Bus stop density was measured within a 1 km radius around the household location. Since bus stops are very dense in Nanjing and most household locations can reach a bus stop within 300 m, we used bus stop density rather than distance to a bus stop.

Household locations with a homogeneous built environment were grouped into clusters. After testing, using the clusters rather than including all the continuous built environment variables helps reduce overfitting and improve the generalizability of the model. It also makes the model easier to explain. Z-score standardization was conducted for all the built environment variables. After standardization, all the variables follow a distribution with a mean of 0 and standard deviation of 1 and have an equal effect on clustering results. K-means, one of the most commonly used clustering algorithms, was used to generate clusters. K-means requires the number of clusters (k value) as an input in advance. Figure 2 is a scree plot used to determine the appropriate k value.

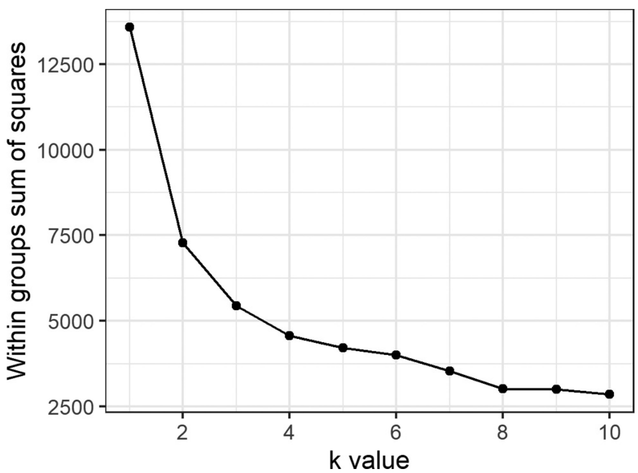

Figure 2. Scree plot of clustering.

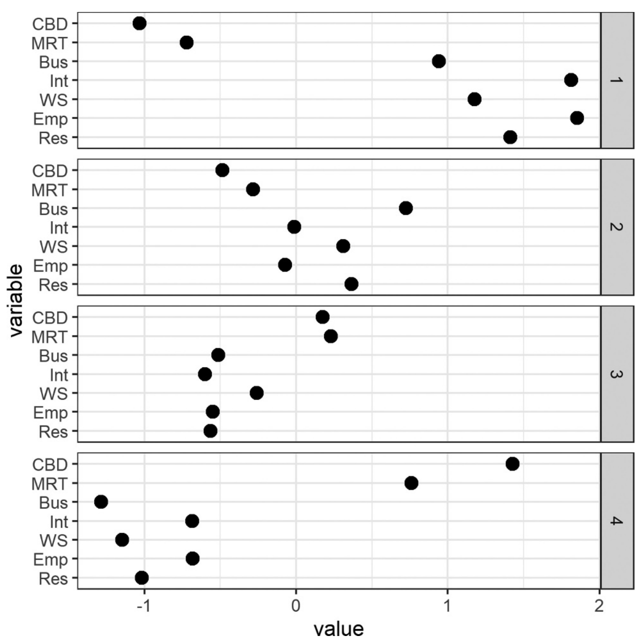

Figure 3. Average characteristics of the clusters.

In general, as the k value increases, the within-groups sum of squares decreases, indicating smaller differences within groups. When it exceeds four, the curve becomes flat. Since a greater number of clusters would complicate cluster results and reduce the interpretability of cluster results, we chose four as the final k value. Figure 3 is a Cleveland dot plot showing the mean characteristics of the four clusters. From the first cluster to the fourth cluster in sequence, the built environment gradually changes from dense, mixed, and highly accessible land use to sparse, monotonous, and lowly accessible land use.

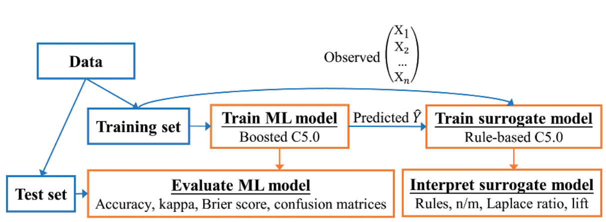

Figure 4. Conceptual diagram.

Methodology

In this section, we first introduced the C5.0 algorithm, which is the basis for the boosted C5.0 algorithm and rule-based C5.0 algorithm introduced later. Then, we proposed the boosted C5.0 algorithm to estimate DAPs and the rule-based C5.0 algorithm as a surrogate model to approximate and interpret the predictions of the boosted C5.0 algorithm. The conceptual diagram is shown in Figure 4.

C5.0 algorithm

The C5.0 algorithm is a thorough decision tree algorithm that splits observed data into homogeneous groups based on recursive partitioning methods (Quinlan 2014). The C5.0 algorithm uses information entropy as the purity index of a data set. In data set D, assuming the proportion of observations in class k is p_k ($k = 1, 2, \ldots, |y|$), the information entropy of D ($Ent(D)$) is the summation of purity indices for all the existing classes, as Eq. (1) shows.

$$\mathrm{Ent(D)} = -\sum_{k=1}^{|y|} p_k \log_2 p_k \tag{1}$$

The smaller the value of $Ent(D)$, the greater the purity of D. With Eq. (1), the purity index for every node in a decision tree can be calculated. Assuming a discrete independent variable a has V possible values $\{a^1, a^2, \ldots, a^V\}$, if a is used to split D, V branch nodes are generated. The branch node v contains all the samples where a equals a^v. The information entropy of D^v can be computed using Eq. (1). Since the sample sizes in the branch nodes are different, each branch node is weighted by the sample size proportion ($|D^v|/|D|$). Eq. (2) shows the information gain for the split on the discrete independent variable a.

$$\mathrm{Gain(D, a)} = \mathrm{Ent(D)} - \sum_{v=1}^{V} \frac{D^v}{D} \mathrm{Ent(D^v)} \tag{2}$$

The C5.0 algorithm splits a node by exhaustively searching over all the possible splits across all the independent variables to obtain the split that can maximize the information gain. Figure 5 shows the pseudocode of the C5.0 algorithm.

To overcome overfitting, the C5.0 algorithm conducts a final global pruning procedure. The overall strategy is to postprune the tree: First, grow a large tree that overfits the training data, then remove branches and nodes having little effect on the classification errors. The C5.0 algorithm can automatically use reasonable defaults for pruning and allows manual tuning.

Boosted C5.0 algorithm

Boosting is rooted in the notion that using a combination of numeric learners with complementary strengths and weaknesses is much more accurate than using any one of the learners alone. The boosted C5.0 algorithm is similar to the AdaBoost algorithm in that the models are fit sequentially, with each iteration adjusting the case weights according to the accuracy of a sample's prediction (Quinlan 2014). The notable differences are as follows: First, the boosted C5.0 algorithm creates successor trees of the same size as the initial tree; second, when combing the constituent trees, each tree computes the confidence values for each class, and an average of these values is calculated without using stage

Input: Train dataset $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$.
Attribute set $A = \{a_1, a_2, \ldots, a_d\}$.
Method: Function $TreeGenerate(D, A)$
1. create node N
2. **if** samples in D are all of the same class C, **then**
3. label N as a leaf node with class C, **return**
4. **end if**
5. **if** $A = \emptyset$ **or** the values of A in D are the same, **then**
6. label N as a leaf node with the majority class in D, **return**
7. **end if**
8. find the best splitting attribute a_* that maximizes the information gain
9. **for** every value a_*^v of a_* **do**
10. create a branch for N, let D_v be the set of data in D with a_* equal to a_*^v
11. **if** $D_v = \emptyset$, **then**
12. attach a leaf node labeled with the majority class in D to N
13. **else**
14. attach the node returned by $TreeGenerate(D_v, A \backslash \{a_*\})$
15. **end if**
16. **end for**

Output: a C5.0 decision tree

Figure 5. The pseudocode of the C5.0 algorithm.

Input: Train dataset $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$.
Iteration K.
Method:
1. Let each sample have the same starting weight $(1/m)$
2. **for** $k = 1\ to\ K$ **do**
4. Fit a C5.0 classifier using the weighted samples
4. Coerce the tree to have the same number of terminal nodes as the initial tree
5. Compute the misclassification error of the tree
6. Update the sample weights to give more weight to the incorrectly predicted samples and less weight to the correctly predicted samples
7. **if** the sum of the weights for the misclassified samples is less than 0.1, **exit**
8. **if** the average weight of misclassified samples is greater than 0.5, **exit**
9. **end for**
10. **Output**: a boosted C5.0 decision tree

Figure 6. The pseudocode of the boosted C5.0 algorithm.

weights. The class with the largest confidence value is the final choice; third, boosting is automatically stopped if the result is either highly effective (e.g., the sum of the weights for the misclassified samples is less than 0.1) or highly ineffective (e.g., the average weight of misclassified samples is greater than 0.5). Figure 6 shows the pseudocode of the boosted C5.0 algorithm.

There are extremely unequal instances in different categories of DAPs. The numbers of single and simple tour DAPs are much greater than those of multiple or complex tour DAPs. Like most other ML algorithms, the boosted C5.0 algorithm is sensitive to unbalanced data as unbalanced data bias the prediction toward the majority classes. To address this problem, we tested several pre-processing techniques such as oversampling, under-sampling, and SMOTE (Chawla et al. 2002), and found that oversampling with square root weighting achieved the best performance, especially in improving the ability to discern minority DAPs and matching the total rates of different categories of DAPs. For square root weighting, the weights of the largest class (SN in this study) were set to 1. Other classes were weighted by Eq. (3).

$$w_x = \sqrt{n_{max}}/\sqrt{n_x} \tag{3}$$

where w stands for weight, n is the class size, x refers to a specific class and max refers to the largest class.

The performance of the boosted C5.0 algorithm is evaluated by accuracy, Kappa, and the Brier score. Accuracy is the percentage of instances that are correctly classified among all instances. Accuracy ranges from 0 (completely predicted wrong) to 1 (completely predicted correct). Kappa measures the agreement between the predicted and observed outcomes. Kappa is normalized at the baseline of random chance and ranges from −1 (complete disagreement) to 1 (complete agreement). The Brier score measures the performance of the model on probability predictions (Brier 1950). Eq. (4) shows the calculation of the Brier score, which is the average squared deviation between the predicted probabilities for a set of alternatives and their outcomes.

$$Brier\ score\ = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2 \tag{4}$$

where p_i is the predicted probability and o_i is the observed value of the instance i (0 if negative and 1 if positive). N is the number of instances. A lower Brier score represents higher accuracy.

Surrogate model and rule-based C5.0 algorithm

The basic idea of the surrogate model (also known as the behavioral model or black-box model) is to find an interpretable model to approximate the behavior of ML as closely as possible. By interpreting the surrogate model, the behavior of ML can be revealed. The surrogate model is very intuitive and has been proven to be an effective way to interpret the decision process of ML (Molnar 2020). Any interpretable econometric model (e.g., the linear regression, decision tree, or rule lists) can be used as a surrogate model as long as it can mimic the ML prediction function well. Moreover, the choice of ML technique and the choice of the surrogate model are independent, which means that a surrogate model can be applied to interpret different ML techniques. This feature is important when different ML techniques need to be tried to achieve more accurate predictions.

The surrogate model is trained on the original training data, but with the ML predictions as the outcome. The performance of the surrogate model can be measured by the error rate for the classification problem. A low error rate means that the surrogate model approximates the ML prediction function well. Of course, we cannot expect that the surrogate model can fully capture the behavior of ML. If this is the case, then it is better to throw away the ML model and use a surrogate model instead.

In this study, the ML technique – boosted C5.0 algorithm is not very interpretable, so we used a rule-based C5.0 algorithm as a surrogate model to approximate the behavior of the boosted C5.0 algorithm. Rules can define the terminal nodes more simply than a tree because one or more conditions can be discarded without changing the subset of observations belonging to the node. The rule-based C5.0 algorithm is based on the C5.0 algorithm we elaborated in the former section and uses the following scheme for deriving rules. First, after creating the initial tree via the C5.0 algorithm, each path through the tree is collapsed into an individual rule. Second, the individual rules are simplified by pruning, and the number of constituent rules is reduced by a global procedure.

Three statistics (i.e., n/m, Laplace ratio, and lift) are used to evaluate the performance of each rule. N is the weighted number of training cases covered by the rule, and m is the weighted number of training cases covered by the rule but do not belong to the class. The Laplace ratio is a confidence estimator of the rule and is calculated as $(n - m + 1)/(n + 2)$. Laplace ratio ranges between 0 and 1. A value close to 1 implies a good confidence of the rule, while a value close to 0 implies the opposite. Lift measures the performance of the rule in predicting cases compared to random chance. Lift ranges between 0 and $+\infty$. A value close to 1 implies that the result is close to random chance and the rule is meaningless, while a value greater than 1 implies that the rule performs well in prediction.

Results

Model results

Following the common procedure for estimating ML models, we split the data into training (70%) and test (30%) sets and used stratified sampling to ensure consistent representation of the

Table 2. Confusion matrices for the boosted C5.0 algorithm.

Training Data	SM	SM-SM	SM-SN	CM	SN	SN-SN	CN	H	Observed Total	Prediction Error
SM	**227**	2	1	3	50	2	2	2	289	1%
SM-SM	5	**20**	0	0	6	0	0	0	31	6%
SM-SN	2	0	**21**	0	0	2	0	0	25	4%
CM	4	1	0	**33**	0	0	0	0	38	5%
SN	40	6	3	2	**342**	24	17	21	455	3%
SN-SN	7	1	1	0	26	**105**	3	2	145	4%
CN	1	0	0	0	8	3	**37**	0	49	20%
H	7	3	0	2	11	3	0	**58**	84	1%
Predict Total	293	33	26	40	443	139	59	83	1116	Avg 6%
Test Data										
SM	**88**	5	4	7	11	1	3	4	123	1%
SM-SM	6	**2**	0	1	1	2	0	0	12	0%
SM-SN	3	0	**2**	1	3	0	1	0	10	20%
CM	9	0	0	**3**	1	0	0	3	16	19%
SN	10	3	2	0	**143**	22	6	8	194	9%
SN-SN	2	0	0	0	31	**22**	4	3	62	13%
CN	1	1	0	0	8	5	**5**	1	21	10%
H	5	1	0	1	13	2	0	**13**	35	9%
Predict Total	124	12	8	13	211	54	19	32	473	Avg 10%

Table 3. Rules generated by the surrogate model.

DAP	No.	n/m	Laplace ratio	Lift	Rule
SM	1	21/3.1	0.82	4.0	Job = Working, Edu = High, DL = Yes, BE = 3
	2	496.4/295.1	0.41	2.0	Job = Working
SM-SM	3	20.3/8	0.60	9.0	Inc = Mid, Job = Working, Edu = Low, DL = Yes, FC = Adult, BE = 3
	4	13.2/5.8	0.55	8.3	Inc = Mid, Job = Working, Edu = High, DL = No, BE = 2
	5	30.4/18	0.41	6.2	Age = Younger senior, Inc = Low, Job = Working, BP = Yes, Car = 0, BE = 1
SM-SN	6	2.7	0.79	12.9	Gen = Female, Inc = Mid, Job = Working, FC = Extended, BE = 2
	7	6.8/1.4	0.73	11.9	Gen = Male, Inc = Mid, Job = Working, Edu = Low, DL = No, BP = No, BE = 2
	8	3.5/0.8	0.68	11.1	Gen = Female, Age = Younger senior, Job = Working, BP = Yes, FC = Elderly, BE = 4
	9	11.6/6.2	0.47	7.7	Inc = Low, Job = Retired, BP = Yes, Car = 0, BE = 4
CM	10	4.9/0.8	0.74	9.4	Inc = Mid, Job = Working, Edu = Low, DL = Yes, FC = Elderly, BE = 2
	11	12.9/4.7	0.62	7.9	Age = Younger senior, Inc = Mid, Job = Working, Edu = Low, DL = Yes, Car = 0, BE = 2
	12	10/5.9	0.43	5.4	Age = Younger senior, Inc = Low, Job = Retired, DL = Yes, FC = Adult, BE in {1, 2, 3}
SN	13	654.6/402.7	0.39	1.5	Job = Retired
SN-SN	14	9.2/1.4	0.79	5.5	Age = Senior, Inc in {Mid, High}, Job = Retired, Edu = High, FC in {Adult, Extended}, BE = 1
	15	21/13.2	0.38	2.7	Age = Younger senior, Inc in {Low, Mid}, Job = Retired, Edu = Low, DL = No, Car = 1, BE = 2
	16	548.9/449.7	0.18	1.3	Gen = Female
CN	17	1.9	0.75	8.9	Age = Younger senior, Job = Retired, DL = Yes, FC = Elderly
	18	8.3/2.5	0.66	7.8	Age = Younger senior, Inc = High, Job = Retired, Edu = Low, BE = 2
	19	6.8/3	0.55	6.6	Age = Senior, Inc = Mid, Edu = High, FC = Elderly, BE in {1, 2}
H	20	16.6/6.3	0.61	5.6	Age = Senior, Job = Retired, BP = No, Bike in {2, 3}, BE = 4
	21	11.9/4.6	0.60	5.5	Inc = Mid, BP = Y, Car = 0, Bike in {0, 1}, BE = 4
	22	12.2/4.9	0.59	5.4	Inc in {Mid, High}, Job = Retired, Car in {1, 2}, Bike in {0, 1}, FC = Extended, BE = 4
	23	8.8/4.4	0.50	4.6	Inc = Low, Job = Retired, Edu = High, BE in {1, 2}
	24	11.3/6.9	0.41	3.7	Inc = Low, Job = Working, Car = 1, BE in {1, 2, 3}

dependent variable in the training and test sets. We examined three training–test splits (60–40%, 70–30%, and 80–20%) and found that the model performance of the boosted C5.0 algorithm and the common structure of the rules of the C5.0 algorithm were stable across all examined training–test splits.

The boosted C5.0 algorithm ran eight iterations and took only a few seconds on the laptop. It achieved an accuracy of 0.755, a kappa of 0.670, and a Brier score of 21.03 on the training set and an accuracy of 0.588, a kappa of 0.430, and Brier score of 21.28 on the test set. Table 2 shows the confusion matrices comparing the predicted DAPs with the observed DAPs. For each class of DAPs, the prediction error (the absolute difference between the predicted total and the observed total divided by the observed total) was calculated. The diagonal cells contain 75.5% of the DAPs in the training set and 58.8% of the DAPs in the test set. These DAPs were correctly predicted by the model. Although large values were observed in the off-diagonal cells (cells that were incorrectly predicted) between SN and SM, SN and SN-SN, SN and CN, and SN and H, these off-diagonal cells were well balanced. The number of observed and predicted DAPs in each class of DAPs is close, which is important for travel demand modeling.

Model interpretation

We estimated a surrogate model using the rule-based C5.0 algorithm. The training set was the same as the original training set, except that the predicted DAPs of the boosted C5.0 algorithm were used as the outcome instead of the observed DAPs. The error rate of the surrogate model is 18.7%, indicating that the surrogate model approximates the predictions of the boosted C5.0 algorithm well. It generates 24 rules, as shown in Table 3.

Two rules (rules 1 and 2) explain SM – the second largest class of DAPs. Work status is the most important variable in both rules. Three rules result in SM-SM (rules 3–5). Work status, household income, and residential built environment appear in all three rules. Education level and driver's license ownership rank second as they appear in two rules (rules 3 and 4). Seniors with middle or low household incomes are more likely to choose SM-SM than seniors with high household incomes (rules 3–5). Four rules explain SM-SN (rules 6–9). Work status and residential built environment are the most influential variables, followed by household income, gender, and bus pass ownership. For working seniors with middle household income and living in a good built environment, if they are female and live in an extended family (rule 6), or if they are male and do not have a bus pass and driver's license (rule 7), they are more inclined to choose SM-SN. For seniors who have a bus pass and live in a bad built environment, if they are female and live in an elderly family (rule 8), or if their household income is low and they do not own a car (rule 9), they are more inclined to choose SM-SN. CM is determined by three rules (rules 10–12). Driver's license ownership, household income, work status, and residential built environment play a role in all three rules. Education level, family composition, and age play a role in two of the rules. In all three rules, seniors all have a driver's license and live in a good built environment, their household income is low (rules 10 and 11) or middle (rule 12), and they do not need to care for grandchildren as they live in an elderly family (rule 10) or an adult family (rule 12).

SN is the most common DAP for retired seniors (rule 13). Three rules result in SN-SN (rules 14–16). Age, household income, work status, residential built environment, and education level all appear in two of the three rules. Retired seniors (aged over 60) with middle or high household income and high education level meet rule 14. Retired younger seniors (aged 50 to 59) with middle or low household income and low education level meet rule 15. Although the household and personal attributes of seniors are opposite between rules 14 and 15, both cohorts live in a good built environment. Three rules determine CN (rules 17–19). Age is the most important variable and work status, household income, residential built environment, and family composition are the second most important variables. For retired younger seniors (aged 50 to 59), if they own a driver's license and live in an elderly family (rule 17), or if they have a high household income and live in a good built environment (rule 18), they are more likely to choose CN. For seniors (aged over 60), if they have a middle household income, live in an elderly family, and reside in a good built environment (rule 19), they are also more inclined to choose CN.

Table 4. Coefficients of the MNL model

Variable	SM	SM-SM	SM-SN	CM	SN	SN-SN	CN
Intercept	4.24***	2.58*	3.83***	2.62*	2.16**	–	–
Job (Ref: Working)							
Retired	-2.49***	-1.88***	-0.95*	-2.28***	–	–	–
Age (Ref: Younger senior)							
Senior	-1.54***	-1.57**	-2.17***	-1.50**	–	0.63*	–
Gen (Ref: Male)							
Female	-0.60*	-2.10***	-1.42***	–	–	–	–
Edu (Ref: Low)							
High	-0.96**	-0.79*	-1.74***	-1.57***	-1.27***	-1.09**	-0.9*
BP (Ref: Yes)							
No	-0.64*	-1.33**	–	-0.84*	–	-1.08**	–
DL (Ref: Yes)							
No	–	–	–	–	0.74*	1.00*	1.09*
Bike (Ref: 0)							
1	–	–	–	–	-1.06*	-1.04*	-1.13*
2	–	–	–	–	-1.07*	-1.04*	–
3	–	–	–	–	-1.20*	–	–
Inc (Ref: Low)							
Mid	–	–	0.90*	0.90*	–	–	0.68*
High	1.23*	–	1.63*	2.48***	–	–	1.12*
FC (Ref: 1)							
2	–	-1.37**	-1.87***	-1.12*	-0.83*	-0.85*	-1.38**
≥3	-2.02***	–	-3.33***	-2.61***	-1.12*	-1.05*	-1.68**
BE (Ref: 1)							
3	–	–	–	–	–	–	0.66*
4	-0.92*	-1.34*	–	–	-0.91*	-0.93*	–
Summary statistics							
Log-likelihood	-1755						
AIC	3776						

***, **, * denote significance at the 0.001, 0.01, and 0.1 levels, respectively.

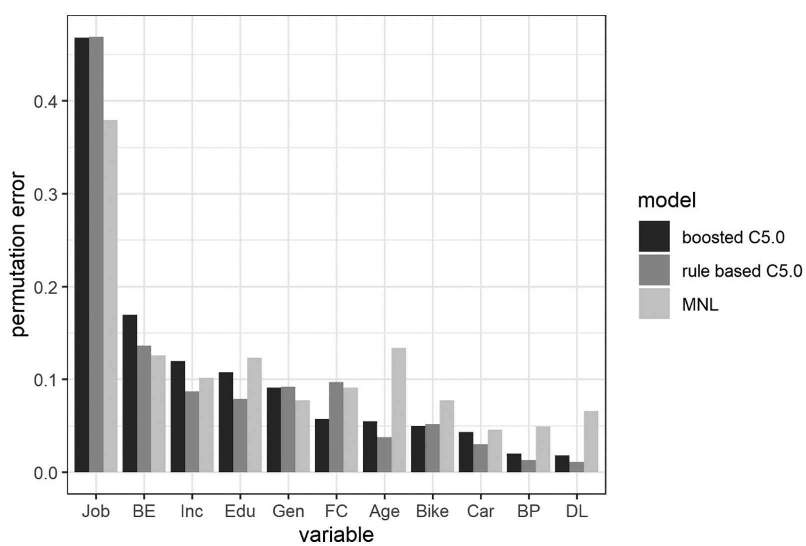

Figure 7. Variable importance of the C5.0 algorithms and MNL model.

H is defined by five rules (rules 20–24). Residential built environment plays a predominant impact on H, followed by household income and work status. These five rules are divided into two branches according to the residential built environment. For seniors living in the worst built environments, if they do not have a bike or car (rule 21), or if they own a car and have a middle or high household income but live in an extended family (rule 22), they stay at home. These two rules reveal that if seniors do not undertake out-of-home activities, it may be due to an unfriendly built environment, lack of transport options, or the responsibility of caring for grandchildren. For seniors who live in a good built environment but have a low household income, they also stay at home (rules 23 and 24).

Model comparison

We used a MNL model as a benchmark method to evaluate the performance of the boosted C5.0 algorithm. The MNL model was estimated with H as the reference. Table 4 shows the coefficients of the MNL model. It achieves an accuracy of 0.637, a kappa of 0.499,

Table 5. Confusion matrices for the MNL model.

Training Data	SM	SM-SM	SM-SN	CM	SN	SN-SN	CN	H	Observed Total	Prediction Error
SM	**224**	8	2	8	38	5	1	3	289	8%
SM-SM	5	**16**	0	2	6	0	1	1	31	29%
SM-SN	16	1	**4**	0	3	0	0	1	25	68%
CM	17	2	0	**13**	5	0	0	1	38	18%
SN	22	8	0	4	**344**	38	20	19	455	9%
SN-SN	6	2	1	1	58	**61**	13	3	145	24%
CN	3	0	1	3	16	4	**21**	1	49	22%
H	20	3	0	0	27	2	4	**28**	84	32%
Predict Total	313	40	8	31	497	110	60	57	1116	Avg 26%
Test Data										
SM	**86**	6	3	11	13	2	1	1	123	9%
SM-SM	8	**1**	0	1	1	0	1	0	12	33%
SM-SN	3	2	**0**	0	5	0	0	0	10	50%
CM	9	1	0	**2**	3	0	0	1	16	6%
SN	19	1	1	1	**127**	25	8	12	194	3%
SN-SN	1	3	1	1	29	**21**	6	0	62	11%
CN	1	2	0	1	10	4	**2**	1	21	14%
H	7	0	0	0	12	3	0	**13**	35	20%
Predict Total	134	16	5	17	200	55	18	28	473	Avg 18%

and Brier score of 21.22 on the training set and an accuracy of 0.533, a kappa of 0.340, and Brier score of 21.79 on the test set. Both for class predictions and probability predictions, the boosted C5.0 algorithm performs better than the MNL model on the training and test sets, respectively.

Although a one-to-one comparison between the rules of the C5.0 algorithm and the coefficients of the MNL model is impossible, we compared the variable importance of the C5.0 algorithms and the MNL model to shed light on the effects of the variables in these models. The variable importance was computed using the permutation feature importance measure first proposed by Breiman (Breiman 2001). We permuted the values of each variable one by one and compared the increase in the model prediction error. One variable is more important than another variable if shuffling the values of the variable increases the model prediction error more than shuffling the values of the other variable because in this case, the model relies more on the variable than the other variable for prediction. The results are shown in Figure 7. We found that the variable importance presented a similar pattern in the boosted C5.0 and rule-based C5.0 algorithms. Generally, Job, BE, Inc, and Edu are the most important variables (permutation error > 0.1), followed by Gen, FC, Age, and Bike (permutation error > 0.05), and Car, BP, and DL are the least important variables. Most of the variables have similar importance in the MNL model and the C5.0 algorithms, and a few variables with lower importance (i.e., Age, BP, and DL) differ more in importance in the MNL model and the C5.0 algorithms, with the former having larger values than the latter.

Table 5 shows the confusion matrices. The average prediction errors of the MNL model were larger than those of the boosted C5.0 algorithm (26% vs. 6% for the training set and 18% vs. 10% for the test set). The same misclassification classes exist in the MNL model as in the boosted C5.0 algorithm, but the former has a larger proportion of misclassifications. The result also confirms the difficulty of characterizing SM-SM, SM-SN, and CM, since these classes of DAPs were rarely observed.

Conclusions and discussion

The aging population has become a pervasive societal phenomenon all over the world. Modeling seniors' DAPs is important for a comprehensive understanding of their activity-travel behavior. In this study, we formulated a DAP choice model and estimated it using the boosted C5.0 algorithm, which produced more accurate predictions than the MNL model. We then applied the rule-based C5.0 algorithm as a surrogate model to interpret the boosted C5.0 algorithm. From a methodological point of view, the idea of the surrogate model was first used to interpret ML in activity-travel behavior studies. Users who wish to try ML but are unfamiliar with ML often need to understand how ML comes to decisions. By transforming ML into an interpretable surrogate model, users without ML knowledge can understand the inner workings of ML. In short, ML is used for prediction, while the surrogate model is used for interpretation. From an empirical point of view, by investigating the rules derived from the surrogate model, this study revealed factors influencing seniors' DAPs and enhanced the understanding of the activity-travel behavior of seniors in the Chinese context. Several implications for policy development and transportation planning for seniors are discussed below.

First, work status is the most influential variable. Retirement implies that seniors have more free time to spend on nonmandatory activities, but aging problems can discourage them from participating in out-of-home activities. Fewer out-of-home activities increase the risk of social exclusion and bring about health issues. Retired men are more likely to lack out-of-home activities than retired women. The findings could support governments in providing more open spaces and facilities to attract retired seniors, especially retired men, to participate in out-of-home activities.

Second, the residential built environment significantly influences the daily activity engagement of seniors. Some studies have found that trip-chaining was more prevalent in a bad built environment, where activities must be organized to avoid long travel distances to reach desired destinations (Elisabeth and Laurie 2015), but we found that multiple tour and complex tour DAPs (i.e., SM-SM, CM, SN-SN, and CN) are more likely to be undertaken in a good built environment. Household car ownership is not significant for most DAPs. Most households in Nanjing had zero or one car, and even if a household had a car, the only car in the household was usually used by working adults. Seniors rely heavily on walking and public transit to get around, while for seniors in Western countries, the car is the main mode of transportation and defines their living space. The effect of driver's license ownership is also small. All these results imply that increasing the accessibility of walking and public transit will increase the out-of-home activities of seniors.

Third, household income and education level differentiate mandatory DAPs. Seniors with a high household income and education level are more inclined to choose the single and simple mandatory tour DAP (i.e., SM), while seniors with middle or low household income and low education level are more inclined to choose the multiple or complex mandatory tour DAPs (i.e., SM-SM, SM-SN, and CM). Some studies have found that financially disadvantaged retirees may encounter social exclusion, and they may participate in out-of-home activities less frequently (Luiu, Tight, and Burrow 2016). However, we found that household income and education level have a mixed effect on nonmandatory tour DAPs. We suspected that this may be related to the generally walkable environment and the policy of free public transit for seniors in Nanjing. As the travel costs are low, the mobility of seniors in a low-income household may not be restricted. Conversely, seniors in a high-income household may need to care for the house and grandchildren to ease the burden on younger earners; their activities may be restricted.

Fourth, although Chinese family composition is special, it explains only a few classes of DAPs. The most obvious result is that seniors in extended families are more likely to stay at home. As Feng et al. show, seniors in extended families often share household responsibilities, such as caring for children and running errands; they do not have much time or energy for out-of-home activities (Feng 2017). The impact of special family composition is mainly reflected in young females' DAPs. The presence of elderly mothers tends to ease the burden of housework on young women, enabling them to participate in labor markets and recreational activities (Feng et al. 2020).

There are some limitations in this study. First, the boosted C5.0 algorithm, like other ML algorithms, requires a larger amount of data. Because the samples of minor DAPs are small, predicting minor DAPs is still difficult, although the boosted C5.0 algorithm improves the predictive accuracy of minor DAPs compared to the MNL model. Second, the household travel survey only recorded a one-day travel diary. The DAPs shown in the off-diagonal cells of the confusion matrices may be seniors' actual DAPs on another day. A multi-day travel diary can provide further insight into the underlying decision processes of DAP arrangements, especially DAP arrangements across multiple days. Third, the DAP choice model does not consider the start and end times of tours. Integrating a time modeling component with the DAP choice model will form a full activity scheduling module in an activity-based modeling framework.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Zhejiang Provincial Philosophy and Social Science Planning Project (19NDQN351YB) and the National Natural Science Foundation of China (51978617).

ORCID

Yiling Deng http://orcid.org/0000-0002-6245-0352

References

Allahviranloo, M., and W. Recker. 2013. "Daily Activity Pattern Recognition by Using Support Vector Machines with Multiple Classes." *Transportation Research Part B: Methodological* 58: 16–43.

Auld, J., and A. Mohammadian. 2009. "Framework for the Development of the Agent-Based Dynamic Activity Planning and Travel Scheduling (ADAPTS) Model." *Transportation Letters* 1: 245–255.

Böcker, L., P. Van Amen, and M. Helbich. 2017. "Elderly Travel Frequencies and Transport Mode Choices in Greater Rotterdam, the Netherlands." *Transportation* 44: 831–852.

Bowman, J. L., and M. E. Ben-Akiva. 2001. "Activity-Based Disaggregate Travel Demand Model System with Activity Schedules." *Transportation Research Part A: Policy and Practice* 35: 1–28.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32.

Brier, G. W. 1950. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78: 1–3.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. "SMOTE Synthetic Minority Over-Sampling Technique." *Journal of Artificial Intelligence Research* 16: 321–357.

Cheng, L., X. Chen, S. Yang, Z. Cao, J. De Vos, and F. Witlox. 2019. "Active Travel for Active Ageing in China: The Role of Built Environment." *Journal of Transport Geography* 76: 142–152.

China's Population Aging Trend Forecast Report. China National Ageing Committee, 2018.

Choo, S., D. Sohn, and M. Park. 2016. "Mobility Characteristics of the Elderly: A Case for Seoul Metropolitan Area." *KSCE Journal of Civil Engineering* 20: 1023–1031.

Davidson, W., P. Vovsha, J. Freedman, and R. Donnelly. CT-RAMP Family of Activity-Based Models. Presented at *Proceedings of the 33rd Australasian Transport Research Forum (ATRF)*, Canberra, 2010.

Dianat, L., K. N. Habib, and E. J. Miller. 2020. "Modeling and Forecasting Daily Non-Work/School Activity Patterns in an Activity-Based Model Using Skeleton Schedule Constraints." *Transportation Research Part A: Policy and Practice* 133: 337–352.

Elisabeth, Z., and B. Laurie. 2015. "Mobility and Out-of-Home Activities of Older People Living in Suburban Environments: 'Because I'm a Driver, I Don't Have a Problem'." *Ageing & Society* 35: 785–808.

Ewing, R., and R. Cervero. 2010. "Travel and the Built Environment: A Meta-Analysis." *Journal of the American Planning Association* 76: 265–294.

Fasihozaman Langerudi, M., T. H. Rashidi, and A. Mohammadian. 2016. "Individual Trip Rate Transferability Analysis Based on a Decision Tree Approach." *Transportation Planning and Technology* 39: 370–388.

Feng, J. 2017. "The Influence of Built Environment on Travel Behavior of the Elderly in Urban China." *Transportation Research Part D: Transport and Environment* 52: 619–633.

Feng, J., X. Chuai, Y. Lu, X. Guo, and Y. Yuan. 2020. "Who Will Do More? the Pattern of Daily Out-of-Home Activity Participation in Elderly Co-Residence Households in Urban China." *Cities* 98: 102586.

Feng, J., M. Dijst, B. Wissink, and J. Prillwitz. 2013. "The Impacts of Household Structure on the Travel Behaviour of Seniors and Young Parents in China." *Journal of Transport Geography* 30: 117–126.

Fernandez-Delgado, M., E. Cernadas, S. Barro, and D. Amorim. 2014. "Do We Need Hundreds of Classifiers to Solve Real World Classification Problems?" *Journal of Machine Learning Research* 15: 3133–3181.

Habib, K. M. N., and V. Hui. 2017. "An Activity-Based Approach of Investigating Travel Behaviour of Older People." *Transportation* 44: 555–573.

Hafezi, M. H., L. Liu, and H. Millward. 2018. "Learning Daily Activity Sequences of Population Groups Using Random Forest Theory." *Transportation Research Record: Journal of the Transportation Research Board* 2672: 194–207.

Hafezi, M. H., L. Liu, and H. Millward. 2019. "A Time-Use Activity-Pattern Recognition Model for Activity-Based Travel Demand Modeling." *Transportation* 46: 1369–1394.

Hahn, J., H. Kim, J. Kim, and G. F. Ulfarsson. 2016. "Trip Making of Older Adults in Seoul: Differences in Effects of Personal and Household Characteristics by Age Group and Trip Purpose." *Journal of Transport Geography* 57: 55–62.

Hou, Y. 2019. "Polycentric Urban Form and Non-Work Travel in Singapore: A Focus on Seniors." *Transportation Research Part D: Transport and Environment* 73: 245–275.

Koushik, A. N., M. Manoj, and N. Nezamuddin. 2020. "Machine Learning Applications in Activity-Travel Behaviour Research: A Review." *Transport Reviews* 40: 288–311.

Kristoffersson, I., S. Berglund, and S. Algers. 2020. "Estimation of a Large-Scale Tour Generation Model Taking Travellers' Daily Tour Patterns into Account." *Transportation Planning and Technology* 43: 158–173.

Lai, X., W. H. K. Lam, J. Su, and H. Fu. 2019. "Modelling Intra-Household Interactions in Time-Use and Activity Patterns of Retired and Dual-Earner Couples." *Transportation Research Part A: Policy and Practice* 126: 172–194.

Luiu, C., M. Tight, and M. Burrow. 2016. "The Unmet Travel Needs of the Older Population: A Review of the Literature." *Transport Reviews* 37: 488–506.

Marquet, O., and C. Miralles-Guasch. 2014. “Walking Short Distances. The Socioeconomic Drivers for the Use of Proximity in Everyday Mobility in Barcelona.” *Transportation Research Part A: Policy and Practice* 70: 210–222.

Martín-Baos, J. Á., R. García-Ródenas, and L. Rodriguez-Benitez. 2021. “Revisiting Kernel Logistic Regression under the Random Utility Models Perspective. An Interpretable Machine-Learning Approach.” *Transportation Letters* 13: 151–162.

Millward, H., M. H. Hafezi, and N. S. Daisy. 2019. “Activity Travel of Population Segments Grouped by Daily Time-Use: GPS Tracking in Halifax, Canada.” *Travel Behaviour and Society* 16: 161–170.

Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, Leanpub, 2020.

Moniruzzaman, M., and A. Páez. 2016. “An Investigation of the Attributes of Walkable Environments from the Perspective of Seniors in Montreal.” *Journal of Transport Geography* 51: 85–96.

Moniruzzaman, M., A. Páez, K. M. Nurul Habib, and C. Morency. 2013. “Mode Use and Trip Length of Seniors in Montreal.” *Journal of Transport Geography* 30: 89–99.

Pitombo, C., A. Sousa, and L. Filipe. 2009. “Classification and Regression Tree, Principal Components Analysis and Multiple Linear Regression to Summarize Data and Understand Travel Behavior.” *Transportation Letters* 1: 295–308.

Pitombo, C. S., A. D. de Souza, and A. Lindner. 2017. “Comparing Decision Tree Algorithms to Estimate Intercity Trip Distribution.” *Transportation Research Part C: Emerging Technologies* 77: 16–32.

Pitombo, C. S., E. Kawamoto, and A. J. Sousa. 2011. “An Exploratory Analysis of Relationships between Socioeconomic, Land Use, Activity Participation Variables and Travel Patterns.” *Transport Policy* 18: 347–357.

Quinlan, J.R. C4.5: Programs for Machine Learning, Elsevier, 2014.

Rashidi, T. H., and A. Mohammadian. 2011. “Household Travel Attributes Transferability Analysis: Application of a Hierarchical Rule Based Approach.” *Transportation* 38: 697–714.

Rasouli, S., and H. Timmermans. 2014. “Activity-Based Models of Travel Demand: Promises, Progress and Prospects.” *International Journal of Urban Sciences* 18: 31–60.

Shamshiripour, A., and A. Samimi. 2019. “Estimating a Mixed-Profile Mdcev: Case of Daily Activity Type and Duration.” *Transportation Letters* 11: 289–302.

Soltani, A., D. Pojani, S. Askari, and H. E. Masoumi. 2018. “Socio-Demographic and Built Environment Determinants of Car Use among Older Adults in Iran.” *Journal of Transport Geography* 68: 109–117.

Szeto, W. Y., L. Yang, R. C. P. Wong, Y. C. Li, and S. C. Wong. 2017. “Spatio-Temporal Travel Characteristics of the Elderly in an Ageing Society.” *Travel Behaviour and Society* 9: 10–20.

Walk Score Methodology. Walk Score®, 2011.

World Population Prospect. United Nations, Department of Economic and Social Affairs, Population Division, 2019.

Zhan, G., X. Yan, S. Zhu, and Y. Wang. 2016. “Using Hierarchical Tree-Based Regression Model to Examine University Student Travel Frequency and Mode Choice Patterns in China.” *Transport Policy* 45: 55–65.

Zhang, Y., E. Yao, R. Zhang, and H. Xu. 2019. “Analysis of Elderly People’s Travel Behaviours during the Morning Peak Hours in the Context of the Free Bus Programme in Beijing, China.” *Journal of Transport Geography* 76: 191–199.

Zhao, X., X. Yan, A. Yu, and P. Van Hentenryck. 2020. “Prediction and Behavioral Analysis of Travel Mode Choice: A Comparison of Machine Learning and Logit Models.” *Travel Behaviour and Society* 20: 22–35.