Scientia Iranica A (2023) 30(3), 936–951

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions A: Civil Engineering
http://scientiairanica.sharif.edu

A new binary genetic programming approach to designing public transportation systems according to transit-oriented development criteria

R. Gholizadeh[a], H. Yılmaz[a,b], and A. Danandeh Mehr[c,*]

a. *Department of Landscape Architecture, Faculty of Architecture and Design, Ataturk University, Erzurum, Turkey.*
b. *Faculty of Agriculture, Kirgizstan Türkiye Manas University, Bishkek, Kirgizstan.*
c. *Department of Civil Engineering, Antalya Bilim University, Antalya, Turkey.*

Received 15 December 2021; received in revised form 12 February 2022; accepted 9 May 2022

KEYWORDS
Urban planning;
Genetic programming;
Transit-oriented development;
Decision tree;
Random forest.

Abstract. This study introduces a new evolutionary approach called Binary Genetic Programming (BGP) to design and assess public transportation systems from a sustainable development perspective. The BGP combines evolutionary system identification techniques with k-fold cross-validation to obtain an accurate model between the land use and transportation parameters from a sustainable urban development point of view. To assess the new model, two public transportation systems including the new tram line of Antalya (Turkey) and the bus rapid transit line of Bhopal (India) were considered. The model was employed to classify the transportation systems into Transit-Oriented Development (TOD) and non-TOD. The solutions generated by the new model were compared with those of classic Decision Tree (DT) as well as the state-of-the-art Random Forest (RF) models evolved as the benchmarks in this study. The results showed that the BGP was highly efficient and might provide less than 5% classification error. It is superior to the DT and RF solutions, which typically require higher datasets to avoid overfitting. Furthermore, the explicit formulation of BGP in combination with the multicriteria evaluation method increases human insight into the factors affecting the design of public transportations from a sustainable urban development point of view.

1. Introduction

Transit-Oriented Development (TOD) is an urban development method that integrates land use with coordinated transportation system [1]. It is a robust approach to design urban areas considering sustainable development perspectives [2,3]. Indeed, TOD enhances the access to public transit via Mixed Use of Land (MUL) features and other urban elements [4]. As a result, it reduces the dependency on cars and positively impacts public life on TOD regions. During the last few decades, TOD has been adopted in several ways as a practical method for sustainable urban development [3,5]. For example, in New Jersey, Portland, and San Francisco, the regenerated areas encompass TOD principles in their planning to ameliorate MUL promote pedestrian facility, affordable housing, and increase public health [6]. In some European cities, such as Paris and Amsterdam, TOD mainly implements

*. *Corresponding author. Tel.: +90 553 4178028; Fax: +90 242 2450045*
E-mail addresses: reza.gholizadeh11@ogr.atauni.edu.tr (R. Gholizadeh); hyilmaz@atauni.edu.tr (H. Yılmaz); ali.danandeh@antalya.edu.tr (A. Danandeh Mehr)

doi: 10.24200/sci.2022.59565.6310

curbing the urban design to make it suitable for transportation modes, mixed uses, and public space within a range [7,8]. In Delhi, Tokyo, and Hong Kong, TOD practice successfully reduced automobile dependency, urban sprawl, and affordable housing [9,10]. Bus-based TOD trick in Ahmedabad (India) curbs sprawl by promoting compact high-density development near public transit [11]. In Brisbane, Melbourne, Perth, and Sydney, the public transport-focused TOD was applied to solve lower-density urban sprawl and connection to the town centers [12]. Loo and du Verle [13] proposed a two-level sustainable mobility strategy for future cities in which the main roles of TOD with respect to both the internal and external movements of people were discussed. The authors in Hong Kong found that people living in TOD area had higher public transport shares than those living in non-TOD neighborhoods. More recently, Yu et al. [14] presented a Latent Dirichlet allocation-based perspective to define thematic characteristics of urban functions for each metro station catchment in Hong Kong. In this connection, the authors demonstrated that the functional patterns within stations' catchment could be considered for formulating more targeted TOD strategies.

To assess the level of TOD, urban designers commonly use TOD indices [1,3,15]. Our review shows that multicriteria analysis [e.g., 10,16] and/or data envelopment analysis were implemented to measure the TODness [e.g., 17,18]. However, the germane literature lacks attempt on the prediction and classification of the areal TODness. This is an important task in sustainable urban development that supports decision-makers to recognize significant hot spots that meet (ideal) TOD standards. Thus, an existing condition/area could be adopted or refused for upcoming investment or complementary infrastructure and services such as developing a new public transport system or building a new stop/station in an existing one. However, it requires a profound understanding of the factors affecting its performance. Therefore, the selection of the parameters according to the geographical context to find the relationship between predictor variables and TODness is an essential step for evaluation.

Despite considerable research on the satisfactory use of Artificial Intelligence (AI) techniques in urban design and planning [e.g., 19,20], classification of TOD zones has not been explored yet. To bridge this research gap, the present study aimed at developing a new AI-baes classification model that integrates the multicriteria analysis with the state-of-the-art Genetic Programming (GP) technique. In previous studies, AI techniques such as Artificial Neural Networks (ANNs), fuzzy logic, Decision Tree (DT), and Support Vector Machines (SVM) are generally used by urban planners to identify the urban evolution pattern or simulation of development alternatives [21–28]. Although GP has been used in different civil engineering disciplines [e.g., 29–33], the use of GP for classification is not yet ubiquitous in civil engineering and urban design [34].

Inspired by the literature review, the main goal of this study is, for the first time, to develop a GP-based model for classification of TOD zones. The proposed model implements multi-source geospatial datasets including built environment and transportation system indicators to classify stations in two public transportation systems (a tramway line in Antalya, Turkey, and a Bus Rapid Transit (BRT) line in Bhopal, India) into TOD and non-TOD zones that are required for sustainable urban development, evaluation, and planning purposes. To verify the efficiency of the introduced Binary GP (hereafter BGP) approach, we compared its classification performance with that attained using the classical DT and state-of-the-art Random Forest (RF) models developed as the benchmarks in this study. To the best of the authors' knowledge, only a few TOD examinations have been conducted for Turkish and Indian cities [e.g., 18,35–38]. The present study is the foremost research that investigates TOD classification in tram and BRT systems. The proposed model is believed to be useful for not only the sustainable development of the case-study areas, but also TOD zones classification applications in other areas and public transportation systems.

2. Methods

2.1. Generic Programing (GP)

The GP is an emerging AI approach that applies evolutionary algorithms to identify an explicit relationship for a given process [39]. It has different variants including (but not limited to) monolithic GP, linear GP, multistage, and multigene GP. However, in all types, a population of random solutions (programs) is formed at the outset and then, the genetic items of each program are progressively changed to achieve the desired solution [40,41]. The computer programs have a tree structure comprising a root/function node, inner nodes, branches, and terminal nodes (leaves). Figure 1 shows a GP tree and the associated mathematical expression.

The main steps required to develop a GP-based

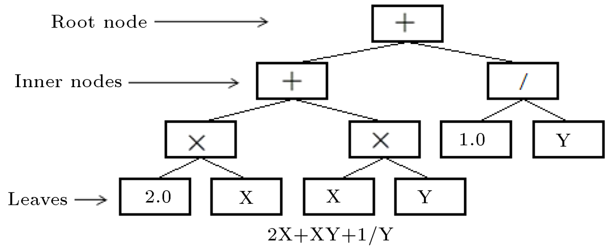


Figure 1. An exemplary genome and its mathematical expression.

decision-making model include (i) the educated selection of dependent variables known as inputs, (ii) wise guess about system functions (mathematical, logical, or Boolean), and (iii) appropriate tuning of evolutionary operators. Skilled decision-making during these steps helps the GP algorithm evolve precise models and reduces the time of computations [42]. Regardless of the type of the problem, regression, or classification, the GP algorithm starts with random establishment of the initial programs known as potential models. At that point, the programs are sorted based on their goodness of fitness, and the ones demonstrating higher suitability would be chosen as parents subjected to the evolutionary operations of crossover and mutation [42]. These evolutionary operators mimic biological evolution processes reflecting "survival of the fittest". To solve classification problems, as it is the case in this study, an additional root node is generally selected so that the numerical model outputs can be categorized into the user-defined classes. This is a challenging task that provides the injection of the modeler's knowledge into the black-box programing technique. Since there is no universal way to determine the root node, one may use a trial-and-error procedure to find the best function. According to the literature, *softmax* and *logistic* functions are frequently used for classification task in engineering problems [43]. The reader is referred to Koza [39] for a more detailed explanation of GP theory and its evolutionary mechanisms.

2.2. Decision Tree (DT)

DT is one of the frequently used classification algorithms in practice. It presents a strategy that advances from top to bottom or from general to specific during the training process. In this strategy which has a kind of tree structure, the attribute value of each node is measured and branches are formed with the achieved results. A typical DT architecture is shown in Figure 2.

DT modeling begins with root node selection accompanied by separation criteria such as information gain. According to the selected criteria, this root node is divided into branches. The separation continues until terminated by leaf nodes. In this structure, the path followed by the data progresses with the answer given in binary (yes/no), categorical, or numerical distinctions depending on if-then-else algorithms. Tree contains all trained cases in the root node and checks for clustering. If all cases in the root node coincide with a single cluster, the solution is achieved. Otherwise, the root node is divided into branches and repeated until the branch is simple enough to decide directly. Based on the size of the dataset, branches and depth of tree are specified by the modeler. To prevent overfitting, pruning algorithms are used, which refers to removing leaf nodes containing a small number of objects from the DT. The reader is referred to Jijo and Abdulazeez [44] for details on DT theory and its types, benefits, and drawbacks.

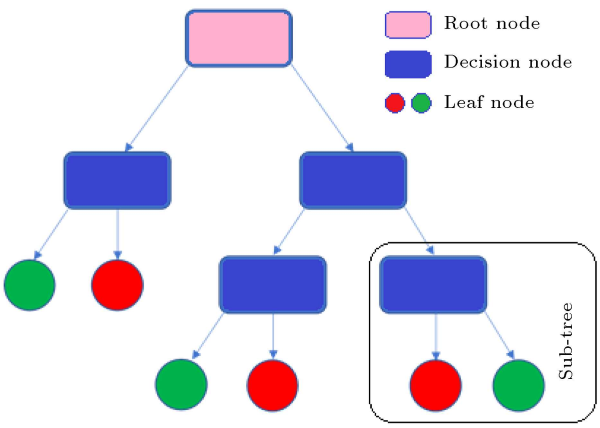


Figure 2. A typical decision tree architecture.

2.3. Random Forest (RF)

RF is a classification and regression method that uses an ensemble of binary DT that has been trained individually, with the conclusion calculated by taking into consideration the findings acquired by each DT [45]. Each tree is built using a distinct bootstrap sample chosen randomly. One-third of the original data is left out to be used for testing the forest. Without the need for any pruning, RF models have the potential to generalize and reduce the risk of overfitting. Also, based on the premise that a combination of predictions is more accurate than using only one prediction, RF seems to give better results than DT. For more details about RF, the reader is referred to Breiman [45].

2.4. The proposed BGP model for TODness classification

Combination of traditional multicriteria evaluation methods with AI techniques can yield accurate models to describe and classify complex urban areas. Although GP is known as a symbolic regression tool, it could be improved in a way that the evolved evolutionary models solve a classification task [46]. For example, to predict the possibility of heavy rainy month in two different cities in Iran, a GP-based classifier was developed by Danandeh Mehr et al. [43] in which the numerical model output was rescaled to be one of categorical outputs of Yes or No. To this end, the authors imposed the Logistic regression function on the best GP tree as a secondary root node. Inspired by that study, a new BGP model that uses *sigmoid* function is proposed as the additional root node capable of transforming the numerical model output into the binominal values of 0.0 and 1.0 that denote non-TOD and TOD areas, respectively. The sigmoid function (Eq. (1)) is a generalization of logistic regression function that turns a vector of K real values into a vector of K real values equaling 1. In this study, a data set of 21 tramway stations (i.e., $X = (x_1, x_2, ..., x_K$ and $K = 21$) with labels $Y = (y_1, y_2)$ where $y_i \in \{0, 1\}$ is considered, indicating a problem of binary classes.

$$S\left(\vec{X}\right)_i = \frac{e^{x_i}}{e^{x_i}+1}, \tag{1}$$

where e^{x_i} is the standard exponential function applied to each element of the input vector. The denominator is the normalization terms such that all outputs can be in the range (0,1), indicating a valid probability.

As illustrated in Figure 3, the initialization of the BGP model starts with the collection of the required data (explained later in the next section) to calculate TODness value of each station. TODness value of each station in this study is determined based on available data and multicriteria analysis (see Section 2.4). Then, the gathered dataset is divided into two parts: training (the first 75%) and testing (the last 25%) sets. As the problem at hand has a limited dataset (21 tram stations for Antalya and 16 BRT stations for Bhopal), a three-fold cross-validation method is applied using the training and testing datasets to ensure model robustness. To this end, the data is randomly partitioned into three equal-size training and testing subsets, and then each subset is used to evolve a BGP model.

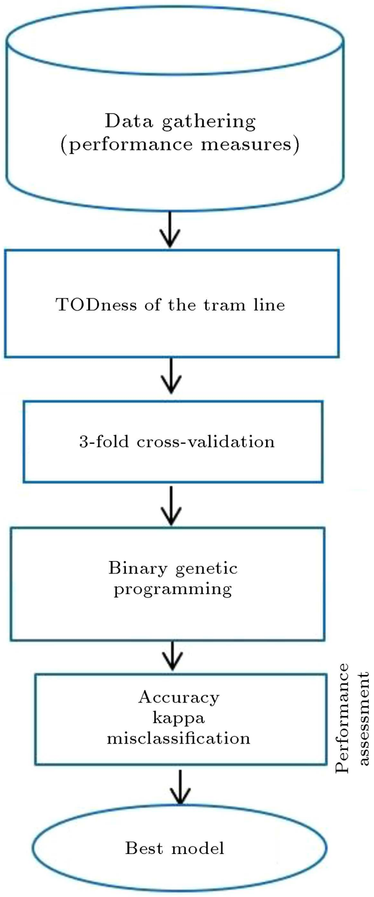


Figure 3. Flowchart of the proposed binary genetic programming approach to evolve TODnees classifier.

The attained accuracy measures are finally averaged to reduce inevitable uncertainty raised from the limited data set. The advantage of this method is that roughly all datapoints are used for both training and validation. To train each set, we used GPdotNet [42] that applied classic GP to formulate each label as a function of input elements (i.e., $Y = f(X)$). The tool supports different fitness functions including total accuracy, Heidke skill score, and pierce skill score. Total accuracy was used as the fitness function to train GP models in this study.

2.5. *Performance appraisal criteria*

To evaluate the accuracy of the new models, three statistical metrices including total Accuracy (AC), Kappa (KA), and Classification Error (CE) were used in this study. While TA is the averaged value of True Positive (TP) and True Negative (TN) class predictions, KA parameter indicates the reliability of the comparative agreement between two classifiers. It is calculated in the range of [0, 1] with higher values for better agreement.

$$AC = \frac{TP + TN}{m} \times 100, \tag{2}$$

$$KA = \frac{O_{AG} + E_{AG}}{1 - E_{AG}}, \tag{3}$$

$$CE = 100 - TA, \tag{4}$$

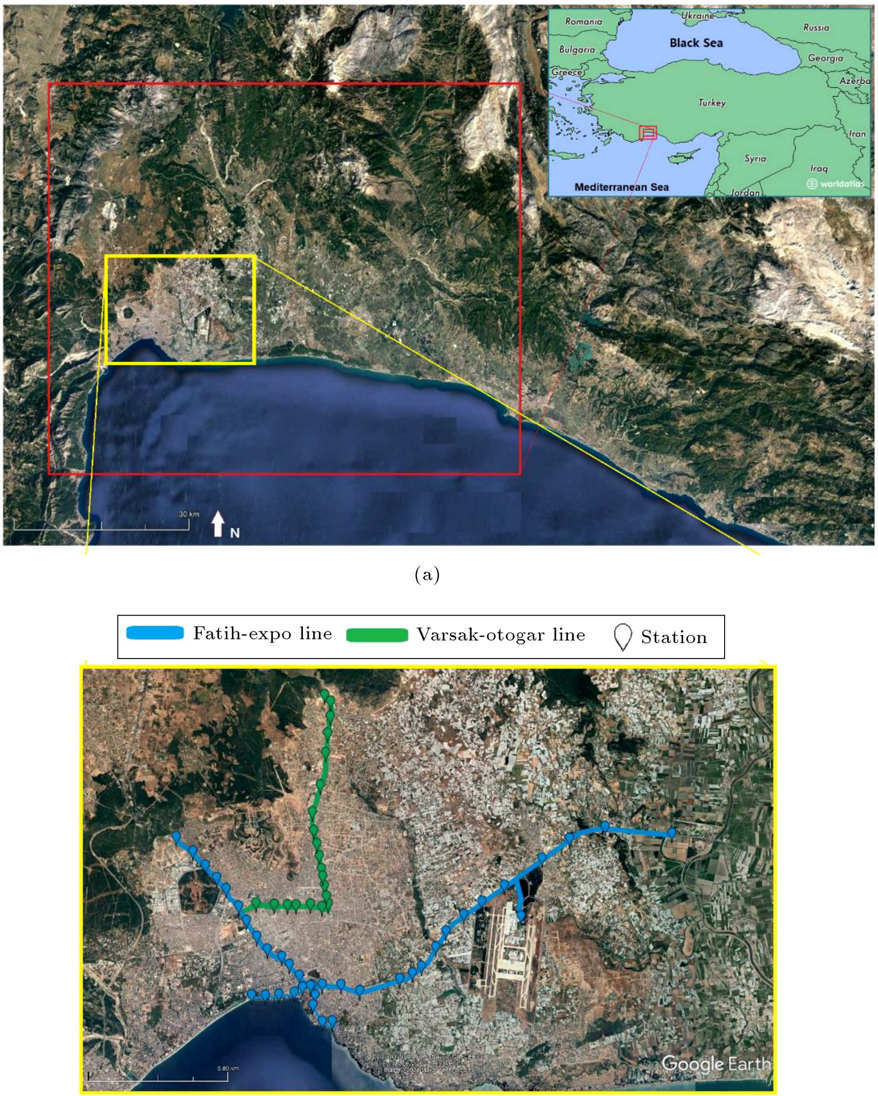
where m is the total sample number, and O_{AG} and E_{AG} are the observed and expected agreements, respectively. All these metrices can be measured using a confusion matrix of the model outputs. Basically, a confusion matrix is a table that summarizes the true and false predictions in an order. It is an $N \times N$ matrix, where N is the total number of target classes (i.e., 0 and 1).

3. Case study area and data

3.1. *Antalya new tram line*

Having a fast population growth, Antalya is located on the Mediterranean coast, south-west of Turkey, between the Torus Mountains and the Mediterranean Sea. The city of Antalya (Figure 4(a)) is of great significance to the truism industry of Turky and its picturesque beaches are the destination of worldwide tourists [47]. The urban population of Antalya is nearly 2.55 million in 2020 and is poised to 2.60 million in 2022. The city has an area of 21000 km^2 and therefore, its population density is about 124 person per square kilometer.

The city has a four-line tram (light rail) system in operation including Museum-Zerdalilik Nostalgic Tram, Fatih-Meydan Tram, and Meydan-Airport-Aksu-EXPO and Varsak-Otogar. As shown in Figure 5, the Varsak-Otogar line, which has 21 stations along



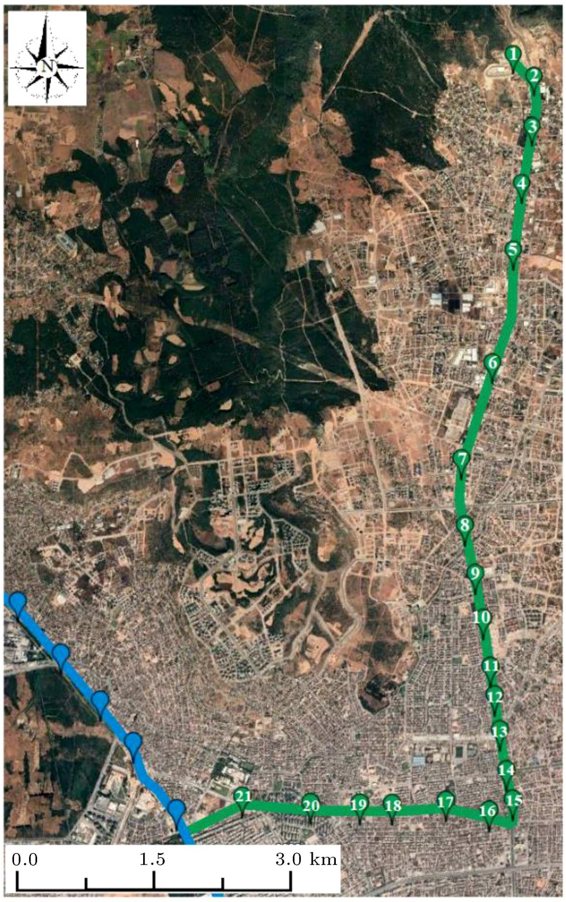
(a)

(b)

Figure 4. Study area displaying (a) Antalya municipal and (b) existing tramway lines.

its course and was partly under construction during this study, is considered for TOD analysis in the present study. The line of interest has a length of approximately 13 km and is located in the northeast part of the city.

The performance measures adopted for TOD index calculation at each station include *Population Density* (PD), *Land Use Diversity* (LD), *Walkable Catchment area* (WC), *Walkable Path* (WP), and *Travel Behavior* (TB). These performance measures were collected as they are frequently suggested in the germane literature [1,48] and available to evaluate the level of TOD among the selected stations. It should be noted that each of these measures is calculated/determined for a buffer zone of 800-meter radius around the station. LD is a major factor affecting non-motorized and public transport-based trips, especially for work purposes, and it creates rational passenger flow of the day consistently. To calculate LD, we, first and foremost, classified the area into residential, commercial, recreational, and industrial areas; then, the Area Index, i.e., the ratio of the work areas in the buffer zone to the work areas in the whole study area, was calculated for each station. Knowledge of the TB of the transit-supportive user is necessary. Therefore, we can find the gap between the user and present transit services [49]. In the present study, TBs including tendency or use of *personal motorized vehicle, public tram, taxi,* and *walking or cycling* were used as a categorical data to distil TOD index of each station. To determine TB of dwellers at 800-meter buffer zone of each station, a total of 100 households were sampled and interviewed. Questionaries were designed so that we could measure the household members with the age of 16 to 69 and report their attitude to each of the TB measurements (i.e., personal motorized vehicle,



Id	Station name
1	Varsak Depo Alani
2	Varsak
3	Kepezpark
4	Aktoprak
5	Aydoğmuş
6	Karşiyaka
7	Şelale
8	Süleyman Demirel
9	Ulubatli Hasan
10	Fevzi Çakmak
11	Kuzeykaya
12	Gazi
13	Sütçüler
14	Gündoğdu
15	Yeşilirmak
16	Kepez Belediyesi
17	Şehitler Parki
18	Erdem Beyazit K.M.
19	Yildirim Beyazit
20	Zafer
21	Atatürk

Figure 5. The first case study area showing tramway stations and their nearby.

public tram, taxi, and walking, or cycling). It is worth mentioning that the members were asked to fulfill the questionaries regardless of the ongoing COVID-19 restrictions during the study.

According to Singh et al. [10], multicriteria analysis was used to combine the data and transfer them into a resulting decision (target binary variable). The main step in multicriteria analysis is standardization of each criterion. It is performed to adjust the units of different indicators. In this study, we simply divided each value by their maximum amount so that all the inputs were unitless. As a result, the classification algorithm gives equal importance to the adopted measures.

3.2. Bhopal BRT

To evaluate the efficiency of the BGP algorithm (see Figure 3) for TODness prediction along a BRT line, the second case study was run using available data from the literature [1]. To this end, we used TODness values (Figure 6) from 16 stations along a well-established BRT system of Bhopal, India. According to Khare et al. [1], the city suffers the problem of urban sprawl and more than 40% of the work trips in the city are made by public transport. Seven criteria including PD, LD, WC, TB, WP, MUL, and Economic Indicator

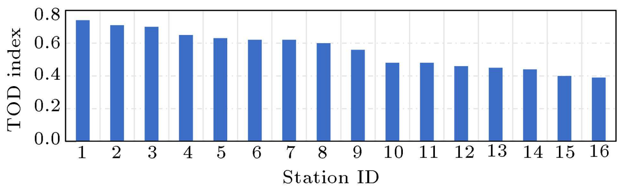


Figure 6. TODness scores of the bus rapid transit stations, Bhopal, India.

(EI) were used to quantify the TOD values at each station. Table A.1 lists the selected stations and the score of each of these performance measures. For a map of the BRT line and details about calculation of each criterion, the interested reader is referred to Khare et al. [1].

4. Results and discussion

As previously described, the novel BGP and classic DT approaches were applied to classify the Antalya tramway and Bhopal BRT stations into TOD and non-TOD areas. In this section, the attained results for each case study area are presented and discussed separately.

4.1. Results of case study#1

To develop BGP and DT models for Antalya tram line, performance measures of 21 tram stations together with the outputs of multiple criteria analysis were considered. Table 1 summarizes the TODness values for all the stations attained through multicriteria analysis and reflects the involvement of indicators in planning for TOD. It is observed from the table that nine stations on the line secure the TOD condition (i.e., TOD index ≥ 0.6). On a scale of one, the maximum TOD index value is 0.772 for Zafer station, while the minimum index value is 0.336 for Varsak Depo Alani station. Since Zafer district is one of the central residential regions and homes for a fame hospital in Antalya, having a high TOD score is in not surprising and agrees with local Moovit (moovitapp.com) records. By contrast, the Varsak Depo Alani station is a non-residential/commercial area with less PD and is mainly used as a warehouse area; thus, it is expected to score very low.

To train and test the models, the datapoints (Table 1) were split into 75% (16 stations) of training and 25% (5 stations) of testing datasets. Then, 3-fold strategy was used to randomly create subsets used for BGP and DT modeling. For example, Table A.2 represents the subsets used in this study. The associated BGP and DT/RF prediction models were evolved using GPdotNet and RapidMiner, respectively. These are well-documented tools with a free license. The BGP model configuration includes the determination of the rate of evolutionary operations, maximum number of generations, functional set, objective function, and additional root node function. In this study, GP engine was set to generate solutions with maximum tree depth of six and improve it up to 500 generations. However, the training process at each fold revealed that the BGP algorithm quickly converged to the best solution after a few number of generations (see Appendix B: Figure B.1). The adopted crossover, mutation, and reproduction rates are 0.9, 0.05, and 0.2, respectively. Figure 7 illustrates the best evolved BGP trees for each fold. The associated DT models are also presented in Figure 8. In both DT and RF modeling, the maximum tree depth was set to 10. According to RF literature,

Table 1. The input and target variables used for the binary genetic programming and decision tree modeling (Population Density (PD), Land use Diversity (LD), Walkable Catchment area (WC), Walkable Path (WP), and Travel Behavior (TB)).

ID	Station name	PD	LD	WC	WP	TB	TOD	Binominal label
1	Varsak Depo Alani	0.033	0.144	0.203	0.302	1.000	0.336	Non-TOD*
2	Varsak	0.066	0.462	0.489	0.438	1.000	0.491	Non-TOD
3	Kepezpark	0.131	0.862	0.195	0.354	0.943	0.497	Non-TOD
4	Aktoprak	0.458	0.531	0.559	0.579	0.957	0.617	TOD**
5	Aydoğmuş	0.479	0.325	0.541	0.824	0.786	0.591	TOD
6	Karşiyaka	0.675	0.502	0.345	0.689	0.743	0.591	TOD
7	Şelale	0.494	0.837	0.744	0.540	0.629	0.649	TOD
8	Süleyman Demirel	0.547	0.364	0.399	0.581	0.800	0.538	Non-TOD
9	Ulubatli Hasan	0.057	1.000	0.351	0.378	0.600	0.477	Non-TOD
10	Fevzi Çakmak	0.451	0.837	0.261	0.724	0.729	0.600	TOD
11	Kuzeykaya	1.000	0.475	0.460	0.640	0.600	0.635	TOD
12	Gazi	0.536	0.585	0.365	0.570	0.586	0.528	Non-TOD
13	Sütçüler	0.419	0.381	0.377	0.481	0.571	0.446	Non-TOD
14	Gündoğdu	0.673	0.517	0.346	0.389	0.643	0.514	Non-TOD
15	Yeşilirmak	0.462	0.642	0.383	0.492	0.557	0.507	Non-TOD
16	Kepez Belediyesi	0.148	0.762	0.218	0.613	0.657	0.479	Non-TOD
17	Şehitler Parki	0.306	0.869	0.853	0.603	0.457	0.618	TOD
18	Erdem Beyazit K.M.	0.298	0.805	0.372	0.498	0.514	0.498	Non-TOD
19	Yildirim Beyazit	0.328	0.787	0.268	0.487	0.586	0.491	Non-TOD
20	Zafer	0.337	0.880	1.000	0.956	0.686	0.772	TOD
21	Atatürk	0.363	0.800	0.758	1.000	0.371	0.659	TOD

Note: *: Negative Label; **: Positive label

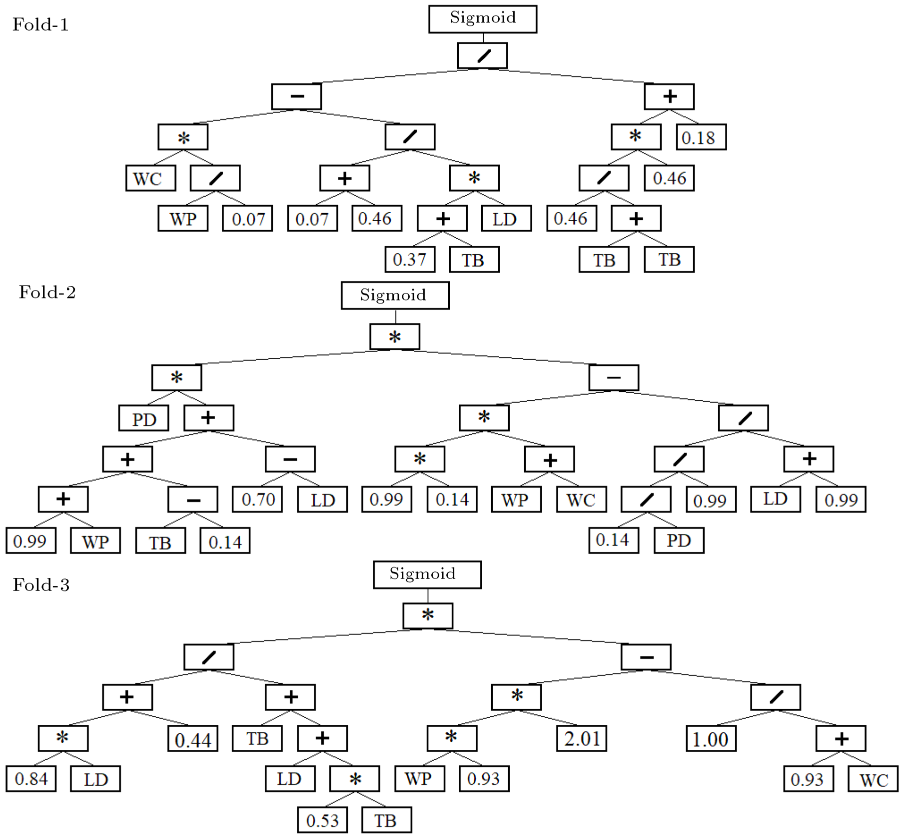


Figure 7. Binary genetic programming models derived for TOD area prediction in Varsak-Otogar tramway line, Antalya (Population Density (PD), Land use Diversity (LD), Walkable Catchment area (WC), Walkable Path (WP), and Travel Behavior (TB)).

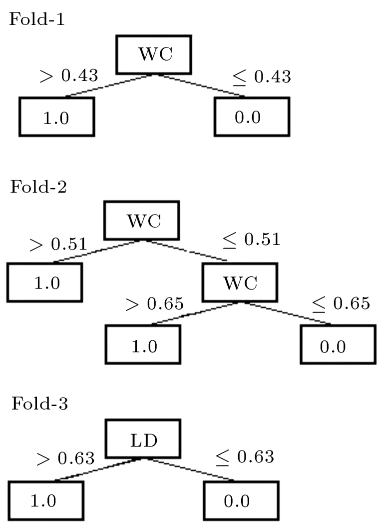


Figure 8. Same as Figure 7, but evolved by decision tree (Land use Diversity (LD), Walkable Catchment area (WC), and Walkable Path (WP)).

the maximum number of trees was limited to 100 so that the models could not be overfitted. Given that 100 trees are generated for each fold, the tree representation of RF models is not depicted in this article.

Through the visual comparison of Figures 7 and 8, it is observed that BGP puts forward more complex structures than DT in which all the performance measures contribute to the model prediction. Conversely, DT predictions were limited to the use of merely WC and WP. To recognize the best structure and attain the models' accuracy, the classification matrices of these structures were compared in Tables 2 and 3, respectively.

Each row in the tables shows which TOD area was predicted by the model. The columns are the ground truth indicating the number events in each class. The values lying across the main diagonal show the success of the model at each class. These matrices together with Eqs. (2) to (4) were utilized to calculate the classification accuracy of the models (see Table 4). It is seen from the table that the highest classification accuracy belongs to the BGP models. In all folds at the training phase, the approach achieves 100% accuracy in terms of all the efficiency criteria. The same accuracy in the testing period, particularly in Folds 2 and 3 implies that the BGP model deserves generalization capability and does not fall into the overtraining trap. By contrast, the benchmark DT and RF model shows a perfect solution merely in the training data sets of Fold 2. Regarding the testing datasets, misclassification rate of the DT (RF) is 20% (20%) at Fold 2, 40% (35%) at Fold 1, and 40% (31.3%)

Table 2. The confusion matrix of the binary genetic programming models evolved based on the presence and absence of TOD.

	Training		Testing	
	True TOD	**True non-TOD**	**True TOD**	**True non-TOD**
Fold-1				
Pred. TOD	7	0	2	0
Pred. non-TOD	0	9	1	2
Fold-2				
Pred. TOD	6	0	3	0
Pred. non-TOD	0	10	0	2
Fold-3				
Pred. TOD	6	0	3	0
Pred. non-TOD	0	10	0	2

Table 3. The confusion matrix of the decision tree models evolved based on the presence and absence of TOD.

	Training		Testing	
	True TOD	**True non-TOD**	**True TOD**	**True non-TOD**
Fold-1				
Pred. TOD	6	0	1	1
Pred. non-TOD	1	9	1	2
Fold-2				
Pred. TOD	10	0	2	1
Pred. non-TOD	0	6	0	2
Fold-3				
Pred. TOD	10	1	2	2
Pred. non-TOD	0	5	0	1

Table 4. Performance results of the Binary Genetic Programming (BGP), Decision Tree (DT), and Random Forest (RF) models at the training and validation stages (total Accuracy (AC), Kappa (KA), and Classification Error (CE)).

Model	**Fold**	**Training**			**Validation**		
		AC (%)	**KA**	**CE (%)**	**AC (%)**	**KA**	**CE (%)**
BGP	Fold 1	100	1	0.0	80	0.615	20
	Fold 2	100	1	0.0	100	1	0.0
	Fold 3	100	1	0.0	100	1	0.0
	Average	100	1	0.0	93.3	0.871	6.7
DT	Fold 1	93.75	0.871	6.25	60.0	0.167	40
	Fold 2	100	1	0.0	80	0.615	20
	Fold 3	93.75	0.862	6.25	60	0.286	40
	Average	95.83	0.91	4.167	66.67	0.356	33.33
RF	Fold 1	95.00	0.891	5.00	65.0	0.178	35.0
	Fold 2	100	1	0.0	80	0.615	20.0
	Fold 3	97.75	0.932	2.25	68.66	0.376	31.33
	Average	97.58	0.941	2.417	71.22	0.390	28.78

at Fold 3. This implies (i) the lower generalizability of RF and DT models if they are not considered over-trained solutions and (ii) RF provides less uncertain solutions than DT. This drawback of DT may rely on the limited number of datapoints available in both training and testing subsets. Thus, the algorithm yields more uncertain classes for testing datapoints. Despite bootstrapping the trees in RF, it cannot increase DT accuracy as much as BGP. The authors believe that the evolutionary feature of GP algorithm that provides more divers with solutions at the training phase of BGP diminishes the negative effect of the limited datasets at the training phase; thus, the algorithm can detect local optimum solutions. According to the mean KA statistics that compare the models in terms of the number of stations and the areas that are correctly categorized, it can be concluded that the BGP classified the stations 100% and 87.1% accurately in the training and testing datapoints. Despite the acceptable classification range in the training datapoints ($KA = 91\%$), the DT and RF provided relatively fair results in the testing datapoints ($KA = 36\%$ and 39%, respectively). As in the literature where the KA values less than 40% are considered as fair and more than 60% as substantial classification [50,51], the proposed BGP model could be interpreted as a substantially satisfactory classifier for TOD area prediction in Antalya tramway system.

4.2. Results of case study#2

Like the evolution of BGP, DT, and RF models for Antalya tram line, normalized performance measures of 16 BRT stations across Bhopal together with their binomial TODness values were employed (see Table A.1). According to Table A.1, half of the BRT stations secure the TOD conditions. On a scale of one, the maximum and minimum TOD index values are 0.74 and 0.39 for New Market station and Ashima Mall, respectively. The TODness range is more or less the same with that of Antalya tram line. In a similar approach explained in the results of the first case study, 3-fold strategy was employed to randomly create subsets and develop BGP, DT, and RF models. Figures 9 and 10 illustrate the best evolved BGP trees and DT models for TODness classification (at each fold), respectively. The associated accuracy results together with those of RF are tabulated in Table 5. Similar to the finding from the first case study, the BGP simulation exhibited fast convergence to the best solution (see Appendix B: Figure B.2).

By comparing Figures 9 and 10, it is seen that BGP evolves more complex structures than DT in which most of performance measures contributed to the model prediction. Inversely, DT predictions were limited to the use of only WC and LD. Given that 100 trees were developed in the best RT model, we did

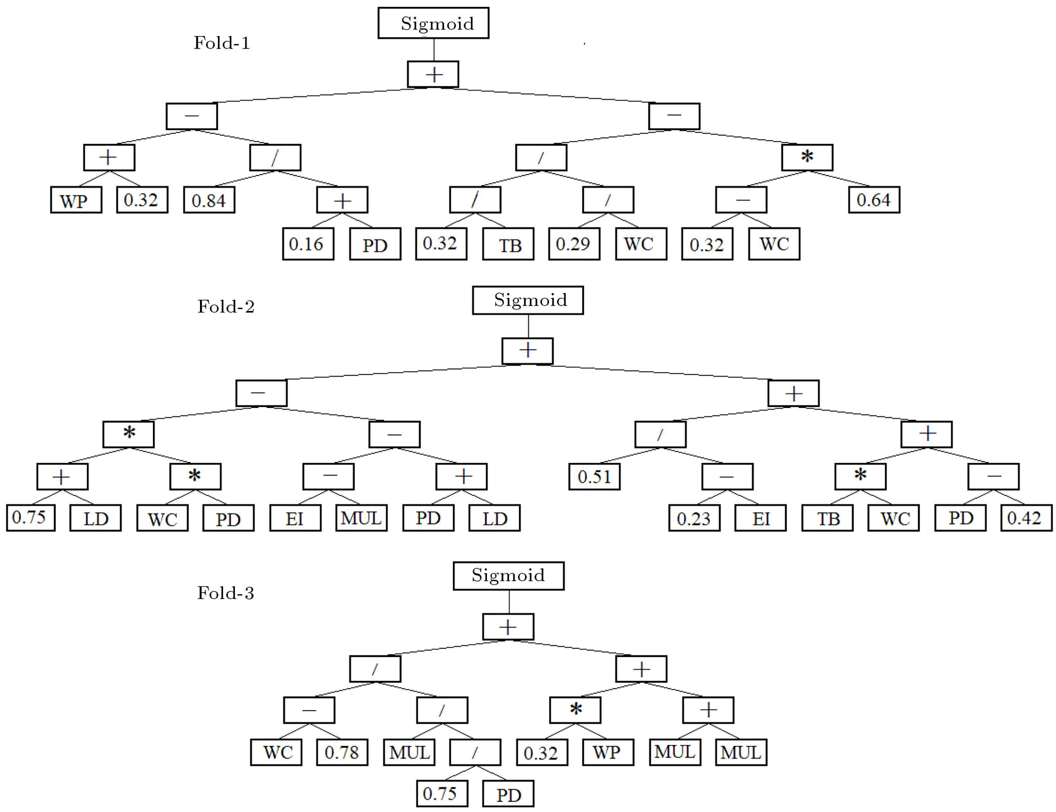


Figure 9. Binary genetic programming models derived for TOD area prediction along Bhopal bus rapid transit line (Population Density (PD), Land use Diversity (LD), Walkable Catchment area (WC), Walkable Path (WP), and Travel Behavior (TB)).

Table 5. Performance results of the Binary Genetic Programming (BGP), Decision Tree (DT), and Random Forest (RF) models at the training and validation stages (total Accuracy (AC), Kappa (KA), and Classification Error (CE)).

Model	Fold	Training (12 stations)			Validation (4 stations)		
		AC (%)	KA	CE (%)	AC (%)	KA	CE (%)
BGP	Fold 1	100	1	0.0	75.0	0.50	25.0
	Fold 2	100	1	0.0	100	1	0.0
	Fold 3	100	1	0.0	100	1	0.0
	Average	100	1	0.0	91.67	0.833	8.33
DT	Fold 1	93.75	0.871	6.25	60.0	0.167	40.0
	Fold 2	100	1	0.0	75.0	0.50	25.0
	Fold 3	91.67	0.833	8.33	100	1	0.0
	Average	95.14	0.90	4.86	78.33	0.56	21.67
RF	Fold 1	100	1	0.0	80	0.615	20
	Fold 2	100	1	0.0	100	1	0.0
	Fold 3	100	1	0.0	75.0	0.50	25.0
	Average	100	1	0.0	85.0	0.705	15.0

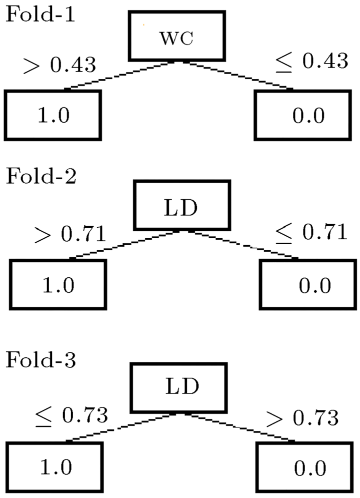


Figure 10. Same as Figure 9 but evolved by decision tree (Land use Diversity (LD) and Walkable Catchment area (WC)).

not provide the models' trees here. However, it can be stated that RF solutions are more complicated than both GP and DT. With respect to the model's average accuracy summarized in Table 5, GP and RF are 100% successful in the training period. However, their accuracy in the testing stations diminished to 92% and 85%, respectively. Despite rather satisfactory results in the training period, the DT shows significantly unreliable predictions in the unseen dataset such that its KA varies in the range of 0.17 to 1.00. Similar magnitudes of performance were also seen when BGP and DT were used to model TODness in Antalya's tramline. Although DT predictions can be improved via an ensemble approach that yields RF model, the solution remains highly complex which makes its application difficult for practical aims. Solutions evolved by GP are not only less complex, but also explicit which can elucidate directional relationships between parameters and TODness. Combining the results from both case studies, we can conclude the superiority of the proposed BGP model over DT and RF. Thus, its application for a sustainable design of public transportation systems is recommended.

Because BGP, DT, and RF rely on various mathematical algorithms, they demonstrate different results for the same datasets. This finding is consistent with the germane literature, proving that the accuracy and performance of predictions from different ML techniques could markedly vary [52]. Our results underlined that outcome from a given model might vary from fold to fold even in case of greater uncertainty, compared to different models. Undoubtedly, more reliable predictions can be achieved when the desired model is trained by the samples representing the transportation process as truly as possible. Therefore, specific attention must be devoted to data partitioning in the future applications of the proposed model. Moreover, future studies are supposed to extend the range of classifiers and ML models used in our study.

5. Summary and conclusion

The combination of traditional multicriteria evaluation methods with novel Artificial Intelligence (AI) techniques can be used to describe and design sustainable public transportation systems. Classification and prediction of Transit-Oriented Development (TOD) zones are of paramount tasks in sustainable urban development as it provides a baseline for urban planners and decision-makers to identify the urban functional area over space and time. This paper, for the first time, suggested a simple an explicit approach by which the state-of-the art Genetic Programming (GP) technique can be further developed for classifying TOD areas along a public transportation system. The new model called Binary Genetic Programming (BGP) was implemented for two use cases: Antalya tramway and Bhopal Bus Rapid Transit (BRT) system. The new model identifies the best relationship between various land use and urban features and TODness values at each station. To verify the efficiency of the BGP model, we further developed traditional DT models, with the same input/target variables as the benchmark solution. While BGP attempted to classify the area into TOD and Non-TOD areas based upon the theory of the "survival of the best", the DT algorithm applies if-then rules to solve the same task. The results of performance analysis in both case studies demonstrated that the suggested model could surely obtain the solution of classification with significantly higher accuracy than DT. Despite the apparent shortage of data, the BGP can produce models with acceptable generalization accuracy. It is well known that the data-driven models are case-sensitive and their structure must be reoptimized when different datasets are used. However, the lack of a consistent framework for quantifying TOD has made the functionality of the concept difficult. Our study was limited to the available data from two cities. Further datasets from other locations could be applied to verify the robustness of the proposed BGP approach.

Acknowledgments

The authors sincerely thank two anonymous reviewers for their constructive comments on the first version of this article. We are also grateful to the Municipality of Antalya for the data provided to accomplish this research.

References

1. Khare, R., Villuri, V.G.K., Chaurasia, D., et al. "Measurement of transit-oriented development (TOD) using GIS technique: a case study", *Arabian J. Geosci.*, **14**(10), pp. 1–16 (2021b).

2. Cervero, R. "Accessible cities and regions: A framework for sustainable transport and urbanism in the 21st century", *WP UCB-ITS-VWP-2005-3, UC Berkeley Center for Future Urban Transport*, Berkeley, CA (2005).

3. Zhou, J., Yang, Y., Gu, P., et al. "Can TODness improve (expected) performances of TODs? An exploration facilitated by non-traditional data", *Transp. Res. Part D Transp. Environ.*, **74**, pp. 28–47 (2019).

4. Ewing, R. and Cervero, R. "Travel and the built environment: A meta-analysis", *J. Am. Plan. Assoc.*, **76**(3), pp. 265–294 (2010).

5. Ibraeva, A., de Almeida Correia, G.H., Silva, C., et al. "Transit-oriented development: A review of research achievements and challenges", *Transp. Res. Part A Policy Pract.*, **132**, pp. 110–130 (2020).

6. Jamme, H.T., Rodriguez, J., Bahl, D., et al. "A twenty-five-year biography of the TOD concept: From design to policy, planning, and implementation", *J. Plan. Educ. Res.*, **39**(4), pp. 409–428 (2019).

7. l'Hostis, A. and Darchen, S., *Characterising Transit Oriented Development in the Paris Metropolitan Region: What Type of TOD Are They?* ffhal-01138076f (2015).

8. Pojani, D. and Stead, D. "Past, present and future of transit-oriented development in three European capital city-regions", In Y. Shiftan & M. Kamargianni, Eds., Preparing for the new era of transport policies: Learning from experience, pp. 93–118, Amsterdam: Elsevier (2018).

9. Loo, B.P., Chen, C., and Chan, E.T. "Rail-based transit-oriented development: lessons from New York City and Hong Kong", *Landscape Urban Plann.*, **97**(3), pp. 202–212 (2010).

10. Singh, Y.J., Fard, P., Zuidgeest, M., et al. "Measuring transit-oriented development: a spatial multi criteria assessment approach for the City Region Arnhem and Nijmegen", *J. Transp. Geogr.*, **35**, pp. 130–143 (2014).

11. Joshi, R., Munshi, T., Joseph, Y., et al., *Links Between Transit Riders and Built Form: The Case of Ahmedabad (West)*, CUE Working Paper Series, Centre for Urban Equity, Ahmedabad, India (2017).

12. Newman, P. "Transit oriented development: An Australian overview. Transit oriented development: making it happen", Australia (2005).

13. Loo, B.P.Y. and du Verle, F. "Transit-oriented development in future cities: towards a two-level sustainable mobility strategy", *Int. J. Urban Sci.*, **21**, pp. 54–67 (2017).

14. Yu, Z., Zhu, X., and Liu, X. "Characterizing metro stations via urban function: Thematic evidence from transit-oriented development (TOD) in Hong Kong", *J. Transp. Geogr.*, **99**, 103299 (2022).

15. Singh, Y.J., Lukman, A., Flacke, J., et al. "Measuring TOD around transit nodes-towards TOD policy", *Transp. Policy*, **56**, pp. 96–111 (2017).

16. Curtis, C. "Delivering the'D'in transit-oriented development: Examining the town planning challenge", *J. Transp. Land Use*, **5**(3), pp. 83–99 (2012).

17. Binglei, X.I.E. and Chuan, D.I.N.G. "An evaluation on coordinated relationship between urban rail transit and land-use under TOD mode", *J. Transp. Syst. Eng. Inf. Technol.*, **13**(2), pp. 9–13 (2013).

18. Khare, R., Villuri, V.G.K., and Chaurasia, D. "Urban sustainability assessment: The evaluation of coordinated relationship between BRTS and land use in transit-oriented development mode using DEA model", *Ain Shams Eng. J.*, **12**(1), pp. 107–117 (2021a).

19. Quan, S.J., Park, J., Economou, A., et al. "Artificial intelligence-aided design: Smart design for sustainable city development", *Environ. Plan. B Urban Anal. City Sci.*, **46**(8), pp. 1581–1599 (2019).

20. Kim, D., Shim, J., Park, J., et al. "Supervised machine learning approaches to modeling residential infill development in the city of Los Angeles", *J. Urban Plan. Dev. Div.*, ASCE, **148**(1), 04021060 (2022).

21. Yeh, A.G.O. and Li, X. "Simulation of development alternatives using neural networks, cellular automata, and GIS for urban planning", *Photogramm. Eng. Remote Sens.*, **69**(9), pp. 1043–1052 (2003).

22. Grekousis, G., Manetos, P., and Photis, Y.N. "Modeling urban evolution using neural networks, fuzzy logic and GIS: The case of the Athens metropolitan area", *Cities*, **30**, pp. 193–203 (2013).

23. Lu, S. and Liu, Y. "Evaluation system for the sustainable development of urban transportation and ecological environment based on SVM", *J. Intell. Fuzzy Syst.*, **34**(2), pp. 831–838 (2018).

24. Mirbagheri, B. and Alimohammadi, A. "Integration of local and global support vector machines to improve urban growth modelling", *ISPRS Int. J. Geo-Inf.*, **7**(9), p. 347 (2018).

25. Karimi, F., Sultana, S., Babakan, A.S., et al. "An enhanced support vector machine model for urban expansion prediction", *Comput. Environ. Urban Syst.*, **75**, pp. 61–75 (2019).

26. Kranjčić, N., Medak, D., Župan, R., et al. "Support vector machine accuracy assessment for extracting green urban areas in towns", *Remote Sens.*, **11**(6), p. 655 (2019).

27. Safari, M.J.S. and Mehr, A.D. "Design of smart urban drainage systems using evolutionary decision tree model", In *IOT Technol. Smart-Cities from Sensors to Big Data, Secur. Trust*, F. Al-Turjman and M. Imran, Eds., pp. 131–149, IET (2020).

28. Ramírez, T., Hurtubia, R., Lobel, H., et al. "Measuring heterogeneous perception of urban space with massive data and machine learning: An application to safety", *Landsc. Urban Plan.*, **208**, 104002 (2021).

29. Cevik, A., Arslan, M.H., and Köroğlu, M.A. "Genetic-programming-based modeling of RC beam torsional strength", *KSCE J. Civ. Eng.*, **14**(3), pp. 371–384 (2010).

30. Shahnazari, H., Dehnavi, Y., and Alavi, A.H. "The next-generation constitutive correlations for simulation of cyclic stress-strain behaviour of sand", *J. Civ. Eng. Manag.*, **21**(1), pp. 31–44 (2015).

31. Tsai, H.C. and Liao, M.C. "Modeling torsional strength of reinforced concrete beams using genetic programming polynomials with building codes", *KSCE J. Civ. Eng.*, **23**(8), pp. 3464–3475 (2019).

32. Danandeh Mehr, A. "Seasonal rainfall hindcasting using ensemble multi-stage genetic programming", *Theor. Appl. Climatol.*, **143**(1), pp. 461–472 (2021).

33. Keshavarz, A. and Tofighi, H. "Gene expression programming models for liquefaction-induced lateral spreading", *Sci. Iran.*, **27**(6), pp. 2704–2718 (2020).

34. Zhang, Q., Barri, K., Jiao, P., et al. "Genetic programming in civil engineering: advent, applications, and future trends", *Artif. Intell. Rev.*, **54**(3), pp. 1863–1885 (2021).

35. Aydemir, P.K., Yilmazsoy, B.K., Akyuz, B., et al. "Kentsel Ulaşımda Yaya Öncelikli Planlama/Tasarım ve Transit Odaklı Gelişimin Metropol Kentlerdeki Deneyimi, İstanbul Örneği", (In Turkish), *Kent Akademisi*, **11**(4), pp. 523–544 (2018).

36. Pilatin, K. "The investigation of marmaray line stations for conformity to transit-oriented development", *Tasarım Kuram*, **17**(32), pp. 150–159 (2021).

37. Ann, S., Yamamoto, T., and Jiang, M. "Re-examination of the standards for transit oriented development influence zones in India", *J. Transp. Land Use*, **12**(1), pp. 679–700 (2019).

38. Jain, D. and Singh, E. "Transit oriented development in India: A critical review of policy measures", *Int. J. Rec. Tech. Eng.*, **7**, pp. 745–751 (2018).

39. Koza, J.R., *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT press (1992).

40. Abooali, D. and Khamehchi, E. "New predictive method for estimation of natural gas hydrate formation temperature using genetic programming", *Neural Comput. Appl.*, **31**(7), pp. 2485–2494 (2019).

41. Mehdizadeh, S., Ahmadi, F., Mehr, A.D., et al. "Drought modeling using classic time series and hybrid wavelet-gene expression programming models", *J. Hydrol.*, **587**, 125017 (2020).

42. Hrnjica, B. and Danandeh Mehr, A. "Optimized genetic programming applications: emerging research and opportunities: emerging research and opportunities", *IGI Global*, PA (2018).

43. Danandeh Mehr, A., Nourani, V., Hrnjica, B., et al. "A binary genetic programing model for teleconnection identification between global sea surface temperature and local maximum monthly rainfall events", *J. Hydrol.*, **555**, pp. 397–406 (2017).

44. Jijo, B.T. and Abdulazeez, A.M. "Classification based on decision tree algorithm for machine learning", *J. Appl. Sci. Technol. Trends*, **2**(01), pp. 20–28 (2021).

45. Breiman, L., *Random Forests. Mach. Learn.*, **45**(1), pp. 5–32 (2001).

46. Morrison, G.A., Searson, D.P., and Willis, M.J. "Using genetic programming to evolve a team of data classifiers", *World Acad. Sci. Eng. Technol.*, **72**, pp. 261–264 (2010).

47. Tur, R., Uzunsakal, L., and Mehr, A.D. "Coastline change determination using UAV technology: a case study along the Konyaaltı coast, Antalya, Turkey", In Drones in Smart-Cities, Elsevier, pp. 123–141 (2020).

48. Vilela, A.P.L., Reboita, M.S., Silva, L.F., et al. "Sustainability neighborhoods in Brazil: a comparison of concepts and applications", *Environ. Dev. Sustain.*, **22**(6), pp. 6001–6028 (2020).

49. Kamruzzaman, M., Baker, D., Washington, S., et al. "Advance transit oriented development typology: case study in brisbane", *Australia. J. Transp. Geogr.*, **34**, pp. 54–70 (2014).

50. Landis, J.R. and Koch, G.G. "An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers", *Biometrics*, **33**, pp. 363–374 (1977).

51. Danandeh Mehr, A. "Drought classification using gradient boosting decision tree", *Acta Geophys.*, **69**, pp. 909–918 (2021).

52. Aram, A., Dalalian, M.R., Saedi, S., et al. "An assessment of data mining and bivariate statistical methods for landslide susceptibility mapping", *Sci. Iran.*, **29**(3), pp. 1077–1094 (2021). DOI: 10.24200/SCI.2021.57334.5240

Appendix A

The original datapoints (Table A.1) and subsets (Table A.2) created via three-fold cross-validation method to develop BGP and DT models in the present study.

Appendix B

The BGP fitness simulation for TOD area prediction in Varsak-Otogar tramway line as well as Antalya (Figure B.1) and Bhopal BRT line (Figure B.2). The maximum fitness (red line) shows that the BGP

Table A.1. The input and target variables used for binary genetic programming (BGP) and Decision tree (DT) modeling of Bhopal bus rapid transit line (population density (PD), land use diversity (LD), walkable catchment area (WC), walkable path (WP), travel behavior (TB), mixed-use of land (MUL), and economic indicator (EI)).

ID	Station name	PD	LD	WC	WP	TB	MUL	EI	TOD	Binominal label
1	New Market	0.28	1	0.91	0.9	0.63	0.51	0.53	0.74	TOD**
2	Board	0.22	0.84	1	0.81	0.71	0.59	0.54	0.71	TOD
3	New Ashok	0.8	0.87	0.22	0.89	0.76	1	0.64	0.7	TOD
4	Piplani	0.19	0.77	0.63	1	0.66	0.59	0.74	0.65	TOD
5	Railway Station	1	0.78	0.21	0.66	0.8	0.51	0.54	0.63	TOD
6	People Mall	0.26	0.53	0.95	0.64	0.75	0.32	0.73	0.62	TOD
7	Karound	0.21	0.78	0.54	0.88	0.78	0.86	0.58	0.62	TOD
8	Sai	0.35	0.76	0.31	0.95	0.67	0.71	0.67	0.6	TOD
9	Habibganj	0.2	0.7	0.48	0.85	0.63	0.47	0.62	0.56	Non-TOD*
10	Gandhi	0.1	0.52	0.38	0.87	0.77	0.35	0.67	0.48	Non-TOD*
11	Kohe Fiza	0.2	0.66	0.21	0.62	0.62	0.5	0.78	0.48	Non-TOD
12	Beema	0.25	0.53	0.09	0.8	0.63	0.83	0.63	0.46	Non-TOD
13	Halalpur	0.08	0.48	0.38	0.71	0.77	0.35	0.69	0.45	Non-TOD
14	Aura	0.23	0.39	0.14	0.85	0.64	0.74	0.69	0.44	Non-TOD
15	C21	0.08	0.55	0.39	0.31	0.77	0.2	0.68	0.4	Non-TOD
16	Ashima	0.09	0.33	0.25	0.68	0.58	0.57	0.68	0.39	Non-TOD

*Negative Label; **Positive label.

Table A.2. The Fold-1 subset used to develop binary genetic programming and decision tree models (an example of Antalya tram line; (population density (PD), land use diversity (LD), walkable catchment area (WC), walkable path (WP), and travel behavior (TB)).

ID	PD	LD	WC	WP	TB	TOD
21	0.363	0.800	0.758	1.000	0.371	1
4	0.458	0.531	0.559	0.579	0.957	1
8	0.547	0.364	0.399	0.581	0.800	0
16	0.148	0.762	0.218	0.613	0.657	0
13	0.419	0.381	0.377	0.481	0.571	0
20	0.337	0.880	1.000	0.956	0.686	1
9	0.057	1.000	0.351	0.378	0.600	0
10	0.451	0.837	0.261	0.724	0.729	1
1	0.033	0.144	0.203	0.302	1.000	0
12	0.536	0.585	0.365	0.570	0.586	0
5	0.479	0.325	0.541	0.824	0.786	1
11	1.000	0.475	0.460	0.640	0.600	1
18	0.298	0.805	0.372	0.498	0.514	0
7	0.494	0.837	0.744	0.540	0.629	1
14	0.673	0.517	0.346	0.389	0.643	0
3	0.131	0.862	0.195	0.354	0.943	0
15	0.462	0.642	0.383	0.492	0.557	0
17	0.306	0.869	0.853	0.603	0.457	1
2	0.066	0.462	0.489	0.438	1.000	0
19	0.328	0.787	0.268	0.487	0.586	0
6	0.675	0.502	0.345	0.689	0.743	1

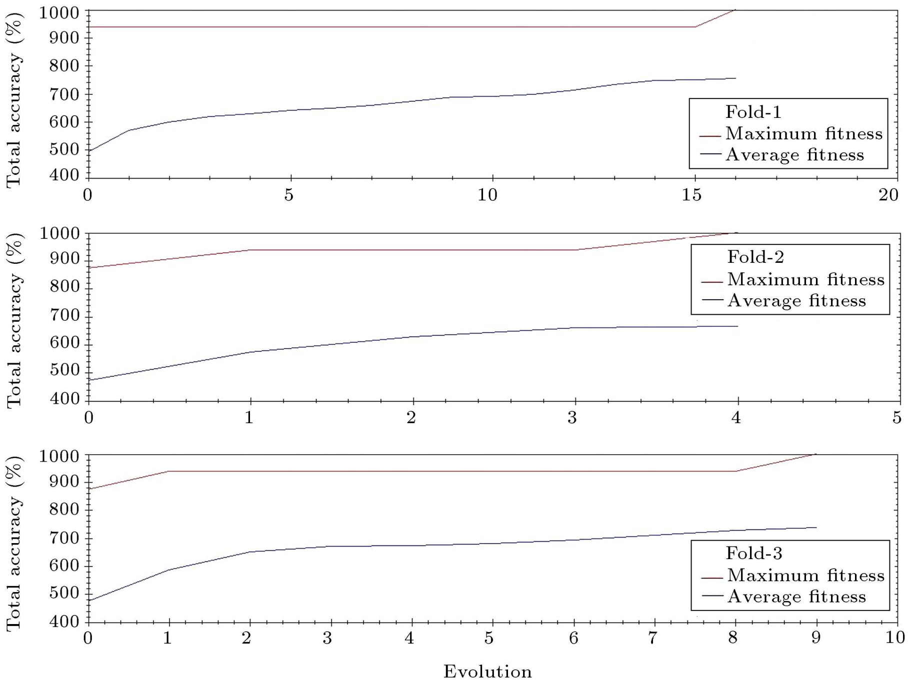


Figure B.1. The binary genetic programming fitness simulation for TOD area prediction in Varsak-Otogar tramway line.

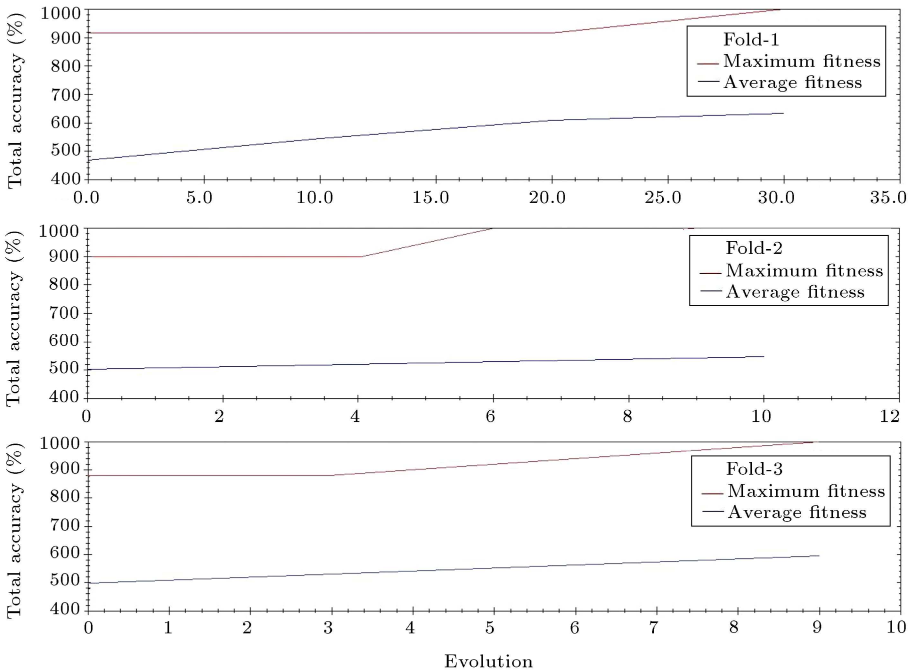


Figure B.2. The binary genetic programming mode; fitness simulation for TOD area prediction along Bhopal bus rapid transit line.

achieved 100% accuracy at the training phase at all three folds after a few generations.

Biographies

Reza Gholizadeh holds a MSc degree in Architecture from the Islamic Azad University of Tabriz. Currently, he is a PhD candidate in Landscape Architecture at Ataturk University, Turkey. He also serves as a full-time faculty at Jolfa International Branch of Azad University, Iran. Mr. Gholizadeh published several scientific papers in national and international conferences. He is interested in urban landscape issues, especially on a neighborhood scale.

Hasan Yılmaz graduated from Atatürk University, Faculty of Agriculture, Department of Horticulture in 1987 and started to work as a Research Assistant in the same year. In 1994, he completed his doctorate at Ege University, Faculty of Agriculture, Department of Landscape Architecture. He received the title of Assistant Professor in 1995, Associate Professor in 2001, and Professor in 2006 from Atatürk University, Faculty of Agriculture, Department of Landscape Architecture. During 2010–2016, Professor. Yilmaz acted as the Dean of Faculty of Architecture and Design at Atatürk University. He also served as a project consultant and international honorary jury member at EXPO2016 Antalya, one of the most important projects in Turkey.

Ali Danandeh Mehr holds a PhD degree in Civil Engineering from Istanbul Technical University (Turkey). Currently, he is an Associate Professor at the Civil Engineering Department of Antalya Bilim University, Turkey. He has also worked as a postdoctoral researcher at the University of Tabriz (Iran) and the University of Oulu (Finland). Dr. Danandeh Mehr has published more than 60 refereed publications and collaborated as an editor/reviewer for different international scholarly journals. His current research focuses on stochastic hydrology, hydroinformatics, and developing evolutionary algorithms to solve different problems in water resources systems.