



Modelling of Activity-Travel Pattern with Support Vector Machine

Anu P. Alex ^{1*}, Manju V. S. ², Kuncheria P. Isaac ³

¹Associate Professor, College of Engineering Trivandrum, Kerala, India-695016

²Professor, College of Engineering Trivandrum, Kerala, India-695016

³Vice Chancellor, Hindustan Institute of Technology and Science, Chennai, India-603103

Abstract

Activity based travel demand modelling involves lot of uncertainty due to the complex and varying decision making behaviour of each individual. This study contributes to the literature by assessing the suitability of Support Vector Machine (SVM) in modelling the activity pattern and travel behaviour of workers. Activity and travel behaviour of workers consists of decision outcomes, which can be modelled as classification and regression problems. SVM is a good classifier and regressor with good testing and learning capability, hence the present study used SVM for modelling. It was found that support vector machine models are well performing to predict the activity pattern and travel behaviour of workers. The SVM models developed in the study predicts the temporal variation of mode wise work activity generation. Prediction of temporal mode share of commuters is advantageous to policy makers to experiment the implementation of temporary Travel Demand Management (TDM) actions effectively.

Keywords: Transportation planning, mathematical modelling, Support Vector Machine, activity-travel pattern, activity based travel demand modelling

1. Introduction

The idea of demand and supply are elementary to economic theory and is extensively used in the field of transport economics. The concept could be applied in the area of travel demand and the supply of transport infrastructure. Modelling of demand implies a procedure for predicting the travel decisions of people. It is used to predict the future travel patterns and demands based on the changes in the transportation system, land use and demographic features. There are mainly two approaches in travel demand modelling; traditional statistically-oriented trip based modelling approach and more behaviourally-oriented activity-based modelling approach. The activity-based approach views travel as a derived demand; derived from the need to pursue activities distributed over space (Axhausen and Gärling, 1992; Kitamura, 1988; Pas, 1985). This approach concentrates on the chains of activities, with time segments as the analysis unit. By predicting what activities are performed when and where, activity based models forecast the trips, their timing and locations. The various approaches for activity-based modelling are

* Corresponding author: Anu P. Alex (anualex@cet.ac.in)

constraints-based modelling, computational process modelling, micro simulation modelling and utility-maximizing models. The roots of constraints-based models are in time geography, whereas, computational process models are motivated by psychological decision process theories and utility-maximizing models are based on microeconomic theory (Timmermans et al. 2002). These studies exposed that activity based travel demand modelling involves lot of uncertainty due to the complex and varying decision making behaviour of each individual. The selection of modelling technique is very important due to this complexity. This leads to a research investigating the modelling of activity based travel demand using machine learning technique, which is very much fruitful to model the human behaviour. Machine learning paradigm, based on the theory of statistical learning, is a rapid advancement in data mining in recent decades. This has focussed engineering research towards modelling of natural phenomena in a realistic nature. Machine learning techniques can learn from data, identify patterns and make decisions with minimum human intervention. The iterative aspect of machine learning is very much important, since they learn from past computations, produce reliable and most accurate decisions. By developing precise models, the system has a better chance of avoiding unknown risks.

Support Vector Machine is a supervised machine learning technique, which has become extremely popular during the past years due to its applicability in classification and regression problems. Kernel trick is the major strength of SVM and it is not solved for local optima as in neural networks. The risk of overfitting is also less in SVM. This technique is successfully applied in various fields viz; protein structure prediction, intrusion detection, handwriting recognition, detecting steganography in digital images and breast cancer diagnosis. It has been successfully applied to model different civil and other engineering systems also (Dibike et al., 2001; Chuan and Yu, 2011; Yang and Zhao, 2013; Kim, 2014; Lam et al., 2010; Liu and Yao, 2017; Gui et al., 2017; Weng, 2018). In this context, the present study contributes to the literature by modelling the activity-travel pattern of workers using SVM, since this modelling consists of classification and regression problems, in which commuters decide the choices among outcomes and travel modes.

2. Activity-travel pattern of workers

The study area selected to model the activity pattern and travel behaviour of workers is Thiruvananthapuram Corporation, the capital city of Kerala, the southernmost state in India. The city has considerable working population and they depend on public transport system and own vehicles to commute. Questionnaire survey was the main source of primary data collection. As it is not feasible to interview the whole population, sampling was resorted to. It was ensured that the sample was a true representation of the population. A total of 9530 persons were interviewed from 2521 households, which includes people from various socio-economic backgrounds. 41% of the sample population were workers, which included the employees in Government sector, private sector and self-employed. The distribution of activity-travel patterns of workers, obtained from the collected data, shows that 97.51% of the workers perform either 'work activity only' or 'work activity + one other activity (before/ after/ during work or other activity stop during work to home commute)' patterns. Hence, only these two activity-travel patterns of workers are modelled in this study. These two patterns are represented in Fig. 1. These patterns are modelled by identifying a number of decision variables as shown in Fig.1. The decision

variables include whether a person goes for out-of-home work activity, if so, the duration and start time of work, mode and duration of commute for work; whether the working person goes to perform other activities, if so, the purpose of other activity, time of occurrence and duration of that activity, mode used to reach the destination of that activity, home stay duration between work activity and other activity; whether the commuter stops during work-to-home commute, if so, the purpose and duration of stop (Alex *et al.*, 2019).

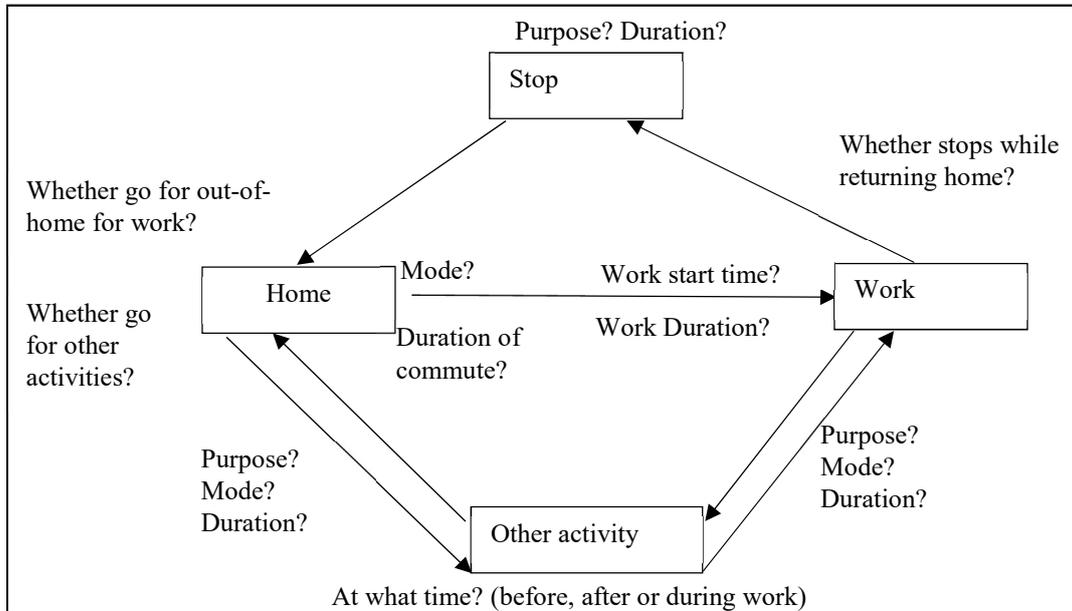


Figure 1: Activity Pattern and Travel Behaviour of Workers

3. Modelling activity-travel pattern using SVM

Support Vector Machine, based on statistical learning theory, was developed by Vladimir N. Vapnik in 1995. It explains the theory of decision planes with boundaries. Decision plane divides a group of objects with various class memberships. SVM employs an iterative training algorithm to minimise an error function. This method is explained in detail in Cortes and Vapnik (1995); Boser *et al.* (1992); and Samui P. (2008). SVM method is usually used for pattern recognition problem for binary outcomes. It can be extended to multi-class problems, where more than two classes are involved. Most of the studies use one-versus-others method for multi-classification problems. SVM can also be used to solve non-linear regression estimation by introducing ϵ insensitive loss function (Vapnik 1995).

This study modelled the activity-travel pattern of workers represented in Fig. 1 using SVM. The decision variables with two discrete outcomes were modelled with binary SVM, that with more than two discrete outcomes were modelled with multi-class SVM, and that with continuous outcome was modelled with SVM regression. The models are discussed in the following sections. There is evidence from the literature regarding the relationship between socio-demographics, activity participation and travel behaviour. Household size, vehicle ownership, employment, age, gender, household income, number of workers in the household, presence of children, level of education, marital status, number of adult members in household, purpose of activity, distance to the activity place, duration of activity, duration of commute are some of the influencing socio-demographic

and activity based factors of activity participation and travel behaviour of individuals (Bhat, 2004; Yagi and Mohammadian, 2009). Variables with high correlations were eliminated and input parameters were finalized based on performance of the models. Notations used for influencing parameters, which are in the developed models are given in Table 1.

Table 1: Parameters and Notations

<i>Notation</i>	<i>Variables</i>	<i>Notation</i>	<i>Variables</i>
H	Household size	mw	Mode of commute for work
vo	Vehicle ownership	dc_w	Duration of commute for work
g	Gender	pur_oth	Purpose of other activities
ag	Age group	stop	Whether stop during work-to-home commute
ms	Marital status	pur_stop	Purpose of stop
emp	Employment	t_oth	Time of occurrence of other activities
edu	Level of education	m_oth	Mode used for other activities
HH	Head of household	d_oth	Duration of other activities
wd	Work duration	st_oth	Home stay duration
wst	Work start time	dur_stop	Duration of stop
d	Distance to work		

3.1 Binary SVM

A classification problem in binary is assessed with a group of training vectors (V) given in Eq. (1), which belongs to two separate classes.

$$V = \{(x_1, y_1), (x_2, y_2), \dots, \dots, \dots, (x_N, y_N)\} \quad (1)$$

$$x \in R^N, y \in \pm 1$$

Where $x \in R^N$ is an N-dimensional data vector with each sample belongs to either of the two classes, $y \in \pm 1$. The objective is to obtain a generalized classifier which can classify the two groups (-1, +1) from the training set. In the current context of classifying decision variables, ‘selection of out-of-home work activity’ and ‘whether stop during work-to-home commute’, two classes were labeled as (-1, +1) and it means (yes, no). The ‘selection of out-of-home work activity’ depends on household size, vehicle ownership, employment, gender, marital status, age group, level of education, and whether the person is head of the household. ‘Whether stop during work-to-home commute’ depends on gender, marital status, mode of commute for work, duration of commute for work, distance to work place, age group, vehicle ownership, employment, level of education. So, $x = [H, vo, g, ag, ms, emp, edu, HH]$ is the input data set used for selection of out-of-home work activity’ and $x = [g, ms, mw, dw, d, ag, vo, emp, edu]$ is the data set used for ‘whether stop during work-to-home commute’. This means a linear hyper plane defined by Eq. (2) can classify the two classes of decision variables, where $w \in R^n$ determines the orientation of the hyperplane and $b \in R$ is a bias. Pictorial representation of a hyperplane is given in Fig. 2.

$$f(x) = w \cdot x + b = 0 \quad (2)$$

The separating hyperplane for the two classes can be defined as Eq. (3).

$$\begin{aligned} y_i(w \cdot x_i + b) &\geq 1 - \xi_i && \text{(for } y_i = 1, \text{ Yes)} \\ y_i(w \cdot x_i + b) &\leq 1 - \xi_i && \text{(for } y_i = -1, \text{ No)} \end{aligned} \quad (3)$$

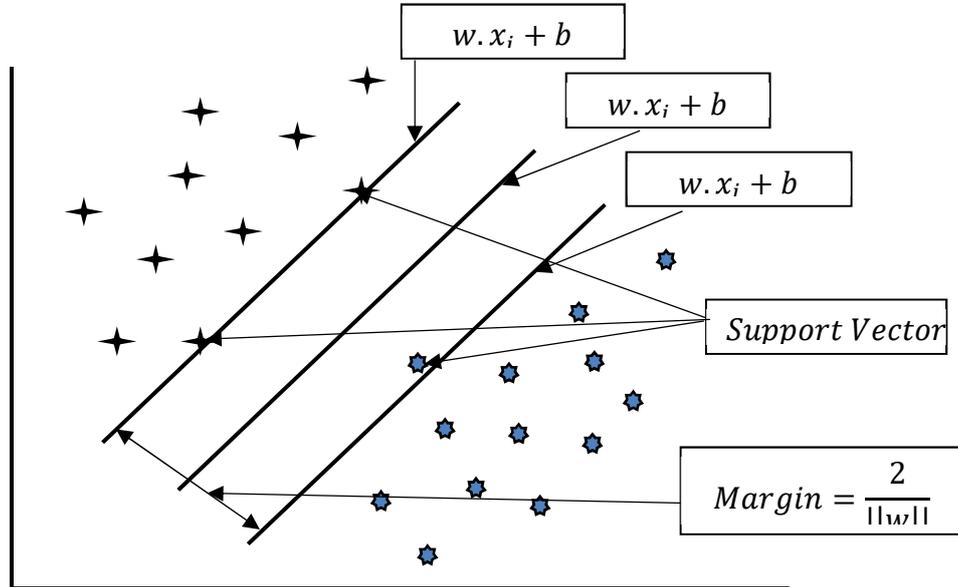


Figure 2: Concept of SVM classification

Where, $\xi_i > 0$ are slack variables, which are used to incorporate the effects of misclassification. The perpendicular distance between origin and the plane $w \cdot x_i + b = -1$ is $\frac{|1+b|}{\|w\|}$ and that to the plane $w \cdot x_i + b = 1$ is $\frac{|b-1|}{\|w\|}$. The margin ($\partial(w,b)$) between the planes is given by (Eq(4)).

$$\partial(w,b) = \frac{2}{\|w\|} \quad (4)$$

The margin between two classes will be maximised and optimal hyperplane is located at the maximum (Fig. 2). Thereby error is minimized. This leads to the constrained optimization problem given in Eq. (5).

$$\begin{aligned} \text{Minimize: } Z &= \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ \text{Subjected to: } &y_i(w \cdot x_i + b) \geq 1 - \xi_i \end{aligned} \quad (5)$$

The constant C, which lies between 0 and ∞ , explains the transaction between the misclassifications in the training set and the maximization of margin. A large value of C minimizes error with lower generalization, whereas a small value of C minimizes error with higher generalization. When C goes to infinitely large, it results in a complex SVM model without allowing the occurrence of any error. If C goes to zero, the result would be less complex model with large amount of errors (Samui P., 2008).

The optimization problem given in Eq (5) can be solved with following Lagrangian construction (Eq (6)) with multipliers α and β .

$$L(w, b, \alpha, \beta, \xi) = \frac{\|w\|^2}{2} + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i \{[(w \cdot x_i + b)]y_i - 1 + \xi_i\} - \sum_{i=1}^N \beta_i \xi_i \quad (6)$$

Lagrangian function has to be minimized with respect to w , b and ξ to obtain the saddle point. By differentiating $L(w, b, \alpha, \beta, \xi)$ with respect to w , b and ξ and equating to zero, three conditions given in Eq. (7) were obtained:

$$\left. \begin{aligned} \frac{\partial L(w, b, \alpha, \beta, \xi)}{\partial w} = 0, & \quad \text{gives} \quad w = \sum_{i=1}^N \alpha_i y_i x_i \\ \frac{\partial L(w, b, \alpha, \beta, \xi)}{\partial b} = 0, & \quad \text{gives} \quad \sum_{i=1}^N \alpha_i y_i = 0 \\ \frac{\partial L(w, b, \alpha, \beta, \xi)}{\partial \xi} = 0, & \quad \text{gives} \quad \alpha_i + \beta_i = C \end{aligned} \right\} \quad (7)$$

Hence from Eqns. (6) and (7) the optimization problem can be expressed as in Eq. (8) (Osuna et al, 1997).

$$\left. \begin{aligned} \text{Maximize: } & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\ \text{Subjected to: } & \sum_{i=1}^N \alpha_i y_i = 0 \text{ and } 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, N \end{aligned} \right\} \quad (8)$$

Lagrange multipliers would be obtained by solving Eq. (8) with constraints. The multipliers may be non-zero or zero. Some of the multipliers will be zero according to the Karush–Kuhn–Tucker (KKT) optimality condition (Fletcher 1987). The non-zero multipliers are called support vectors, which are the data points lie on the optimal hyperplane and are difficult to classify. The values of w and b are obtained from $w = \sum_{i=1}^N y_i \alpha_i x_i$ and $b = -\frac{1}{2} w [x_{+1} + x_{-1}]$, where x_{+1} and x_{-1} are the support vectors of classes +1 (Yes) and -1 (No), respectively. The classifier can be formed as in Eq. (9):

$$f(x) = \text{sign}(w \cdot x + b) \quad (9)$$

Where $\text{sign}(\cdot)$ is the signum function. It gives -1 (No), if the element is less than zero and +1 (Yes), if it is greater than or equal to zero. If linear hyper plane is inappropriate, SVM maps input data into a high-dimensional feature space through some non-linear mapping (Boser et al. 1992). When x is replaced by its mapping in the feature space $\phi(x_i)$, the optimization problem becomes Eq. (10).

$$\left. \begin{aligned} \text{Maximize: } & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (\phi(x_i) \cdot \phi(x_j)) \\ \text{Subjected to: } & \sum_{i=1}^N \alpha_i y_i = 0 \text{ and } 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, N \end{aligned} \right\}$$

Instead of feature space $\Phi(x_i)$, Kernel function $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ has been used to lessen computational requirement (Cristianini and Shawe-Taylor 2000; Cortes and Vapnik 1995). Radial basis function was used as kernel functions in the current study. The above discussed binary classification model was implemented in this study to model the decision outcomes of activity pattern and travel behaviour of workers. i.e; to predict whether a person goes for out-of-home work activity (yes/no) and to predict whether he/she stops during work to home commute (yes/no). The decision variables and classes are given in Table 2.

Table 2: Decision Variables and Outcomes of Binary SVM Models

<i>Sl. No.</i>	<i>Decision Variables</i>	<i>Outcome</i>	<i>Classes</i>	<i>SVM Model Type</i>
1	Whether go for out-of-home work activity	Yes No	+1 -1	Binary Classification
2	Whether stop during work-to-home commute	Yes No	+1 -1	Binary Classification

Separate training data and testing data sets were used for constructing and testing the performance of the models. Different values of hyper-parameters were tried and the parameters that yielded the best accuracy on the models were used. The hyper-parameters which gave better cross validation results and minimum number of support vectors were capacity constant, $C=10$ and $\epsilon=0.10$. Kernel function used for transferring the data points to feature space was Radial Basis Function (RBF). The input parameters which gave better validation results were also selected by trial and error. The selected input parameters are shown in Table 4.

3.2 Multi-Class SVM

Higher level multi-class SVM methods are usually used for more than two classes, which utilizes binary classification as the basics. ‘One-versus-others method’ was used in this study. It is an effective and simple method for multi-class classification problems (Dubchak *et al.*, 1999). Assume there are M classifications in the problem. It is converted into a binary classification problem, i.e; one class contains datapoints in one ‘true’ class and the ‘others’ class combines data in all other categories. A binary classifier is trained for this two-class problem. Then the M classes are partitioned into another binary classification problem. The same process is replicated for each of the M classes, which results in M binary trained classifiers.

Multi-class SVM method is used in the present study to model the decision variables with more than two outcomes. i.e; to predict the mode of commute for work, purpose of other activities, time of occurrence of other activities, mode used for other activities and purpose of stop. The decision variables and classes used for each are given in Table 3. As in binary classification, a training data set and testing data set were used for constructing and testing the models. The hyper-parameters which gave better cross validation results and minimum number of support vectors were capacity constant, $C=10$ and $\epsilon=0.10$. Kernel function used for transferring the data points to feature space was Radial Basis Function (RBF). The input parameters which gave better validation results are shown in Table 4.

Table 3: Decision Variables and Outcomes of Multi-class SVM Models

<i>Sl. No.</i>	<i>Decision Variables</i>	<i>Outcome</i>	<i>Classes</i>	<i>SVM Model Type</i>
1	Mode of commute for work	Walk / cycle	0	Multi-class SVM
		Two wheeler	1	
		Car	2	
		Bus	3	
		Train	4	
2	Purpose of other activity	No activity	0	Multi-class SVM
		Personal business and recreation	1	
		Shopping	2	
		Eat out	3	
3	Purpose of stop	Care for children/spouse/elderly	0	Multi-class SVM
		Shopping	1	
		Personal business and recreation	2	
4	Time of occurrence of other activities	Before work	0	Multi-class SVM
		After work	1	
		Work based	2	
5	Mode used for other activities	Walk / cycle	0	Multi-class SVM
		Two wheeler	1	
		Car	2	
		Bus	3	

Table 4: Input Variables and Performance of SVM Models

<i>Sl. No</i>	<i>Attributes</i>	<i>Input Variables</i>	<i>Training Perf. (%)</i>	<i>Testing Perf. (%)</i>
1	Whether go for out-of-home work activity	H, vo, g, ag, ms, emp, edu, HH	99.23	97.15
2	Work duration	H, vo, g, ag, ms, emp, edu	82.35	81.65
3	Work start time	wd, g, ag, edu	88.42	87.11
4	Mode of commute for work	vo, g, ag, ms, emp, HH, wd, wst	73.25	70.41
5	Duration of commute for work	wd, wst, d, vo, g, ag, emp, mw	75.36	74.13
6	Purpose of other activities	wd, dc_w, d, H, vo, emp, mw, HH, edu, ag, g, ms	75.80	74.25
7	Whether stop during work-to-home commute	g, ms, mw, dc_w, d, ag, vo, emp, edu	95.12	94.47
8	Purpose of stop	dc_w, emp, ag, H, g, ms	92.45	91.26
9	Time of occurrence of other activities	wd, wst, vo, ms, emp, edu, mw, pur_oth	85.28	84.89
10	Mode used for other activities	vo, g, ag, pur_oth, t_oth	75.87	73.57
11	Duration of other activities	wd, pur_oth, m_oth, t_oth	84.76	83.14
12	Home stay duration	t_oth, wst, dc_w, pur_oth, H, wd, vo, g, d_oth, mw, emp, ag, ms	82.51	81.90
13	Duration of stop	wd, dc_w, d, vo, g, ag, edu, pur_stop	90.52	89.48

3.3 SVM Regression

The dependence of a variable y on a set of independent variables x can be estimated in SVM regression using an alternative loss function (Smola, 1996). Regression estimates are found out by a set of linear functions. These linear functions are defined in a high dimensional space in the first phase (Samui, P., 2008). Using Vapnik's ϵ -insensitive loss function, risk is measured and it is minimized in the second step for carrying out estimation of regression. In the third step, regression SVM uses a risk function, which consists of a regularization term and empirical error. It is obtained from structural risk minimization (SRM) principle (Samui P., 2008). The concept of SVM regression is shown in Fig. 3.

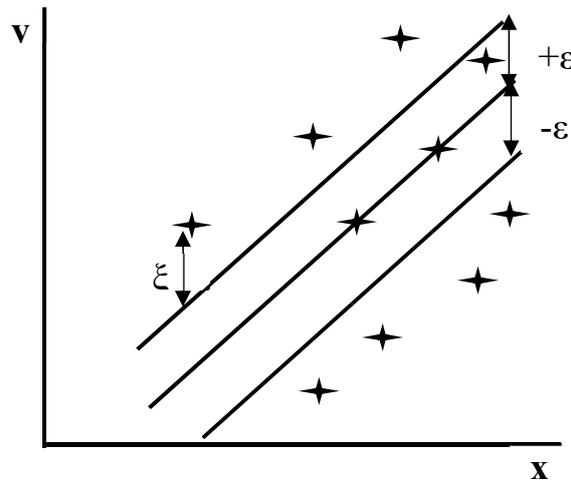


Figure 3: Concept of SVM Regression

Assume that the dataset consists of $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, where, x is the input and y is the output. Let us assume a linear function Eq. (11).

$$f(x) = (w \cdot x) + b \tag{11}$$

where, $w \in R^N$ and $b \in r$; $R^N = N$ -dimensional vector space; b = scalar threshold; w = adjustable weight vector; and r = one dimensional vector space.

The objective of SVM regression is to find a function $f(x)$ that gives a difference ϵ from the true output (y). The optimization problem involves Eq. (12) and (13):

$$\text{Minimize: } Z = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \tag{12}$$

$$\left. \begin{aligned} \text{Subjected to: } & y_i - \{(w \cdot x_i + b)\} \leq \epsilon + \xi_i, i = 1, 2, 3, \dots, N \\ & \{(w \cdot x_i + b)\} - y_i \leq \epsilon + \xi_i^*, i = 1, 2, 3, \dots, N \\ & \xi \geq 0 \text{ and } \xi_i^* \geq 0, i = 1, 2, 3, \dots, N \end{aligned} \right\} \tag{13}$$

The optimization problem given in Eq. (12) can be solved with the following Lagrangian construction (Eq (14)) with multipliers α, α^*, γ and γ^* .

$$\begin{aligned} L(w, \xi, \xi^*, \alpha, \alpha^*, \gamma, \gamma^*) = & \frac{\|w\|^2}{2} + C \left[\sum_{i=1}^N (\xi_i + \xi_i^*) \right] - \sum_{i=1}^N \alpha_i [\epsilon + \xi_i - y_i + (w \cdot x_i + b)] \\ & - \sum_{i=1}^N \alpha_i^* [\epsilon + \xi_i^* + y_i - (w \cdot x_i) - b] - \sum_{i=1}^N \gamma_i \xi_i + \gamma_i^* \xi_i^* \end{aligned} \tag{14}$$

Lagrangian function has to be minimized with respect to w , b , ζ and ζ^* to obtain the saddle point. By differentiating $L(w, \xi, \xi^*, \alpha, \alpha^*, \gamma, \gamma^*)$ with respect to w , b , ζ and ζ^* and equating to zero, the following four conditions were obtained Eq. (15):

$$\left. \begin{aligned} \frac{\partial L}{\partial w} = 0, \text{ gives } w &= \sum_{i=1}^N x_i(\alpha_i - \alpha_i^*) \\ \frac{\partial L}{\partial b} = 0, \text{ gives } \sum_{i=1}^N \alpha_i &= \sum_{i=1}^N \alpha_i^* \\ \frac{\partial L}{\partial \xi} = 0, \text{ gives } \sum_{i=1}^N \gamma_i &= \sum_{i=1}^N (C - \alpha_i) \\ \frac{\partial L}{\partial \xi^*} = 0, \text{ gives } \sum_{i=1}^N \gamma_i^* &= \sum_{i=1}^N (C - \alpha_i^*) \end{aligned} \right\} \quad (15)$$

Hence from Eqs. (14) and (15) the optimization problem becomes Eq. (16) and (17):

$$\text{Maximize: } -\varepsilon \sum_{i=1}^N (\alpha_i^* + \alpha_i) + \sum_{i=1}^N y_i (\alpha_i^* - \alpha_i) - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i^* - \alpha_i)((\alpha_j^* - \alpha_j)(x_i \cdot x_j)) \quad (16)$$

$$\text{Subjected to : } \sum_{i=1}^N \alpha_i = \sum_{i=1}^N \alpha_i^*, 0 \leq \alpha_i \leq C, \text{ and } 0 \leq \alpha_i^* \leq C \quad i = 1, 2, \dots, N \quad (17)$$

The coefficients α_i and α_i^* can be found out by resolving the optimization problem in Eq. (16). Zero Lagrange multipliers means sparseness in the training objects. Training objects with non-zero Lagrange multipliers are the support vectors. Prediction errors of these objects will be larger than $\pm\varepsilon$. In the case of inappropriateness of linear hyper plane, some non-linear mapping is done in SVM, where input data is mapped into a high-dimensional feature space (Boser *et al.* 1992). Optimization problem becomes Eq. (18) and (19), when feature space $\phi(x_i)$ replaces x by its mapping.

$$\begin{aligned} \text{Maximize: } -\varepsilon \sum_{i=1}^N (\alpha_i^* + \alpha_i) + \sum_{i=1}^N y_i (\alpha_i^* - \alpha_i) \\ - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i^* - \alpha_i)((\alpha_j^* - \alpha_j)(\Phi(x_i) \cdot \Phi(x_j))) \end{aligned} \quad (18)$$

$$\text{Subjected to : } \sum_{i=1}^N \alpha_i = \sum_{i=1}^N \alpha_i^*, 0 \leq \alpha_i \leq C, \text{ and } 0 \leq \alpha_i^* \leq C \quad i = 1, 2, \dots, N \quad (19)$$

Kernel function $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ is introduced as a replacement for feature space $\Phi(x_i)$ to lessen the computational requirement (Cortes and Vapnik 1995; Cristianini and Shawe-Taylor 2000). Fig. 4 shows the architecture of SVM regression.

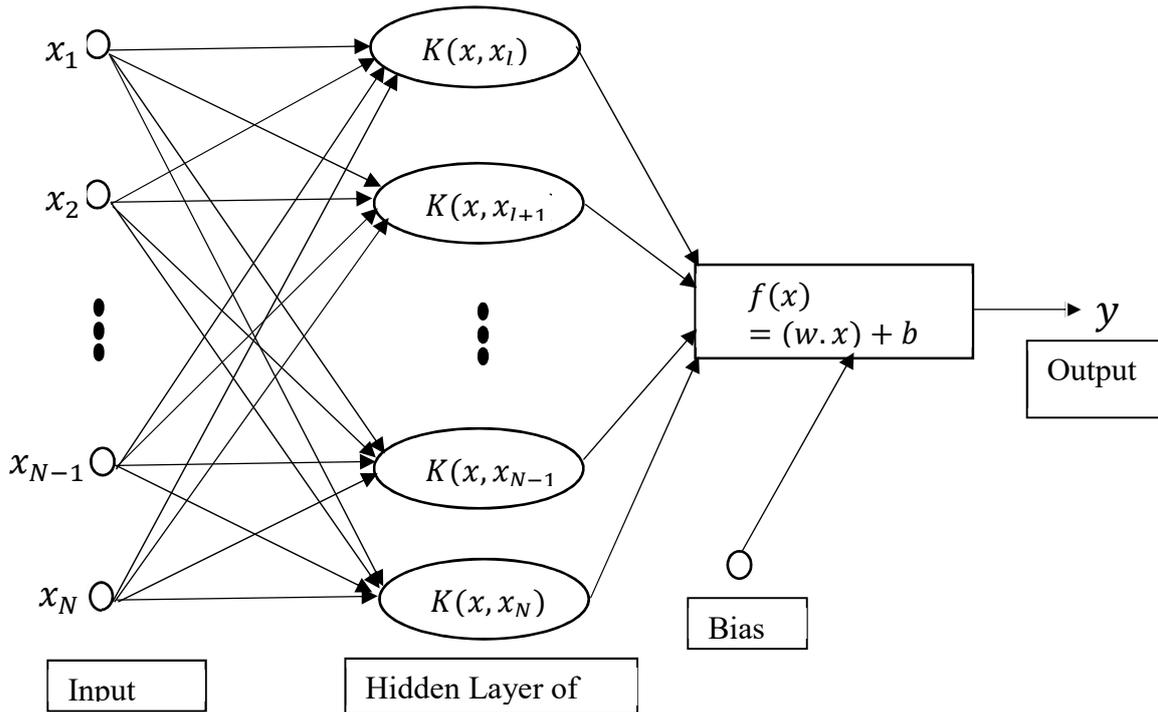


Figure 4: Architecture of SVM Regression

The above discussed SVM regression model was applied in this study to model the work duration, work start time, duration of commute for work, duration of other activities, home stay duration and duration of stop. A training data set and testing data set are used for constructing the model and testing the performance. The hyper-parameters which gave better cross validation results and minimum number of support vectors were capacity constant, $C= 10$ and $\epsilon = 0.10$. Kernel function used was Radial Basis Function (RBF). The input parameters which gave better validation results are shown in Table 4.

4. Results and discussion

Model performances are represented in terms of training performance and testing performance in SVM models. The training and testing performance of the classification models are found out using Eq. (20) and those of regression models are calculated using Eq. (21). Table 4 represents the performance of the models. Higher training and testing performance shows that the models are good. The SVM classification and regression models were validated, to evaluate their performance, with new set of data, which are not used for training. The performance of the SVM classifications models, given in Table 5, shows that the prediction accuracy of the SVM classification models are above 70%. The performance of regression models are shown in Figs. 5 to 10, in which actual and predicted values are plotted. The correlation coefficient (r) obtained in all the cases are greater than 0.90. Hence it is proved that the developed models are capable of predicting the activity-travel pattern of workers. This is due to the training and learning capability of SVM.

$$\begin{aligned} &\text{Training or Testing performance (\%)} \text{ of classification models} \\ &= \left(\frac{\text{Number of data predicted accurately}}{\text{Total number of data}} \right) * 100 \end{aligned} \quad (20)$$

Training or Testing performance (%) of regression models

$$= 100 - \left(\frac{\text{Actual value} - \text{Predicted Value}}{\text{Actual Value}} \right) * 100 \quad (21)$$

Table 5: Results of Validation of SVM Classification Models

Attributes	Sample Size	SVM Models	
		Correctly Predicted	Prediction Accuracy (%)
Whether go for out-of-home work activity	777	753	96.88
Mode of commute for work	700	494	70.57
Purpose of other activities	700	523	74.67
Whether stop during work-to-home commute	700	650	92.87
Purpose of stop	306	281	91.89
Time of occurrence of other activities	259	215	83.12
Mode used for other activities	259	189	73.12

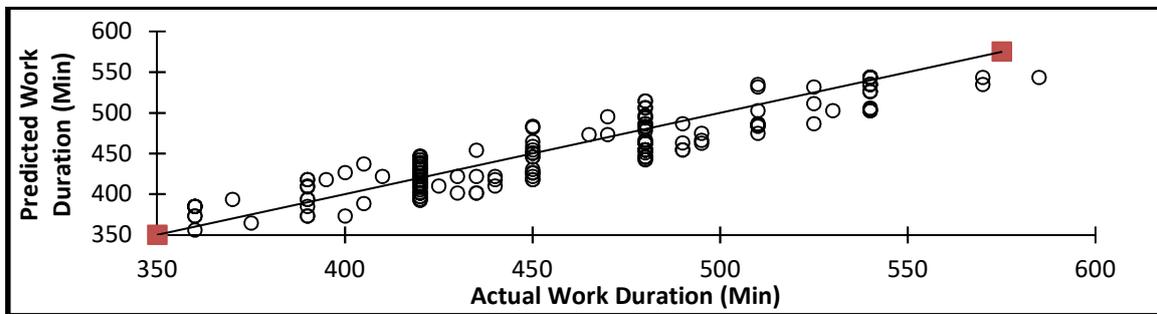


Figure 5: Performance of Work Duration SVM Model

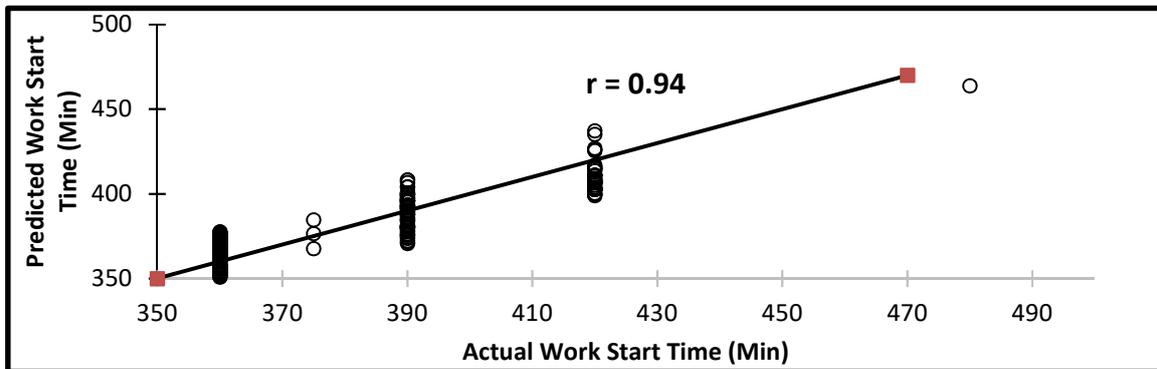


Figure 6: Performance of Work Start Time SVM Model

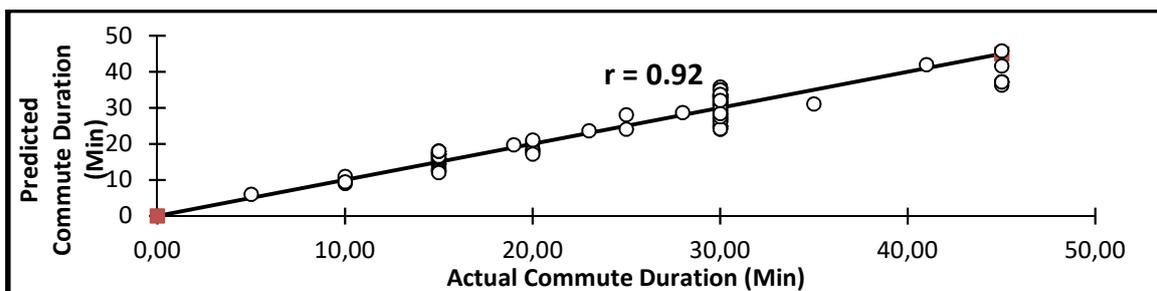


Figure 7: Performance of Duration of Commute for Work SVM Model

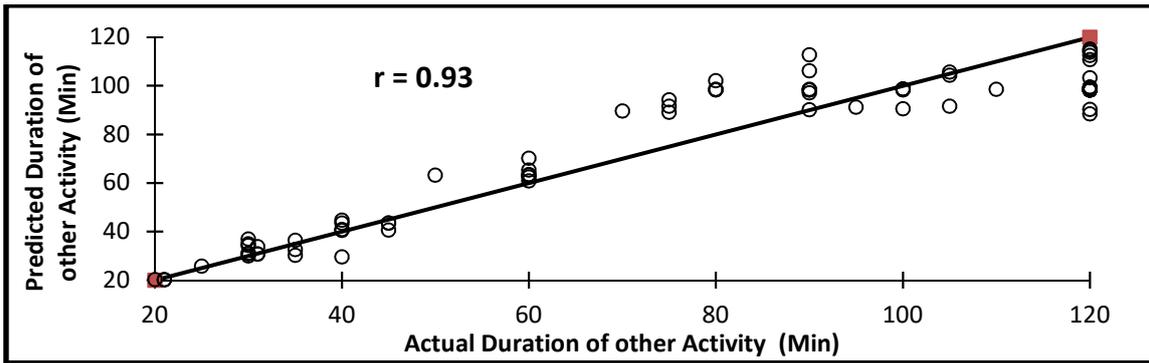


Figure 8: Performance of Duration of Other Activity SVM Model

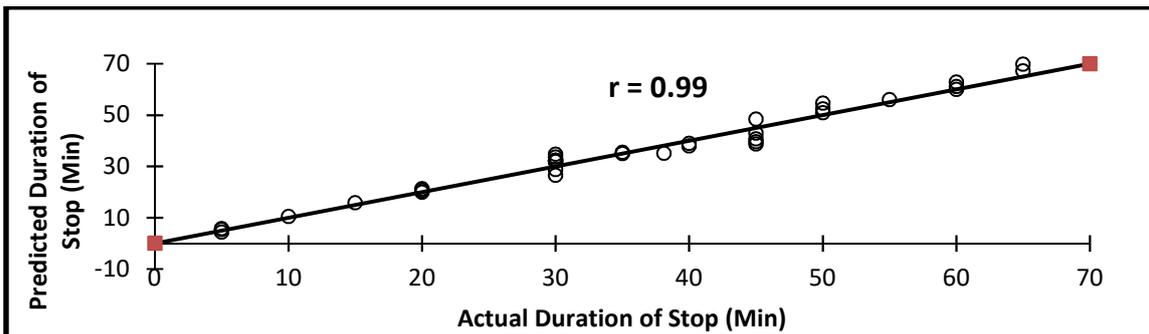


Figure 9: Performance of Duration of Stop SVM Model

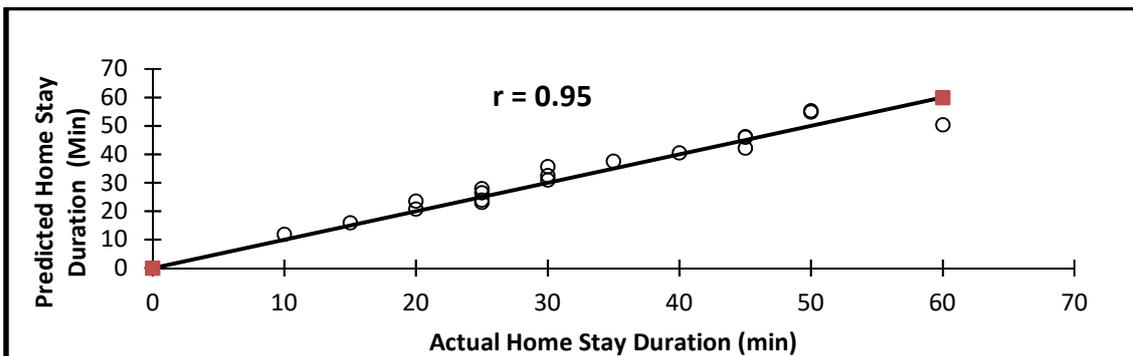


Figure 10: Performance of Home Stay Duration SVM Model

5. Conclusion

Policy makers are now more concerned about implementing short term demand management measures like staggering of school/work hours, promoting cycling etc. Hence, the need for realistic representation of behavior in travel demand modelling is vital today to understand the travel behaviour responses for such short-term congestion management policies. More realistic representation of travel behaviour is possible by activity-based modelling approach. The modelling of activity-travel behaviour of commuters is a complex process and involves lot of uncertainty. It is a pattern recognition problem, where commuters select the choices among decision outcomes and travel mode alternatives. Hence this study has experimented SVM classification and SVM regression for modelling the activity-travel pattern of workers. Different values of hyper-parameters were tried and the parameters that yield the best accuracy on the models were used. The Kernel function used as for transferring the data points to feature space was Radial Basis Function (RBF).

The present study contributes to the literature by developing support vector machine models for activity-travel pattern. It is proved that SVM classification and regression models developed for predicting the activity-travel pattern of workers are well performing. The prediction accuracy of SVM classification models are more than 70% and SVM regression models gave good results ($r > 0.90$). This is due to the training capability and learning ability of SVM. The study proves that SVM is a powerful tool for modelling the activity pattern and travel behaviour of workers. Even though statistical check is not possible for SVM models, it is easy to transfer the model. It can be concluded from the study that the developed SVM models represent real life activity-travel pattern of workers. Hence it can be used for predicting the activity-travel pattern and thereby the modes used by workers. This will eventually help to analyse the temporal mode share of workers.

References

- Alex, A.P., Manju V. S., Isaac, K. P. (2019) “Modelling of Travel Behaviour of Students using Artificial Intelligence”, *Archives of Transport*, Vol. 51. Issue 3, pg: 7 – 19. doi: 10.5604/01.3001.0013.6159
- Axhausen, K. and Gärling, T. (1992) “Activity-based approaches to travel analysis: conceptual frameworks, models and research problems”, *Transport Reviews*, 12, 324-341. doi:10.1080/01441649208716826
- Boser B. E., Guyon I. M., Vapnik V. N. (1992) “A training algorithm for optimal margin classifier” In: Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory, Pittsburg, PA, USA. 27–29
- Chuan, J. and Yu, X. (2011) “Travel mode choice analysis using support vector machines”, 11th International Conference of Chinese Transportation Professionals (ICCTP). pp: 360-371.
- Cortes, C., Vapnik V. N. (1995) “Support-vector networks”. *Machine Learning* 20(3):273–97. doi:10.1007/bf00994018
- Cristianini, N., Shawe-Taylor (2000) *An introduction to support vector machine*. London: Cambridge University press
- Dibike, Y. B., Velickov, S., Solomatine, D. and Michael B. (2001) “Abbott model induction with support vector machines: introduction and applications”, *Journal of Computing in Civil Engineering*, Vol 15 issue 3. doi:10.1061/(asce)0887-3801(2001)15:3(208)
- Dubchak, I., Muchnik, I., Mayor, C., Dralyuk, I., Kim, S. H. (1999) “Recognition of a protein fold in the context of the SCOP classification”, *Proteins: Structure, Function, and Genetics*, vol. 35, no. 4, pp. 401-407. doi:10.1002/(sici)1097-0134(19990601)35:4 < 401: aid-prot3>3.0.co;2-k
- Fletcher, R. (1987) *Practical Methods of Optimization*, John Wiley and Sons, New York.
- Gui, G., Pan, H., Lin, Z., Li, Y., and Yuan, Z. (2017) “Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection”, *KSCCE Journal of Civil Engineering*, 21(2), 523–534. doi:10.1007/s12205-017-1518-5
- Kim, J., Kim, S., and Tang, L., (2014) “Case Study on the Determination of Building Materials Using a Support Vector Machine”, *Journal of Computing in Civil Engineering*, Volume 28, Issue 2 (315-326). doi:10.1061/(asce)cp.1943-5487.0000259

- Kitamura, R. (1988) “An evaluation of activity-based travel analysis.” *Transportation*, 15, 9-34. doi:10.1007/bf00167973
- Lam, K. C., Lam, M. C. K., and Wang, D. (2010) “Efficacy of using support vector machine in a contractor prequalification decision model”, *Journal of Computing in Civil Engineering*, Volume 24, Issue 3 (273 - 280). doi:10.1061/(ASCE)cp.1943-5487.0000030
- Liu, S., and Yao, E. (2017) “Holiday passenger flow forecasting based on the modified least square support vector machine for the metro system”, *Journal of Transportation Engineering, Part A: Systems*, Volume 143(2). doi:10.1061/jtepbs.0000010
- Osuna E., Freund R., and Girosi F. (1997) “An improved training algorithm for support vector machines”. In Principe J., Gile L., Morgan N., and Wilson E. (Eds.), *Neural Networks for Signal Processing VII—Proceedings of the 1997 IEEE Workshop*, pp. 276–285, NY.
- Pas, E. I. (1985) “State of the art and research opportunities in travel demand: another perspective”, *Transportation Research: Part A*, 19A: 460-464. doi:10.1016/0191-2607(85)90048-2
- Samui, P. (2008) “Support vector machine applied to settlement of shallow foundations on cohesionless soils”. *Computers and Geotechnics*, 35(3), 419–427. doi: 10.1016/j.compgeo.2007.06.014
- Smola A. J., Scholkopf B. (2004) “A tutorial on support vector regression”. *StatComput*;14:199–222. doi:10.1023/b:stco.0000035301.49549.88
- Timmermans, H., Arentze, T., Joh, C-H. (2002) “Analysing space-time behaviour: new approaches to old problems.” *Progress in Human Geography* 26(2), 175–190. doi:10.1007/bf00165706
- Vapnik V. N. (1995) *The nature of statistical learning theory*. New York: Springer
- Weng, J., Tu, Q., Yuan, R., Lin, P., (2018) “Modeling mode choice behaviors for public transport commuters in Beijing”, *Journal of Urban Planning and Development*. Volume 144, Issue 3. doi:10.1061/(asce)up.1943-5444.0000459
- Yagi, S., Mohammadian, A. (2009) “An activity-based microsimulation model of travel demand in the Jakarta metropolitan area.” *Journal of Choice Modelling*, 3(1), 32-57. doi:10.1016/s1755-5345(13)70028-9
- Yang, J., and Zhao, J. (2013) “Road traffic safety prediction based on improved SVM”, *Fourth International Conference on Transportation Engineering*. pp:107-114. doi:10.1061/9780784413159

Acknowledgements

The authors acknowledge the support provided by Kerala State Council for Science, Technology and Environment (KSCSTE) Government of Kerala, under the grant ETP/09/2012.