**Journal of Theoretical and Applied Information Technology**
15th May 2020. Vol.98. No 09


ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# AN IMPROVED MACHINE LEARNING-BASED APPROACH FOR PREDICTING TRAVELERS MODE CHOICE IN MOROCCO

**[1]MOHAMED EL HADRAOUI, [2]FOUZIA GHAITI**

[1,2]Modeling and Decision Support in Systems Team, Mohammadia School of Engineers,

Mohammed V University,

Rabat, Morocco

E-mail: [1]med.elhadraoui@gmail.com, [2]ghaiti@emi.ac.ma

## ABSTRACT

Predicting the travel mode choice is an important task of transportation planning and policy making to understand inter-urban mobility. It enables the enhancement of the third step of the widely used four-step model. While advances in machine learning have led to numerous powerful predictive models, their usefulness for modeling travel mode choice remains none widely explored. The aim of this paper is to fill in this gap by proposing an advanced machine learning approach tailored to this problem. That is, using extensive Moroccan travel diary data in the year 2016, enriched with numerous individual and household features, our contribution consists of investigating the importance of applying the feature selection approach while using support vector machines (SVM) as a predictive model. The experimental results show that the adopted approach outperforms both native SVM and the artificial neural network, which are the most common data-driven techniques of dealing with such a problem.

**Keywords:** *Travel mode choice prediction, Support vector machines, Feature selection, Inter-urban mobility*

## 1. INTRODUCTION

For developing countries such as Morocco, the prediction of the travel mode choice is crucial for an accurate modeling and understanding of the inter-urban mobility system. Indeed, the travel mode choice is an important issue for transportation planning and policy making to predict travel demand and understand the underlying factors, by conducting data-based assessments of different programs, policies and infrastructure projects. More precisely, it enables to predict the national mobility within the country and therefore give decision-makers better visibility to decide on the priority of transport infrastructure projects, applicants in terms of investments, and know how to manage different crises (e.g. financial crisis, COVID-19 pandemic …).

Therefore, the performance of travel mode prediction has received growing attention in the last few years. In fact, recent advances in travel data processing have powered many of the recent innovations in inter-urban transportation of different countries. In particular, Morocco has made tremendous progress in the transport sector, as it is one of its main areas of interest.

The integration of machine learning for predicting model choice can be included in various systems, including the one adopted in Morocco, and our goal in this study is to include machine learning to improve the existing system, which is adopted by the Moroccan Department. In fact, the Department of Equipment and Transport makes use of a trip based travel demand model, which is often called a four-step model. That is, the Department has justified its choice as this model is based on socio-economic variables and it can therefore help to simulate the evolution of its parameters and see their impact on the mobility and the modal choice. Looking at this model, in general, the procedure takes a top-down approach, from the decision to travel, to a destination and mode choice, up to ending with the road choice. The four steps are named respectively trip generation, trip distribution, mode choice and trip assignment. This paper is a part of a research project which aims to model the inter-city national mobility by focusing on this third step.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The incorporation of machine learning is crucial, as it is the basis of this third step of this model. In fact, this step, which consists of predicting the choice of travel mode for each traveler based on their personal data in addition to his target. Other works done for instance in France adopt the traditional logit model. However, this approach needs a precise database suitable to this kind of models. However, due to the hardness of obtaining such data in the case of Morocco, and therefore the necessity of taking advantage of machine learning approaches that could enable us to deal with the available data in the most accurate way. Moreover, the issue of predicting the choice of mode of travel is complicated by the fact that it is influenced by various factors, including the characteristics of the individual and the household. Therefore, the construction of an accurate model with traditional methods such as the multinomial logit model (MNL) can be difficult because of the variety and non-linearity of these factors [17]. Thus, methods based on machine learning techniques like artificial neural network [19], and support vector machine (SVM) [29] can give better results than traditional methods [21].

Our aim in this paper is therefore to propose an alternative approach to the Logit model. In fact, as indicated before, the adoption of the MNL approach as done in other countries is demanding in terms of data reliability and in particular the step of modal choice which consists in predicting which mode the traveler will choose to move while basing himself on his personal data (gender, socio-professional category, having a vehicle or not,…) as well as those of the journey (mileage, reason for the trip, journey time,…).

More in practice, we propose in this paper to apply an advanced machine learning approach to tackle this prediction problem. That is, we aim in the experiment to investigate if the proposed improved machine learning approach give better results in terms of predicting the modal choice for a Moroccan traveler than those reached by classical machine learning approaches.

The rest of the article is organized as follows. In the next section, we describe the related works. In Section 3, we present the concepts related to our improved machine learning approach. Sections 4 and 5 present the experimental setup and the results obtained. Finally, we conclude and present the perspectives of our work.

## 2. LITERATURE REVIEW

In this section, we give a brief review on related works to our subject. Our aim is to present a review of different methods used for travel mode prediction in order to justify our choice.

First of all, we must note that the purpose of transport demand models is to simulate the behavior of travelers during the period under consideration. A travel is the movement of persons from an origin to a destination to perform certain activities, using particular means of transport. Travels are made according to several reasons, namely: work, studies, leisure, services, etc. In this article, we are interested in the four-step model, since it is the one adopted by the Moroccan Department of Equipment and Transport. The aim of this model introduced in 1960 [11] is to predict the number of trips for different travel modes and routes taken between different origin and destination zones.

Various approaches have been adopted to model the choice of the travel mode. For examples, the authors of the paper [12] have adopted an approach of dividing individual travelers into a number of groups based on their individual characteristics. It was performed using cluster analysis through a statistical analysis system software. Trips to a central business district, the city of Nanjing (China) were taken as a case study. The traditional way to deal with such a problem is the multinomial logit model [2], it has been evaluated for the first time in 1981 on data provided by the Chicago Area Transportation Study (CATS). In addition, in [8], a developed heteroscedastic model, in addition to the multinomial logit model, and the nested logit model have been applied to the estimation of the corridor of Toronto-Montreal.

These models have been applied to different case studies, regarding the four-step model, it has been applied in different countries. For instance, [1] describes the process of applying such a model using a simplified transport network in Dhaka City, Bangladesh.

We can see that a number of recent empirical studies have shown that machine learning can outperform logit models in terms of predictive accuracy. See for instance [36]. Thereby, in this paper, we interest especially on how to integrate machine learning to tackle this issue. Machine

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

learning techniques have gained prominence in recent years. It has successively solved various problems in pattern recognition [18], natural language processing [10], recommendation [22], and automatic control of algorithms [14].

The most adopted machine learning methods for travel mode choice prediction are artificial neural network (ANN) and support vector machines (SVM) [17]. In the following, we highlight some of the recent papers that used and compared these two approaches.

Firstly, regarding the underlined paper ([17]), it compared and analyzed the performance of these two methods and analyzed the impact of different variables (characteristics). According to this article, the support vector machine model should be used as an alternative approach for modeling the choice of travel mode because of its promising performance and simple implementation, comparing to multinomial logit. With respect to the comparison with the feed-forward multilayer neural network, SVM showed better results and the neural network was more prone to over-fitting. Also, in [34], a support vector machine model was tested and outperforms a multinomial logit model and a feedforward neural network model, based on a collected data in California.

On the other hand, Gazder [15] proposed a combination of logit model and ANN, for mode choice modeling. This model performed better than logit models and artificial neural networks. The paper in [35] was intended to apply probabilistic ANN for the prediction of travel mode choice. In that paper, the network structure has been established on the basis of data obtained from the resident displacement survey, then the K-means cluster algorithm was applied to optimize the number of hidden nodes. In [30], a deep learning approach was adopted based on a reported preference survey conducted in Singapore. Their approach has generated reasonable economic information at the aggregate level, either by the model set or by the average of the population. In [28], a survey of reported preferences was conducted among a number of participants. Information on three route attributes, including journey time, travel time fluctuations, and fuel cost, was provided to participants to enable them to make decisions on route selection. The results showed that SVM has a much higher computer efficiency than neural networks. The authors of the paper [20] showed that asymmetric data sets, linear and semi-linear models tended to work better, while nonlinear-models such as SVM is adequate for more balanced data sets. Their proposed method was tested on the new Chicago Travel Tracker Survey dataset, and prediction performance was evaluated across different data-mining algorithms. In [31], based on data obtained from Beijing public transit smart cards, an SVM classification scheme was established to identify commuter mode choices.

Other modeling techniques based on machine learning can be used to solve this problem. For example, in [32], a heterogeneous Hidden Markov Modeling (HMM) approach has been proposed to model dynamic discrete choices. The approach has been evaluated and demonstrated on a travel mode choice application using the ten-wave Puget Sound transport panel. HMM is also a powerful machine learning tool that has been applied to other transportation issues, such as airline scheduling [7]. Markov process theory, in general, has been proven for application in airline transport like in [5] and [13]. Notably, the use of Markov model has been applied in [4] to deal with gate disturbances in airport. A most recent machine learning models for transport systems can be found also in [3] and [6]. However, in this article, we are interested in predicting mode choice and therefore we advocate the use of data-based approaches such as SVM and ANN.

Regarding SVM, it has successfully tackled different problems, including demand forecasting and electric load forecasting. Regarding the former, it has been shown in [33] that the use of support vector regression (SVR), the regression form of SVM can be useful to have an accurate prediction for this problem. Concerning the later, it has been presented in [25] that the experimental results on two widely used electric load dataset show that support vector machines (in particular SVR) can yield better results for this problem. This paper shows also that feature selection has improved the performance of support vector machines. More details on feature selection and its importance can be found in the following section. We can also note in this part that support vector machines has been hybridized with other statistical methods such as ARIMA to deal with forecasting problems (see for instance [24]).

At the end of this section, we can conclude that machine learning is suitable to deal with the problem of travel mode prediction, that SVM and ANN are the most common approaches to deal with this problem while using the four-step model.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

## 3. THE PROPOSED APPROACH FOR TRAVEL MODE CHOICE PREDICTION

### 3.1 Support Vector machines

The Support Vector Machine (SVM) is a recent tool from the field of machine learning based on a robust mathematical theory. It has been successfully applied in many areas and has recently attracted increasing interest from researchers. It has been first introduced by Vapnik et al. (1992) [9] and was applied firstly to pattern recognition (classification) problems. Recent research has given extensions to regression problems, including time series forecast. That is, an SVM version for regression (SVR) was proposed by Vapnik et al. in 1997 [29].

This subsection introduces briefly the concept of SVM. In the soft margin version of SVM, we have to optimize this problem:

$$\min_{w,b,\xi_i} \frac{1}{2}||w||^2 + C\sum_{i=1}^{N}\xi_i \quad (1)$$

$$s.t \begin{cases} y_i \cdot (< w, \phi(x_i) > +b) \leq (1-\xi_i) & i = 1,..,n \\ \xi_i \geq 0 \end{cases} \quad (2)$$

Where the $y_i$ are either -1 or 1, each indicating the class to which the point $x_i$ belongs. Each $x_i$ is a p-dimensional real vector.

The constant $C \geq 0$ determines the compromise between the flatness of the model and the error on the training set. $\xi_i$ denotes the training error and represents the number of samples. $\varphi$ correspond to the kernel trick which is briefly illustrated below (more details on these parameters can be found in [26]).

The kernel trick avoids having a direct mapping that is needed to get linear learning algorithms to learn a decision limit. In the input space, some functions can be expressed as an inner product in another higher- dimensional space (this space can be infinite). The most well-known kernel is the radial basis function (RBF). Indeed, the RBF is the core widely used because it offers both good performance and low complexity. We have therefore adopted in this paper the RBF kernel, which is defined as follows:

$$K(x_i, x_j) = exp(-||x_i - x_j||^2/2\sigma^2) \quad (3)$$

As mentioned earlier, in SVM, we adopted kernels to transform a nonlinear classification problem into a linear problem. In particular, we use the RBF kernel as follows:

$$\phi(w) \cdot \phi(x) = K(w, x). \quad (4)$$

where w is the weight vector

We want then a linear classifier in an infinite-dimensional kernel space:

$$g(x) = \text{sign}(\phi(w) \cdot \phi(x) + b), \quad (5)$$

On the other hand, this problem can be reformulated as a Lagrangian, which as above, we need to minimize with respect to w, b and ξi and maximize:

$$L_p = \frac{1}{2}||w||^2 + C\sum_{i=1}^{L}\xi_i - \sum_{i=1}^{L}\alpha_i[y_i(x_i \cdot w + b) - 1 + \xi_i] - \sum_{i=1}^{L}\mu_i\xi_i \quad (6)$$

We can compute then w and b according to the following equations:

$$w = \sum_{i=1}^{L}\alpha_i y_i x_i \quad (7)$$

$$y_i(w^T x_i + b) = 1 \quad (8)$$

### 3.2 Feature selection

Nowadays, choosing the most suitable subset among a large number of datasets is one of the most interesting and difficult problems in solving different prediction problems. Therefore, feature selection (FS), also known as variable or attribute selection, aims to select the most relevant input features within a dataset. It has many advantages because it can improve the prediction ability of the algorithm by eliminating irrelevant inputs, reducing data for accelerated training, and increasing processing efficiency [27]. It is usually utilized to identify a subset where the meanings of features are important.

There are three main approaches for feature selection [16]. Filter ones are based on information theory, wrapper and embedded approaches utilize a machine learning algorithm to score features subsets based on their predictive ability. In embedded approaches, the selection is done in the training

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

process while in wrappers, the machine learning classifier is used as a black box.

Most feature selection algorithms perform a search in space subsets of features. Some characteristics affect the nature of the research: the most important is the organization of the research (the heuristic strategies are generally more feasible and adaptable to this problem) and the evaluator (we can distinguish two main families of methods: filter and wrapper). In this article, we use a wrapper approach based on the predictive accuracy of support vector machines. You can find more details on these approaches in [26].

### 3.3 Artificial neural network

In this paper, we compare support vector machines with artificial neural networks (ANN). Therefore, this section gives a brief description of the idea beyond ANN. Neural networks are a set of algorithms, loosely modeled based on human brain inspiration, as support vector machines, it was designed primarily to recognize patterns. ANN interpret different data according to a sort of perception of the machine, labeling the raw inputs. Neural networks can be used for both prediction and classification.

## 4. EXPERIMENTAL SETUP

### 4.1 Data engineering

In this study, we use a provided dataset that illustrates the choice of mode of travel for 1146 Moroccan passengers in 2016. In these data, the number of samples is 1146, and the number of features is 58. The features concern socio-economic parameters. Such as the type of housing, vehicle ownership, origin and destination of the trip. Our goal is to apply machine learning algorithms to predict whether the traveler will use public transport or their own car. We should also note that we only included observations with complete records in our study. In addition, a challenging issue with such a dataset is to be normalized. This dataset is described within the next paragraphs.

### 4.2 Categorical features

The first issue within this dataset is to manage categorical features. In other words, many machine learning models, such as SVM, are algebraic. This means that their entry must be numeric. To use these models, categories must first be transformed into numbers before we can apply these machine learning algorithms to them. The approach adopted for feature engineering involves some form of transforming the categorical values of features into numeric values, and then applying a coding scheme to those values. For that, we adopted the following Python code:

```python
def handle_non_numerical_data(df):
    columns = df.columns.values
    for column in columns:
        text_digit_vals = {}
        def convert_to_int(val):
            return text_digit_vals[val]
        if df[column].dtype != np.int64 and df[column].dtype != np.float64:
            column_contents = df[column].values.tolist()
            unique_elements = set(column_contents)
            x = 0
            for unique in unique_elements:
                if unique not in text_digit_vals:
                    text_digit_vals[unique] = x
                    x+=1

            df[column] = list(map(convert_to_int, df[column]))

    return df
```

*Figure 1: Converting categorical features in Python*

### 4.3 Data splitting and over-fitting

Another essential issue while pre-processing the dataset is to divide the dataset. Indeed, one of the challenging issues while using machine learning is over-fitting. This occurs when a model learns the details and noises in the learning data as they negatively impact the performance of the model with new data. Therefore, to obtain an accurate assessment of machine learning models, they must be evaluated using invisible data. This is why it is necessary to divide the dataset into training and testing data. The training data and the testing data: the training is used to build the model and the testing equipment to evaluate it on new data.

In our experience, 80% of the samples are used for training, while the remaining 20% are used for testing. To summarize our approach, we depict in

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the following flowchart the main steps for our approach:

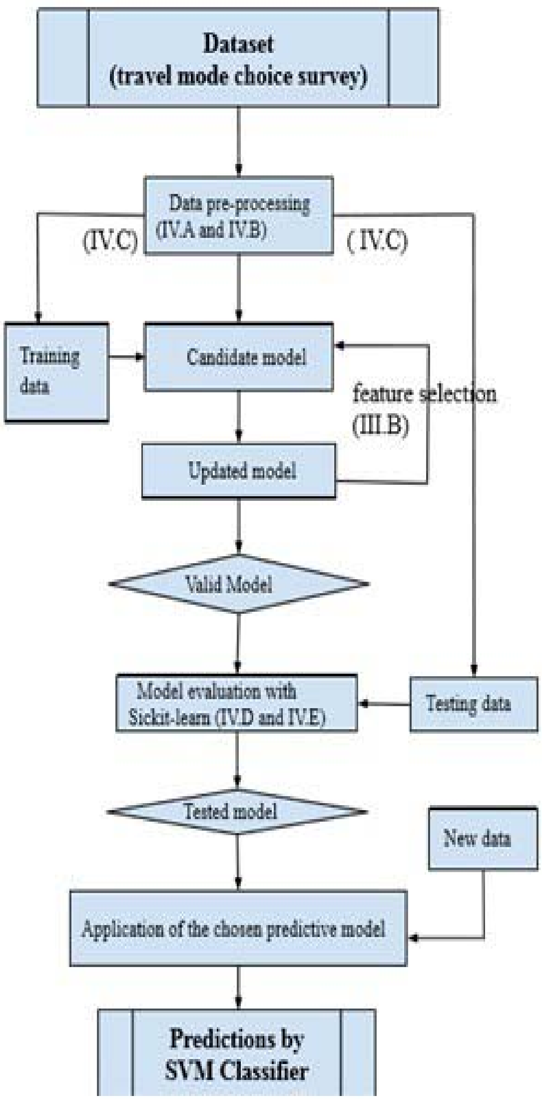



*Figure 2: The Flowchart of the adopted approach*

## 4.4 The adopted package

To compare and conduct the experiment, we have adopted the Sickit-learn package [23]. Scikit-learn is a Python module integrating a wide range of advanced machine learning algorithms for supervised and unsupervised problems. It has been very successful in recent years and is one of the most adopted packages these days.

## 4.5 Evaluation measurement

The typical way to compare machine learning models is to compare their accuracy on the test set. However, we can extend the comparison using the confusion matrix.

This matrix (which is often used for an unbalanced dataset) is a summary of the prediction results on a classification problem. The purpose of the confusion matrix is to show the number of correct and incorrect predictions summarized with count values and broken down by class. It illustrates how your classification model is confused when making predictions. This means that it gives us more errors made by a classifier, the types of errors made. The three elements of the confusion matrix are the precision, the recall and the measure F (F1-Score).

In order to define the elements of this matrix, we describe below the necessary definitions and abbreviations:

*Table 1: Abbreviation Meaning.*

| Abbreviation | Meaning |
|---|---|
| Positive (P) | Observation is positive |
| True Positive (TP) | Observation is positive, and is predicted to be positive |
| False Negative (FN) | Observation is positive, but is predicted negative |
| True Negative (TN) | Observation is negative, and is predicted to be negative |
| False Positive (FP) | Observation is negative, but is predicted positive |

We these definitions, we can define the accuracy, precision, recall and F-measure with the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FP} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + TN} \tag{11}$$

Journal of Theoretical and Applied Information Technology
15th May 2020. Vol.98. No 09


ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$F - measure = \frac{2 * Recall * Precision}{Precision + Recall} \quad (12)$$

## 5. EXPERIMENTAL RESULTS

### 5.1 Comparison with conventional machine learning techniques

In this section, we compare the proposed approach to the main approaches used to solve this problem, namely support vector machines and the artificial neural network, as previously emphasized.

First, we compare the accuracy of the algorithms on the testing set as shown in the following table:

*Table 2: COMPARISON OF THE PERFORMANCE USING THE CONFUSION MATRIX.*

| Data sets | SVM-FS | SVM | ANN |
|---|---|---|---|
| Accuracy | 0.76623376 62337663 | 0.7229437 229437229 | 0.70995 670995671 |

In addition, in the following table, we depict the values of the three elements of the confusion matrix. (precision, recall and F1-score):

The following table shows different performance measurement obtained: ("SVM-FS" corresponds to our approach, "SVM" means that we use the native model and "ANN" is the artificial neural network)

We can see from the previous table that our model (SVM-FS) with FS have better values SVM without FS. We can conclude then that our approach may improve the performance of the SVM model.

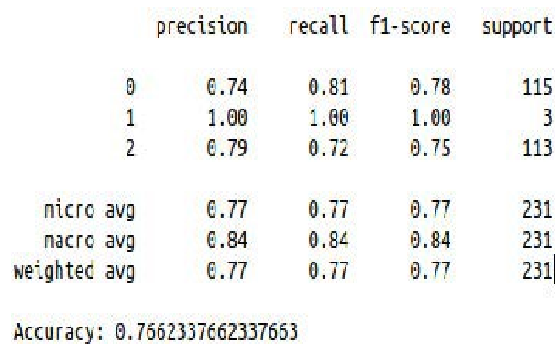
*Table 3:* COMPARISON OF THE PERFORMANCE USING THE CONFUSION MATRIX.

| Data sets | SVM-FS | SVM | ANN |
|---|---|---|---|
| Precision | **0.74** | 0.73 | 0.71 |
| Recall | **0.81** | 0.70 | 0.72 |
| F1-score | **0.78** | 0.72 | 0.71 |

Furthermore, we can see that the feature selection process has improved the prediction capacity of SVM (by comparing SVM-FS with SVM) which means that the eliminated attributes do not have a great impact on the travel mode choice. They can be then replaced by other attributes who can have a more impact on travel mode choice.

### 5.2 Analysis of the approach results

In this section, we conduct a further examination of the performance. For that, we depict in the following figure 3 and 4 the confusion matrix of SVM while using feature selection (more details on the explanation of the various parameters can be found in [23]).



| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.74 | 0.81 | 0.78 | 115 |
| 1 | 1.00 | 1.00 | 1.00 | 3 |
| 2 | 0.79 | 0.72 | 0.75 | 113 |
| micro avg | 0.77 | 0.77 | 0.77 | 231 |
| macro avg | 0.84 | 0.84 | 0.84 | 231 |
| weighted avg | 0.77 | 0.77 | 0.77 | 231 |

Accuracy: 0.7662337662337663

*Figure 3: The confusion matrix of the adopted approach*

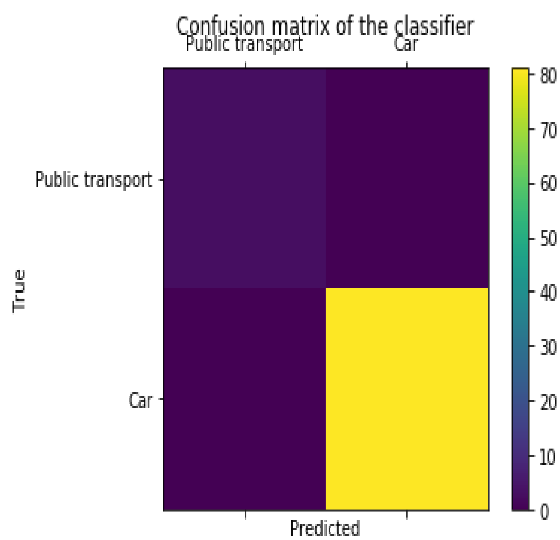



*Figure 4: Plotting the confusion matrix*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

We can conclude at the end of this section that the proposed approach can yield good results both in terms of precision and recall.

## 6. CONCLUSION

In this paper, we investigated the effectiveness of using feature selection to improve the performance of the support vector machine for the travel mode prediction. Experimental results show that our proposed approach (SVM-FS) offers better computational performance than native SVM and the artificial neural network described above. We can conclude that the selection of the most relevant feature and the model selection of SVM can enhance the accuracy for travel mode choice prediction. This result can be useful, especially in the case of large datasets with many features or variables.

These results show that the proposed approach could improve the prediction of the traveler's mode choice, for the case of Morocco. Therefore, it could improve the quality of the assignment on the transport networks and subsequently give better visibility for decision-makers to optimize transport infrastructure projects and guarantee better mobility to citizens. More in general, this work could pave the way for different researches tailored to model the interurban mobility in Morocco through the four-step model.

But, to take advantage from this, future research should attempt to compare our approach with other state of the art algorithms and to validate it. Furthermore, more details have to be provided on how to integrate this approach into the existing system used by the Moroccan Department of Equipment and Transport. In addition, we can investigate the use of other methods instead of the four-step model.

## REFERENCES:

[1] Bayes Ahmed. The traditional four steps transportation modeling using a simplified transport network: A case study of dhaka city, bangladesh. International Journal of Advanced Scientific Engineering and Technological Research, 1(1):19–40, 2012.

[2] Alex Anas. The estimation of multinomial logit models of joint location and travel mode choice from aggregated data. Journal of regional science, 21(2):223–242, 1981.

[3] Oussama Aoun and Abdellatif El Afia. Application of multi-agent markov decision processes to gate assignment problem. In Information Science and Technology (CIST), 2014 Third IEEE International Colloquium in, pages 196–201, Oct 2014. doi: 10.1109/CIST.2014.7016618.

[4] Oussama Aoun and Abdellatif El Afia. A robust crew pairing based on multi-agent markov decision processes. In Complex Systems (WCCS), 2014 Second World Conference on, pages 762–768, Nov 2014. doi: 10.1109/ICoCS.2014.7060940.

[5] Oussama Aoun and Abdellatif El Afia. Using markov decision processes to solve stochastic gate assignment problem. In Logistics and Operations Management (GOL), 2014 International Conference on, pages 42–47, June 2014. doi:10.1109/GOL.2014.6887445.

[6] Oussama Aoun and Abdellatif El Afia. Time-dependence in multi-agent MDP applied to gate assignment problem. International Journal of Advanced Computer Science and Applications, 9(2), 2018. doi: 10.14569/IJACSA.2018.090247.

[7] Oussama Aoun, Malek Sarhani, and Abdellatif El Afia. Investigation of hidden markov model for the tuning of metaheuristics in airline scheduling problems. IFAC-PapersOnLine, 49(3):347–352, 2016. doi: 10.1016/j.ifacol.2016.07.058.

[8] Chandra R Bhat. A heteroscedastic extreme value model of intercity travel mode choice. Transportation Research Part B: Methodological, 29(6):471–483, 1995.

[9] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classiffers. 5th Annual ACM Workshop on COLT, Pittsburgh PA, pages 144–152, 1992.

[10] Ronan Collobert and Jason Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In Proceedings of the 25th international conference on Machine learning, pages 160–167. ACM, 2008.

[11] Juan de Dios Ortuzar and Luis G Willumsen. Modelling transport. John Wiley & Sons, 2011.

[12] Ling Ding and Ning Zhang. A travel mode choice model using individual grouping based on cluster analysis. Procedia engineering, 137:786–795, 2016.

[13] Abdellatif El Afia and Oussama Aoun. Data-driven based aircraft maintenance routing by markov decision process model. In Proceedings of the 2nd international Conference on Big Data,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Cloud and Applications, page 74. ACM, 2017, doi: doi.org/10.1145/3090354.3090430.

[14] Abdellatif El Afia, Malek Sarhani, and Oussama Aoun. Hidden markov model control of inertia weight adaptation for particle swarm optimization. IFAC-PapersOnLine, 50(1):9997–10002, 2017.

[15] U Gazder. Application of neural networks for mode choice modeling and modal traffic forecasting. PhD thesis, PhD dissertation, Department of Civil and Environmental Engineering, King . . . , 2014.

[16] Isabelle Guyon and Andre Elisseeff. An introduction to variable and feature selection. Journal of Machine Learning Research, 3:1157 – 1182, 2003.

[17] Julian Hagenauer and Marco Helbich. A comparative study of machine learning classifiers for modeling travel mode choice. Expert Systems with Applications, 78:273–282, 2017.

[18] Yann LeCun, L´eon Bottou, Yoshua Bengio, Patrick Haffner, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.

[19] M. Minsky and S. Papert. An introduction to computational geometry. MIT Press, ISBN 0-262-63022-2, 1969.

[20] Elke Moons, Geert Wets, and Marc Aerts. Nonlinear models for determining mode choice. In Portuguese Conference on Artificial Intelligence, pages 183–194. Springer, 2007.

[21] Hichem Omrani. Predicting travel mode of individuals by machine learning. Transportation Research Procedia, 10:840–849, 2015.

[22] Michael J Pazzani and Daniel Billsus. Content-based recommendation systems. In The adaptive web, pages 325–341. Springer, 2007.

[23] Fabian Pedregosa, Ga¨el Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. Journal of machine learning research, 12(Oct):2825–2830, 2011.

[24] Malek Sarhani and Abdellatif El Afia. Forecasting demand with support vector regression technique combined with x13-arima-seats method in the presence of calendar effect. International Journal of Applied Logistics (IJAL), 5(2):74–86, 2014.

[25] Malek Sarhani and Abdellatif El Afia. Feature selection and parameter optimization of support vector regression for electric load forecasting. In 2016 International Conference on Electrical and Information Technologies (ICEIT), pages 288–293. IEEE, 2016.

[26] Malek Sarhani, Abdellatif El Afia, and Rdouan Faizi. Facing the feature selection problem with a binary pso-gsa approach. In Recent Developments in Metaheuristics, pages 447–462. Springer, 2018.

[27] S.Piramuthu. Evaluating feature selection methods for learning in data mining applications. European Journal of Operational Research, 156:483–494, 2004.

[28] Bingrong Sun and Byungkyu Brian Park. Route choice modeling with support vector machine. Transportation research procedia, 25:1806–1814, 2017.

[29] V. Vapnik, S. Golowich, and A. Smola. Support vector method for function approximation regression estimation and signal processings. MIT Press Cambridge, 9:144–152, 1992 7.

[30] Shenhao Wang and Jinhua Zhao. Using deep neural network to analyze travel mode choice with interpretable economic information: An empirical example. arXiv preprint arXiv:1812.04528, 2018.

[31] Jiancheng Weng, Qiang Tu, Rongliang Yuan, Pengfei Lin, and Zhihong Chen. Modeling mode choice behaviors for public transport commuters in beijing. Journal of Urban Planning and Development, 144(3):05018013, 2018.

[32] Chenfeng Xiong, Xiqun Chen, Xiang He, Wei Guo, and Lei Zhang. The analysis of dynamic travel mode choice: a heterogeneous hidden markov approach. Transportation, 42(6):985–1002, 2015.

[33] Yi Yang, Rong Fuli, Chang Huiyou, and Xiao Zhijiao. Svr mathematical model and methods for sale prediction. Journal of Systems Engineering and Electronics, 18(4):769–773, 2007.

[34] Yunlong Zhang and Yuanchang Xie. Travel mode choice modeling with support vector machines. Transportation Research Record, 2076(1):141–150, 2008.

[35] Dan Zhao, Chunfu Shao, Juan Li, Chunjiao Dong, and Yinhong Liu. Travel mode choice modeling based on improved probabilistic neural network. In Traffic and transportation studies 2010, pages 685–695. 2010.

[36] Xilei Zhao, Xiang Yan, Alan Yu, and Pascal Van Hentenryck. Modeling stated preference for mobility-on-demand transit: A comparison of machine learning and logit models. arXiv preprint arXiv:1811.01315, 2018.