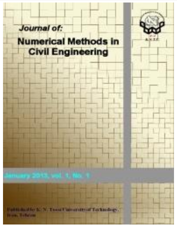


Numerical Methods in Civil Engineering

Predicting the Traffic Crashes of Taxi Drivers by Applying the Non-Linear Learning of ANFIS-PSO with M5 Model Tree

Ehsan Abbasi* and Mansour Hadji Hosseinlou**

ARTICLE INFO

Article history:
Received:
September 2018.
Revised:
December 2018.
Accepted:
February 2019.

Keywords:
General health questionnaire, Traffic accident rates, M5 tree, Adaptive neuro-fuzzy inference systems, Nonlinear-learning

Abstract:
As an essential issue in traffic crashes, human factor plays an indispensable role. This study uses the general health questionnaire (GHQ-12) within some socio-demographic and also a number of daily exercise related questions for prediction of traffic crashes among taxi drivers in the City of Tehran. A novel technique is been developed by applying nonlinear-learning of composition model of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Particle Swarm Optimization (PSO) with M5 model tree. In order to improve the generalization ability of a single data driving algorithm, a cluster of ANFIS models with different nodes and hidden layers are implemented to extract the inherent relationship between traffic accident rates and human factors. The predictions of ANFIS models are combined applying a nonlinear weighted average method composed of M5 tree, and the PSO is utilized to find the optimal parameters of ANFIS models. The performance of the proposed model is evaluated in a case study and the relevant data are collected from a large group of Taxi drivers in the City of Tehran, Iran; as it was carried out to predict the associated accident rates. The Nash-Sutcliffe coefficient (NSE) and different error criteria are utilized to assess the prediction efficiency of the associated Hybrid model. Results indicate that the M5 tree is capable in enhancing the prediction accuracy of the novel model applied in the prediction of the accident rates as in comparison with the popular ANFIS model. Moreover, the linear relationships generated by M5 tree show the sensitivity of ensembled model accuracy on the single ANFIS models, which have a partial tendency in underestimation of the traffic crashes.

1. Introduction

Traffic accident rates in Iran is twenty times the global average, which makes the country suffering from the extensive consequences of traffic injuries, deaths and the associated costs. According to the Iranian experts~~,~~ this anomaly is somehow related to the countrywide psychological and health related problems [1]. Therefore, more investigations regarding the possible relationships of these problems with traffic crashes might be helpful in order to address an appropriate policy for reducing the traffic accident rates in the country. In other hand, there is a growing tendency to investigate the relationships between crash predictions and traffic operating characteristics such as road environment , traffic and weather conditions [2]; while there are great interests to determine the vast causes of crashes based on the human factors as the most important determined elements in the analysis of the traffic crashes [3, 4]. While the traffic crashes are the main reasons for injuries and consequential disabilities, and sometimes the mortality in Iran [5], thus it is crucial to comprehend as many as factors which could influence in these type of crashes.

Several studies utilized crash severity data for modeling the crashes severity by applying different types of the family of regressions, logit models, Artificial Neural Networks (ANN), Fuzzy Models, and Time-Series Models [6–9]. The study of relationship between mental health and driving behavior of taxi drivers in the City of Kerman, Iran, compared to Manchester's Driving Behavior Questionnaire (MDBQ) and General Health Questionnaire (GHQ) revealed that there is a meaningful positive relationship between mental health and driving behavior. If drivers possessed a proper mental health, their driving behavior would also be acceptable. In addition to that, the physical activity of the drives is proved to be a reducing factor of hypertension among the taxi drivers in Brazil [10].

Ph.D. student, Department of Civil Engineering, K. N. Toosi University of Technology, Tehran, Iran.
**Corresponding Author: Associate Professor, Department of Civil Engineering, K. N. Toosi University of Technology, Tehran, Iran. Email: mansour@kntu.ac.ir*

There are many studies which have focused on determining the epidemiology of urban traffic crashes of Tehran [11] and Crash generation concept in the city of Mashhad [12]. More recently, a study has been done which investigated the social determinants of risky driving on the intercity roads of Tehran province showed that people with a driving job, chronic disease, poor socio-economic status, and having only a family dispute, without a religious attitude, under medical supervision, secondary education, or being a woman, and applying drugs had more road traffic crashes. This study also concluded that among all significant aforementioned factors, those factors were related to socioeconomic status, health condition, and family structure had a greater role [13].

Based on literature review, it seems that the preceding studies have considered many vehicles, roads, environmental and human factors for prediction of traffic crashes. Moreover, very few of the previous works have attempted to consider human factors, which could influence on the accident rate in a realistic area of study. In order To accurately realize the relative impact of these human factors; it is necessary to develop a prediction model which can present the linear relations between the inputs and outputs. To the best of our knowledge, insufficient attempts was done to address the influence of exercise and mental health on the accident frequencies. Accordingly, the main objective of the present study was to examine the relationships of these human factors on the property damage only crashes among taxi drivers of Tehran.

2. Methodology

2. 1. Adaptive Neuron Fuzzy Inference System (ANFIS)

The concept of a fuzzy set was introduced in 1964 by Zadeh [14] who was working on the problem of computer's compiling capabilities which is not easily transformed into the absolute terms of 0 and 1. Therefore, Fuzzy logic is intended to model the logical reasoning with vague or imprecise statements. The most common fuzzy methodology is Mamdani's fuzzy inference method which was proposed in 1975 by Ebrahim Mamdani [15] to control a combination of steam engine and boiler by synthesizing a set of linguistic control rules obtained from experienced human operators. Fig 1. (a) shows the general structure of a typical fuzzy logic system.

Regarding the safety and reliability analysis, a membership function (M) is defined by the typical convex functions as triangular, trapezoidal, rectangular and Gaussian type that defines how each point in the input space is mapped to a membership value between 0 and 1. The shape of the membership function commonly does not affect the final results. We use Gaussian type in this study due to the fact that based on the two variables, number of minutes of exercise and mental health, the Gaussian type might be more appropriate.

The ANFIS is an adaptive neuro fuzzy inference system that is based on Takagi-Sugeno fuzzy inference system [16]. It has the capability to approximate nonlinear functions [17]. Fig. 1. (b) shows the structure of Sugeno fuzzy inference system. The rules in Sugeno can be written as: "If x_1 is A_{i1}, and x_2 is A_{i2}, then y_i is $f_i(x)$, were x_1 and x_2 are the input variables, A_{in} are the linguistic variables and y_i is the consequent part of ith rule and w_i are the weights used for calculating $\hat{y}_i$, which is the estimate of y_i ". For more detailed reading about Takagi-Sugeno structure, please refer to Ismail et al [18].

2. 1. 1. Particle Swarm Optimization (PSO)

The PSO is a popular meta-heuristics optimization algorithm presented by Kennedy and Eberhart in 1995 [19]. It uses three key components including velocity, fitness function and positon to find the optimal solution. Fitness of each solution is the utility of objective function such as minimum or maximum values. The position shows the unknown solution of the model and the velocity specifies the speed of positions' variations. Particles are p dimensional vectors where fitness function will be computed for these particles. A particle is randomly generated within the search area and a random velocity is assigned to it. The number of initial particles in the search area is named as "generation". A fitness function is a criteria of excellence of particles. Usually, the PSO tries to find the minimum value of the fitness function. it uses the local rules to find the global solutions of the optimization algorithm. Generations' clusters in the point that it has a great fitness function. However, the optimization process can be terminated when the maximum value of generation reached or the termination criteria desired. In the PSO algorithm, the position of initial particles is changed as below [20]:

$$x_{k+1}^i = x_k^i + u_{k+1}^i \tag{1}$$

where, velocity u_{k+1}^i is calculated by

$$u_{k+1}^i = wu_k^i + c_1 r_1 \left(p_k^i - x_k^i\right) + c_2 r_2 \left(p_k^g - x_k^i\right) \tag{2}$$

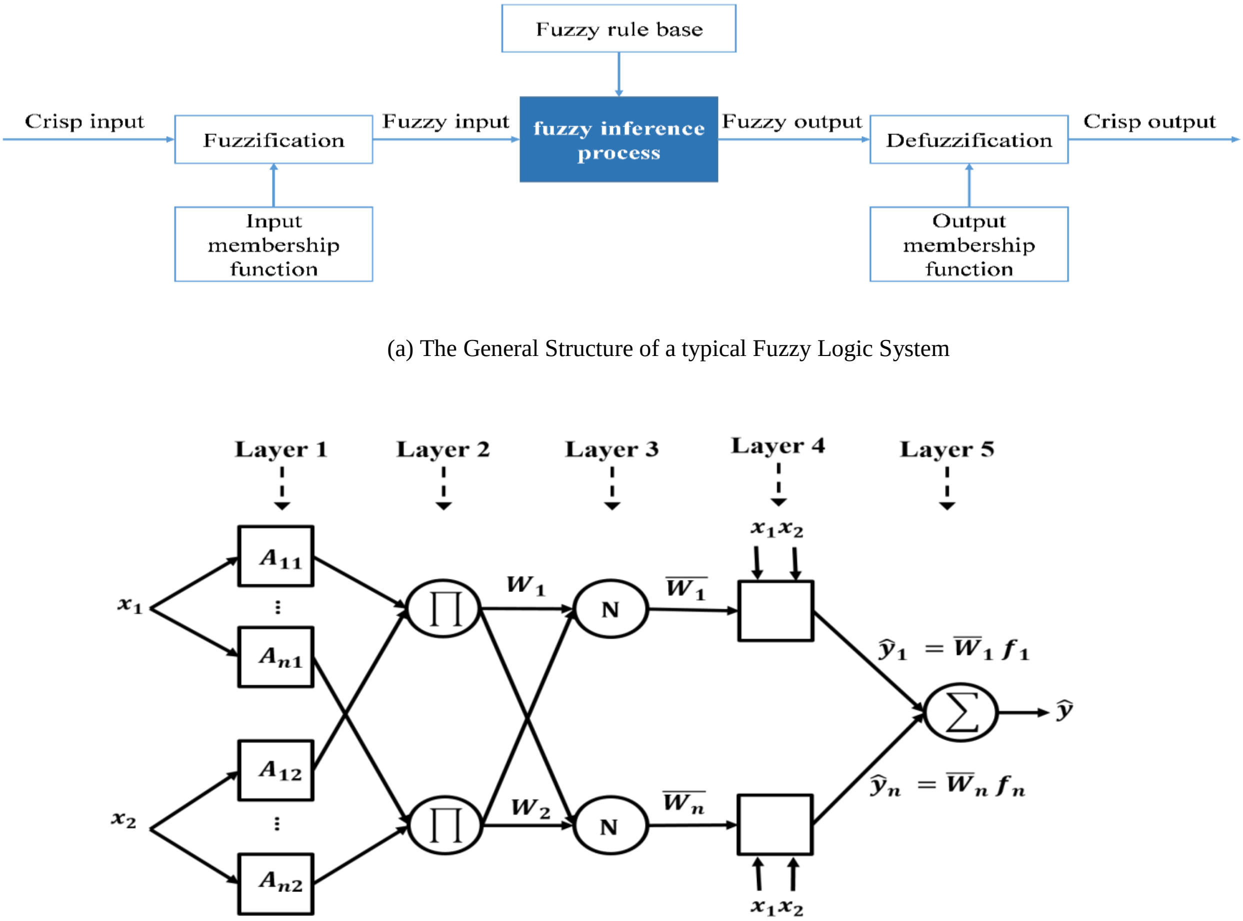
in which, x_{k+1}^i is the updated positon of each particle, x_k^i represents particle position, u_k^i is the velocity of particle, p_k^i is the best position of particle, p_k^g is the best position of swarm, u_{k+1}^i is the new value of particle, c_1 and c_2 are the constant coefficients.

2. 2. Regression Method

The regression method investigates the relationship between a dependent variable and independent variables. In the other words, the dependent variable is modeled as a function of the independent variables as follows [21]:

$$\mathrm{Y} = f(\mathrm{X}.\beta) + \varepsilon \tag{3}$$

Where Y is a dependent variable, X is an independent variables, β is for the unknown parameters and ε is an error term. If the regression function is unknown, the function must initially be guessed and a trial and error process must be applied to find the best function. The regression function can be linear, exponential, power, logarithmic, polynomial and so on [21].



(a) The General Structure of a typical Fuzzy Logic System

(b) Sugeno-Type Fuzzy Inference

Fig. 1: (a) The General Structure of a typical Fuzzy Logic System, (b) Sugeno-Type Fuzzy Inference [18]

2. 2. 1. Poisson Regression Model (PRM)

PRM or a log-linear model is a generalized linear model form of the regression analysis used to predict a dependent variable that consists of the count data. The count data is a type of data in which the observations can only take the non-negative integer. The Poisson regression assumes that the response variable Y has a Poisson distribution, and assumes the logarithm of its expected value can be modeled by a linear combination of unknown parameters (21).

$$\ln[E(Y)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \qquad (4)$$

The mean accident frequency is $\lambda = E(Y)$ which can be interpreted by the Poisson distribution function. The mean and variance of the Poisson distribution is equal (21):

$$E(Y) = \mathrm{var}(Y) = \lambda \qquad (5)$$

Since the mean value is equal to the variance, any factor that affects one will also affect the other. Therefore, if the observed variance is greater than the mean (known as over dispersion) negative binomial (NB) model is used. Also, the Poisson and the negative binomial models are not applicable if many zero crashes are observed. Thus in this case, the Zero-altered probability processes such as zero-inflated Poisson (ZIP) model and zero-inflated negative binomial (ZINB) model must be used [21].

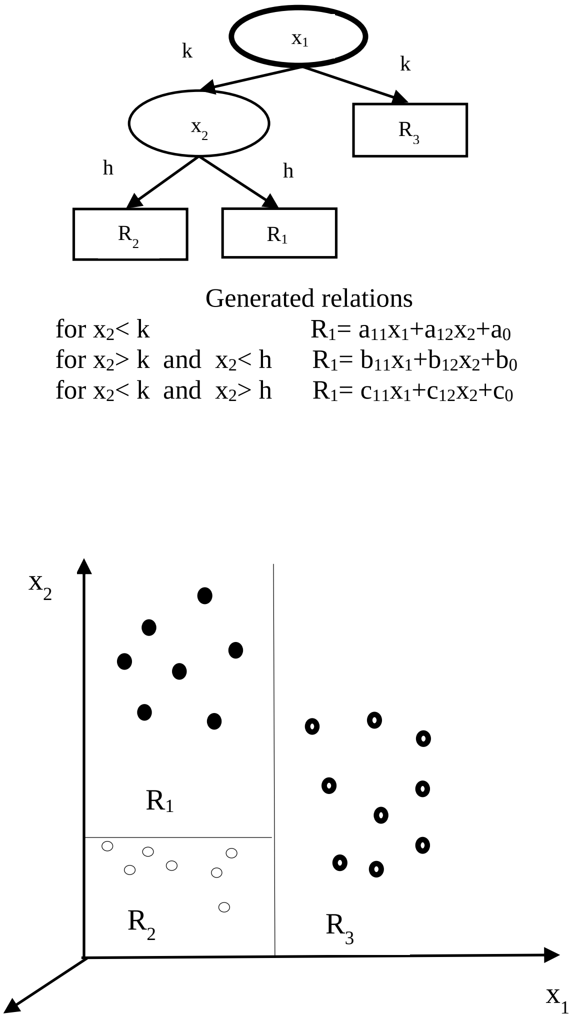
2. 3. M5 Tree

The decision tree is a new machine learning algorithm presented by Quinlan (1992) [22] that it has been used for prediction in various fields, recently. In the decision tree, a node is located on top of the tree and many leaves are located below. A sample is entered in the node and it is examined to determine if this sample belongs to branch. Many methods are developed for this examination, however, the objective of all methods is unique. The selection rules should be determined according to structure of the learning model [23]. Selection process is continued to reach the best roles. In the M5 tree model, a complex problem can be divided in simple sub networks and linear relationship is allocated to them. Therefore, the M5 tree can be used for complex nonlinear problems. The M5 tree consists of three steps including building of tree, pruning and smoothing [24]. In the first step of the building tree, the features with more ability for splitting should be determined. The split rules are determined based on the standard deviation reduction (SDR) index [25].

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \times sd(T_i) \tag{6}$$

Where, T = number of the samples in split nodes, Ti = number of the new nodes created after splitting top nodes, based on the SDR index, sd = standard deviation, |Ti|/|T = criterion for the prediction error after splitting top node.

If the value of the SDR index is less than 5%, the splitting process is finished (26). Fig. 2 shows the methodology developed in this study for prediction of the maximum discharge rates. As demonstrated in Fig. 2, the M5 tree is trained with training samples that are generated by the simulation-optimization model. In each iteration, the error between the simulated and predicted values of the maximum discharge rate is computed. If the error is more than the threshold value of 0.072, the simulation–optimization model is operated for a new aquifer dimensions and the M5 tree retrained with the new samples.



Generated relations

for $x_2 < k$	$R_1 = a_{11}x_1 + a_{12}x_2 + a_0$
for $x_2 > k$ and $x_2 < h$	$R_1 = b_{11}x_1 + b_{12}x_2 + b_0$
for $x_2 < k$ and $x_2 > h$	$R_1 = c_{11}x_1 + c_{12}x_2 + c_0$

Fig. 2: The structure of M5MT tree

2. 4. Nonlinear-learning of ANFIS with M5 model tree

Bootstrap aggregation or bagging is a novel approach for integration of many predictors to create a global model whose efficiency is better than individual predictors [27]. In this technique, the weighted average of the training function is resulted in canceling the variance of each model and enhance the performance of predictions. Considering many predictors applying simple linear regression, where the M dataset is to predict y applying input data (x). The outputs of the linear regression model in the bagging technique versus the single regression can be expressed as below:

$$y_{RSM}(x) = \frac{1}{M} \sum_{m=1}^{M} y_m\,(x) \tag{7}$$

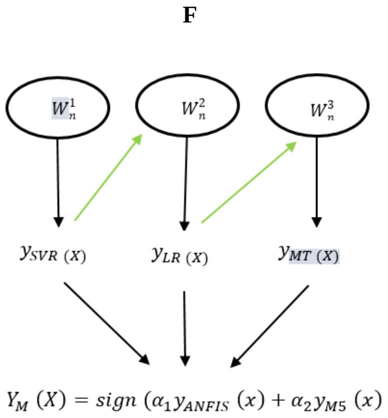
$$y_m(x) = h\,(x) + \,e_m\,(x) \tag{8}$$

Where, m = 1, . . . , M is number of dataset and h(x) the result of prediction by simple regression.

A comparison of error between predicted and observed data for bagging (E_{bag}) and simple regression (E_r) can be written as below [27]:

$$E_{bag} = \frac{1}{M}\,E_r \tag{9}$$

It can be indicated that the generated bagging error will not exceed the generated error of the single regression models. In this study, the bagging technique is extended to improve the performance of a set of decision tree. Additionally, the boosting technique [28] is used for combining the individual ANFIS and M5 tree models (see Fig. 3). In the boosting technique, each single predictor y_m is trained applying weighted result of many dataset, where this weights w(m) are function of efficiency of each predictors. The training process is applied for each one of the predictor and finally, all in the model are integrated.



F
W_n^1
W_n^2
W_n^3
$y_{SVR}\,(X)$
$y_{LR}\,(X)$
$y_{MT}\,(X)$
$Y_M\,(X) = sign\,(\alpha_1 y_{ANFIS}\,(x) + \alpha_2 y_{M5}\,(x)$

Fig. 3: The structure of proposed model

3. Evaluation of prediction performance

For the evaluation the efficiency of the hybrid model statistical indexes such as correlation coefficient and relative error statically Nash–Sutcliffe criterions are used.

$$R^2 = \frac{E_0 - E}{E_0} \times 100 \tag{10}$$

$$E_0 = \sum_i (X_{observed} - \bar{X}_{predicted})^2 \tag{11}$$

$$E = \sum_i (X_{observed} - \hat{X}_{predicted})^2 \tag{12}$$

$$\bar{X}_{predicted} = \left(\sum_{i=1}^{n} X_{observed}\right)/n \tag{13}$$

where $X_{observed}$ = value of observed accident rate, $\bar{X}_{predicted}$ = average of predicted accident rate, $\hat{X}_{predicted}$ = value of predicted accident rate and n is the number of samples.

R^2 denotes the correlation coefficient and its value is between 0 and 1. Furthermore, the root mean squared error (RMSE) and the relative error (RE$_m$) are used for the evaluation of the error between the simulated and predicted samples. The RE$_m$ criterion is an indicator of error between the maximum value of the simulated and predicted samples.

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(X_{observed} - X_{predicted})\right]^{\frac{1}{2}} \tag{14}$$

$$RE_m = \frac{(X_{observed} - \hat{X}_{predicted})}{X_{observed}} \times 100 \tag{15}$$

4. Case study

This descriptive-analytical cross-sectional study was done during October and November 2017. All of the taxi drivers on the urban streets of Tehran were considered as a target population. A total number of 294 taxi drivers in 6 major taxi stations (located in the north, south, east, west and two at the center of Tehran) were chosen applying the proportional allocation sampling and a random systematic sampling was conducted on each one of the taxi stations. Among them, 259 taxi drivers had fully completed the questionnaires. The study objectives were verbally explained to the participants via face-to-face interviews. They were assured that the gathered information would remain confidential and/or anonymous. The participants completed the questionnaires and returned them immediately after completion on site. Data collection tools included socio-demographic data questionnaires (age and educational level), history of physical disease (brain and cardiovascular diseases, liver and kidney diseases, gastrointestinal disease, and musculoskeletal disorder), disabilities, if they are smoking or not, and number of minutes of daily exercise. Also, the number of traffic crashes that they had during past few months were included in the questionnaire with the general health questionnaire (GHQ-12).

The 12-item General Health Questionnaire (GHQ-12) which is the shortest version of this questionnaire (the original GHQ has 60 items), is a widely used screening instrument which was developed by Goldberg in the 1970s to assess the current mental health of individuals (29). Also, this version is used in many countries and languages [30]. Each item is rated on a four-point scale (less than usual, no more than usual, rather more than usual, or much more than usual), applying one of two most common scoring methods: dichotomous (0-0-1-1) or Likert type (0-1-2-3) (14). We used the Likert type scoring method which has ranges from 0 to 36 that the minimum value (0) illustrates that the individual has no mental health problems and the maximum value indicated that a serious symptoms of mental health problem exists.

5. Results and discussion

The input parameters of hybrid model are including university degree (UD), smoking rate (SR), age (A), disease rate (DR) and exercise per week (EW). To implement decision tree applying dimensionless parameters, the efficient machine learning software WEKA [31] is employed. According to the lack of the training data, the hold out method is used to split the test data from the training data. The decision tree is trained with 70% of the obtained samples from the study area. The number of the tree parameters are determined by the trial and error procedure. The tree with low Nash–Sutcliffe error is applied for the prediction. The decision tree calculates the RMSE for each iteration and updates itself with the new samples generated by the ANFIS model. When RMSE is less than threshold value, the training process is stopped and parameters determined for current samples. The depth of the M5 tree and number of nodes for the prediction step are 2 and 4, respectively. In this study, the ANFIS algorithm used the fuzzy concept with membership functions [0, 1] and the Gaussian MF function. Fig. 4. Illustrates the range of the membership function (MF) parameters ~~that~~ needed for optimization before the training process.

Table. 3. Values of statistical index for proposed hybrid model

Factor	R^2	RE_m	$RMSE$
PRM	0.65	-3.15	0.025
$M5$	0.91	-1.8	0.012
ANFIS	0.88	-2.14	0.008
Hybrid model	0.96	-1.52	0.006

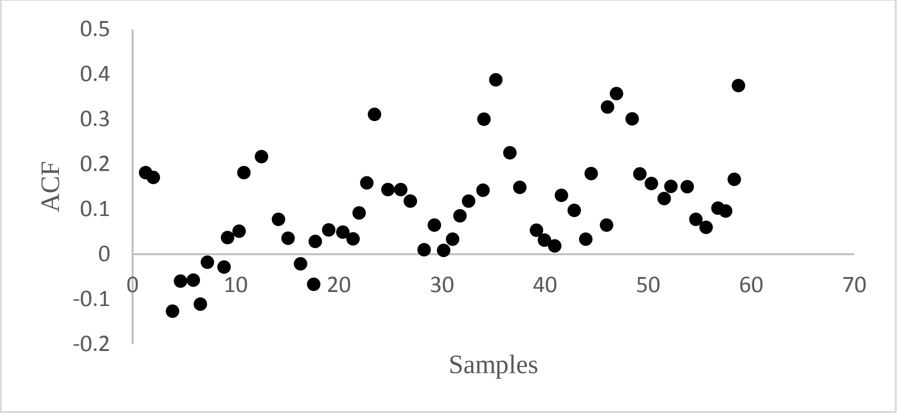


Fig. 4: Auto-correlation function (ACF) of the samples for nonlinear learning of ANFIS models

The linear equations generated by M5 tree for accident rate is presented below:

LM1: If A >38 and EW > 2 Then, Accident rate = 0.0914 × (A) + 0.0012 (DR) – 0.5
LM2: If A >44 and EW > 3 Then, Accident rate = 0.082 × (A) + 0.022 (EW) – 0.2
LM3: If A >55 and EW > 4 Then, Accident rate = 0.0914 × (A) + 0.0012 (DR) + 0.0012 (SR) – 0.5

As proposed by the LM1, when A >38 and EW > 2 (LM1), the predicted value of the accident rate is more effected by the age of drivers, and also DR and EW have low impact. Additionally, with increasing the DR value, the accident rate increases. In other words, for drivers which have the exercise rate more than 2 days per week the accident rate is influenced by the age value, while the LM2 proposes that for A >44, the accident rate has more impacted EW and age has negligible effect. Also, for A >55, the accident rate is influenced by SR and DR. This means that the effect of smoking for A >55 is critical. Interestingly, all relations indicate that the university degree has no effect on accident rate. A quantitative judgment applying above mentioned statistical indexes is applied on the performance of four data driven models (PRM, ANFIS and M5) and the proposed bagging technique. The tendency to a linear equation between predicted and observed result is investigated in term of coefficient of determination (R). Additionally, the error between the predicted and observed data is evaluated applying the $\boldsymbol{RE_m}$ and RMSE indexes.

The efficiency criteria for 200 training and 140 validation samples of the abovementioned algorithms is summarized in Table. 3. Result indicate that the value of R for three models ranged between 0.65 and 0.96. Also, the value of RMSE changed from 0.0061 to 0.025 for Hybrid and PRM methods respectively. Interestingly, the maximum and minimum correlation value among four approach is belonged to Hybrid model (0.96) and PRM (0.65) respectively. Moreover, the negative value of Bias RE_m for all methods show that this models underestimate the accident rate. Between four techniques to forecast accident rate, the Hybrid technique shows the lowest value of RMSE (27.01) and AIC (26.09). The poor performance of single ANFIS model can be attributed to large number of parameters used for training kernel function. The scatter plots of predicted accident rate by all approaches and the observed data are shown in Fig. 5. These plots comprise the goodness of fit between predicted an actual value of accident rate for test samples.

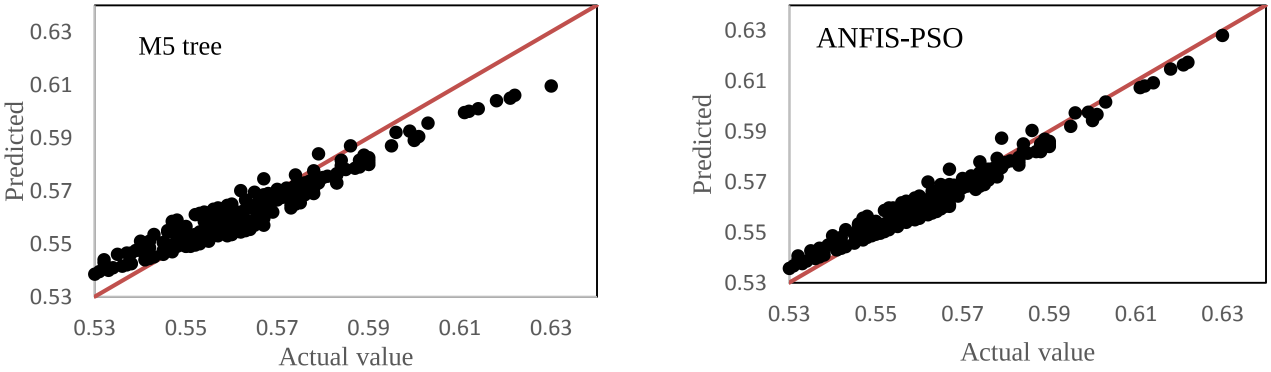


Fig. 5: Comparison of the results of ANFIS-PSO and M5 tree with observed test dataset

Fig. 6. (a) and (b) represents the results of prediction performance of Hybrid model for training (70%) and testing (30%) data respectively. In each part the mean and standard deviation for residuals, the RMSE and the differences between the targets and output of Hybrid model plotted for comparison.

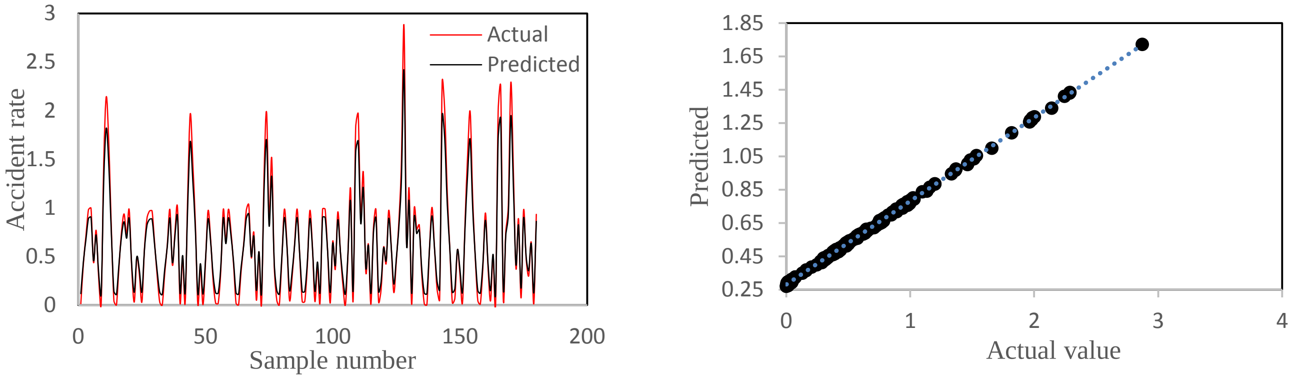


Fig. 6: Comparison between predicted and actual value of accident rate for different samples

As shown in Fig. 6 for small values of accident rate, hybrid model can be trained with more training samples while, large values of accident rate show high correlation between simulated and predicted samples. Figs. 6 illustrate that the accident rate forecasted by the Hybrid model shows more similarities with the M5 results and generate less residual errors. To improve the generalization ability of Hybrid model, the different ensemble of the M5 tree was examined. Finally, better forecasting efficiency of Hybrid model than the ANFIS, PRM and M5 tree is verified in Fig. 5, which indicate the great learning ability of the nonlinear-learning method. Despite the non-significance of socio-economic variables, this study revealed an interesting result that numbers of minutes of daily exercise and general health play significant role in prediction of accident rate. The results of the Hybrid model and the M5 tree confirms that with increasing minutes of daily exercise and decreasing the mental health problems, the number of accident rate decreases among all taxi drivers. This findings in line with previous studies which indicated driving performance is affected by health related changes. Also, physical activity was shown to be a protection factor for hypertension among taxi drivers in Brazil, even considering the deleterious effect of time as a taxi driver.

6. Conclusion

In this study, a novel technique is developed for the prediction of traffic crashes among taxi drivers of Tehran. Additionally, to improve the generalization ability of a single data driven algorithm, a cluster of the ANFIS models with different nodes and hidden layers are implemented to extract the inherent relationship of traffic accident rates. The frequency of traffic crashes determined to be a good measure for studying traffic crashes and many attempts by different types of models conducted to understand the factors influencing traffic crashes. Result indicate that between four techniques (PRM, ANFIS, M5 and Hybrid) to forecast the accident rate, the Hybrid technique show the lowest value of RMSE (27.01) and AIC (26.09). The maximum and minimum correlation value among four approach is belonged to Hybrid (0.96) and PRM (0.65), respectively. The findings also showed that the Hybrid model could be effectively implemented in the accident frequency studies and the policy makers may simply make an intervention to encourage taxi drivers who spent a lot of their time on traffic, to easily can have exercise and to encourage them by various policy making to do more daily exercises. Therefore, daily exercise would not only bring wellbeing for them but also reduces the crashes risks. Also, improving mental health by providing training programs for taxi drivers, may result in less traffic crashes. Consequently, Taxi drivers are vulnerable to many risks and needs more attention. As proposed by the M5 tree relations, when age >38 and exercise per week > 2, the predicted value of accident rate is more effected by age of drivers, and the disease rate and exercise per week have low impact. It will be recommended that the future studies to consider the cultural approaches in additions of mental and physical health. Furthermore, previous study announces that the number of crashes in Tehran is different for different seasons of the year. Therefore, it is appropriate to implement the future studies for each one of the seasons and conclude based on the differences between seasonal trends. One has to be aware that the data of this study is based on the self-reported crashes and the results may be data specific. Moreover, the physical activity and general health may be influenced by age, educational level, income and other demographic variables that needs further investigation. Finally, this study focuses on the PDO, and the other types of traffic crashes need more investigations and the results may be different for those type of crashes.

Acknowledgements

We express our warmest gratitude toward those taxi drivers of Tehran who generously accepted to participate in this study and complete the instrument toll by their patience and careful readings.

References

[1] Enayat H. Driving Culture in Iran: Law and Society on the Roads of the Islamic Republic. Iranian Studies 2017; 50(2):323–6.

[2] You J, Wang J, Guo J. Real-time crash prediction on freeways applying data mining and emerging techniques. J. Mod. Transport. 2017; 25(2):116–23.

[3] Pakgohar A, Tabrizi RS, Khalili M, Esmaeili A. The role of human factor in incidence and severity of road crashes based on the CART and LR regression: A data mining approach. Procedia Computer Science 2011; 3:764–9.

[4] Lajunen T, Summala H. Can we trust self-reports of driving?: Effects of impression management on driver behaviour questionnaire responses. Transportation Research Part F: Traffic Psychology and Behaviour 2003; 6(2):97–107.

[5] Noori Hekmat S, Dehnavieh R, Norouzi S, Bameh E, Poursheikhali A. Is There Any Relationship between Mental Health and Driving Behavior of Taxi Drivers in Kerman? GJHS 2016; 9(2):294.

[6] Kockelman KM, Kweon Y. Driver injury severity: An application of ordered probit models. Accident Analysis & Prevention 2002; 34(3):313–21.

[7] Abdel-Aty MA, Abdelwahab HT. Predicting Injury Severity Levels in Traffic Crashes: A Modeling Comparison. Journal of Transportation Engineering 2004; 130(2):204–10.

[8] Delen D, Sharda R, Bessonov M. Identifying significant predictors of injury severity in traffic crashes applying a series of artificial neural networks. Accid Anal Prev 2006; 38(3):434–44.

[9] Teymuri GH, Sadeghian M, Kangavari M, Asghari M, Madrese E, Abbasinia M et al. Review the number of crashes in Tehran over a two-year period and prediction of the number of events based on a time-series model. Electron Physician 2013; 5(3):698–705.

[10] Marcelo CV, Sandro S, Arianne CR. Physical activity overcomes the effects of cumulative work time on hypertension prevalence among Brazilian taxi drivers. THE JOURNAL OF SPORTS MEDICINE AND PHYSICAL FITNESS; 2016.

[11] Rabiei R, Ayatollahi H, Rahmani Katigari M, Hasannezhad M, Amjadnia H. Epidemiology of Urban Traffic Crashes: A Study on the Victims' Health Records in Iran. GJHS 2016; 9(5):156.

[12] Naderan A, Shahi J. Aggregate crash prediction models: introducing crash generation concept. Accid Anal Prev 2010; 42(1):339–46.

[13] Moradi A, Salamati P, Vahabzadeh E. The Social Determinants of Risky Driving on the Intercity Roads of Tehran Province, Iran: A Case-Cohort Study. Arch Trauma Res 2016; 6(1).

[14] Zadeh LA. Fuzzy sets. Information and Control 1965; 8(3):338–53.

[15] Mamdani EH, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies 1975; 7(1):1–13.

[16] Jang J. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans. Syst., Man, Cybern. 1993; 23(3):665–85.

[17] Abraham A. Adaptation of Fuzzy Inference System Applying Neural Learning. In: Kacprzyk J, Nedjah N, Macedo Mourelle Ld, editors. Fuzzy Systems Engineering. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005. p. 53–83 (Studies in Fuzziness and Soft Computing).

[18] 2016 Federated Conference on Computer Science and Information Systems: IEEE; 2016. (Annals of Computer Science and Information Systems).

[19] Kennedy, J., Eberhart, R.C., 1995. Particle Swarm Optimization. Proc. IEEE Int'l. Conf. on Neural Networks IV: 1942e1948. IEEE Service Center, Piscataway.

[20] NJ. Kennedy, J., Eberhart, R.C., 2001. Swarm Intelligence. Academic press, CA, USA.

[21] Zheng X, Liu M. An overview of accident forecasting methodologies. Journal of Loss Prevention in the Process Industries 2009; 22(4):484–91.

[22] Quinlan, R. J. Learning with Continuous Classes (1992) In: 5th Australian Joint Conference on Artificial Intelligence.

[23] Schrider, D. R., & Kern, A. D. (2016). S/HIC: Robust identification of soft and hard sweeps applying machine learning. PLoS Genet, 12(3), e1005928.

[24] Oliver, J. J., & Hand, D. J. (2016, January). On pruning and averaging decision trees. In Machine Learning: Proceedings of the Twelfth International Conference (pp. 430-437).

[25] Jung, N. C., Popescu, I., Kelderman, P., Solomatine, D. P., & Price, R. K. (2010). Application of model trees and other machine learning techniques for algal growth prediction in Yongdam reservoir, Republic of Korea. Journal of Hydroinformatics, 12(3), 262-274.

[26] Norouzi, M., Collins, M., Johnson, M. A., Fleet, D. J., & Kohli, P. (2015). Efficient non-greedy optimization of decision trees. In Advances in Neural Information Processing Systems (pp. 1729-1737).

[27] Hothorn, T., & Lausen, B. (2003). Double-bagging: Combining classifiers by bootstrap aggregation. Pattern Recognition, 36(6), 1303-1309.

[28] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and regression trees. CRC press.

[29] Gao F, Luo N, Thumboo J, Fones C, Li S, Cheung Y. Does the 12-item General Health Questionnaire contain multiple factors and do we need them? Health Qual Life Outcomes 2004; 2:63.

[30] Gouveia VV, Barbosa GA, Andrade EdO, Carneiro MB. Factorial validity and reliability of the General Health Questionnaire (GHQ-12) in the Brazilian physician population. Cad. Saúde Pública 2010; 26(7):1439–45.

[31] Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.

[32] Yasa, R., & Etemad-Shahidi, A. (2014). Classification and regression trees approach for predicting current-induced scour depth under pipelines. Journal of Offshore Mechanics and Arctic Engineering, 136(1), 011702.