Research Article

Journal of Algorithms & Computational Technology

Journal of Algorithms & Computational Technology
2018, Vol. 12(1) 20–29
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1748301817729953
journals.sagepub.com/home/act
SAGE

Traffic fatalities prediction using support vector machine with hybrid particle swarm optimization

Xiaoning Gu[1], Ting Li[2], Yonghui Wang[2], Liu Zhang[3], Yitian Wang[2] and Jinbao Yao[4]

Abstract

Road traffic safety is essential, therefore in order to predict traffic fatalities effectively and promote the harmonious development of transportation, a traffic fatalities prediction model based on support vector machine is established in this paper. The selection of parameters greatly affects the prediction accuracy of support vector machine. Introducing particle swarm optimization can find the optimal parameters and improve the prediction accuracy of support vector machine by parameter optimization. However, standard particle swarm optimization is easy to trap into the local optimum, so that the best parameter solutions cannot be found. Therefore, the mutation operation of the genetic algorithm is introduced into particle swarm optimization, particle swarm with mutation optimization is generated. It expands the search space and makes parameter selection more accurate. This paper predicts fatalities of traffic accident using small samples and nonlinear data. The results show that compared with particle swarm with mutation optimization back propagation neural network prediction model, particle swarm optimization-support vector machine model, support vector machine, back propagation neural network, K Nearest Neighbor (K-NN), and Bayesian network, the prediction model of traffic fatalities based on particle swarm with mutation optimization-support vector machine has higher prediction precision and smaller errors. It is feasible and effective to use particle swarm with mutation optimization to optimize the parameters of support vector machine, and this model can predict the accident more accurately.

Keywords

Traffic accident, support vector machine, particle swarm optimization, mutation operation, prediction model, optimal parameters

Date received: 8 June 2016; accepted: 29 July 2017

Introduction

Background

In recent years, vehicle number and highway mileage are increasing along with the continuous improvement of road infrastructure construction of China. This has contributed to the economic development but also had some negative effects: frequent road traffic accidents. Among all kinds of traffic accidents, the harm of traffic fatality to social life is extremely serious; it is always threatening our personal safety and has become a serious social problem, which is worthy of our attention. However, traffic fatality has strong randomness. The randomness is affected by the factors such as driver and passenger characteristics, vehicle types, traffic conditions, and geometric design characteristics. However, the complex relationship between traffic fatalities and various influencing factors is nonlinear.

[1]Automotive Engineering College, Dalian University of Technology, Dalian, PR China
[2]Transportation Management College, Dalian Maritime University, Dalian, PR China
[3]School of Transportation Science and Engineering, Beihang University, Beijing, PR China
[4]School of Civil Engineering, Beijing Jiaotong University, Beijing, PR China

Corresponding author:
Jinbao Yao, School of Civil Engineering Beijing Jiaotong University, Beijing, PR China.
Email: bao_yaojin@163.com

As the various factors influence each other, it is difficult to use a single factor to explain traffic fatalities. Therefore, it is necessary to summarize and analyze the traffic safety data and find out the inherent laws of traffic fatality. It is of practical significance to forecast the development trend of traffic fatality under existing road traffic conditions, and it provides the basis for further formulating the road traffic safety plan or making the decision.

Literature review

At present, many methods used in traffic fatality prediction have different application conditions and modeling mechanisms. Binomial regression, Bayesian approach, back propagation neural network (BPNN) models, and some new methods are used to fit the accident data. Poch and Mannering[1] estimated a negative binomial regression of the accidents frequency at intersection approaches. Clarke et al.[2] created decision trees with the use of a machine learning method. It distinguishes between personal injury caused by the accident and damage in the general sense. Abdel-Aty and Haleem[3] explored to combine multivariate adaptive regression splines with another machine learning technique (random forest). Xu et al.[4] aimed to build the genetic programming model for real-time crash prediction on freeways and evaluated the application of the model. Ramani and Selvaraj[5] optimized the aggregated feature selection with voting algorithm. An optimal number of significant features with majority votes were selected. Other traffic accident prediction methods can be found in these literatures.[6–11]

Some studies have also proposed new models of traffic accident prediction. Yasdi[12] and Quek et al.[13] used artificial neural network (ANN) for traffic prediction and applied the model on the road. Xie et al.[14] evaluated the application of Bayesian neural network model in vehicle crash accident prediction. Kunt et al.[15] used 12 accident-related parameters in the genetic algorithm (GA), pattern search, and BPNN modeling methods. These models are used to predict the severity of the highway traffic accident. Deublein et al.[16] used an improved Bayesian network to assess the traffic risk accidents in Switzerland. The number of accidents involving personal injuries on Swiss roads was verified and the forecast tolerance was 25%. This shows the prediction model is effective and efficient, and it provides a theoretical basis for the road network planning and decision-making process. Kunt et al.[17] predicted the severity of freeway traffic accidents in Iran, Tehran in a GA, pattern search, and ANN modeling methods. The prediction model is established by the parameters, which includes the age and sex of the driver, the type of vehicle, the road speed ratio, and collision type. ANN, GA and the combination model of GA and PS are also used in traffic accident predictions. The prediction results of the three models are compared. Although the ANN can distinguish complex nonlinear system, other problems such as the slow convergence speed, overlearning, and local extreme value still exist. The existence of these problems has an impact on their prediction accuracy.

Support vector machine (SVM) has begun to be used in traffic accident prediction in recent years. SVM can be self-learning and optimized based on variable data.[18] Small sample, nonlinear, and local extreme problems can be solved by it. Li et al.[19] used the SVM model to predict motor vehicle collisions. The study results show that the SVM model is more accurate than the traditional negative binomial model in predicting collision data. Li et al.[20] developed a SVM model for predicting the severity of the injury caused by different accidents. They also compared the performance of the model and the ordered probability model. Yang and Zhao[21] introduce the accident rate per 10,000 cars and the accident rate per 10,000 people in the paper. The SVM model needs to be modified and improved. For example, the performance of the SVM depends on the parameters. Before the training phase, there are three parameters (C, v, γ) that need to be determined. An improved SVM model is proposed based on the theory of SVM. Many of the literature suggest that heuristic algorithms have been widely used in solving many complex problems,[22–24] their algorithms are tested to be effective by the results. In addition, in most cases especially in real-life optimization problems, the best results can be obtained through these algorithms. To select the parameter values of SVM automatically, Particle swarm optimization (PSO) is applied in this paper. There are many researches on optimizing parameters using PSO. Li et al.[25] also used PSO to search for the optimal parameters of SVM to predict traffic fatalities. The results show that PSO-SVM is more scientific and has high accuracy; it effectively improved the prediction accuracy than single SVM. Cai et al.[26] proposed a PSO-SVM to detect the spectrum in the study of cognitive radio systems and obtain a nonlinear threshold. Simulation results show that the performance of the algorithm is better than that of traditional energy detection. Xiao et al.[27] studied the causes of early failure of large-scale doubly fed wind turbines (DFWT). PSO is used to optimize the acquisition of the corresponding characteristic signals of the DFWT transmission system and improve the accuracy of the input of SVM. The results show that this optimization can effectively improve the accuracy, so that DFWT misalignment type recognition is more accurate. However, there are some defects in using PSO to optimize the parameters of SVM. The convergence speed of PSO is fast, it is easy to fall into local optimum, and search for the "false" optimal solution. Therefore, the idea of GA is introduced in PSO to expand the solution space. Ding et al.[28] proposed a hybrid particle swarm GA to solve the classification problem. The classification results of the data sets are compared with those of other algorithms, and the experimental results show the effectiveness of the algorithm. There is a good balance between the speed of convergence and the diversity of the population, and a better classification accuracy is obtained. Yang et al.[29] combined the crossover variation of GA with PSO to optimize the water injection system for multisource water, and the experiment proves that the optimization efficiency is increased. The mutation operation of GA is also used to improve the performance of PSO in this paper.

Contributions

There are two main contributions in this paper: First, the traffic accident prediction model based on particle swarm with mutation optimization (PSOM)-SVM is proposed. PSO with mutation operation is introduced to find the optimal parameter combination of SVM. Since the introduction of mutation operation, the PSOM effectively enlarges the search range of optimal solution and avoids the local optimal situation. Traffic fatalities involving death are the most harmful traffic accident and have been highly valued. The statistics of traffic fatalities are more comprehensive, so this paper takes the traffic fatalities indicator as the most comparable indicator in traffic accident. Highway mileage, vehicle number, and population size are put into the model to get the number of traffic fatalities, which is the most comparable indicator in traffic accident; second, the performance of the PSOM-SVM, PSOM-BPNN, PSO-SVM, SVM, BPNN, K-NN, Bayesian network, and neural network prediction model is compared. The ability to fit and predict the model is evaluated by calculating the magnitude of error values.

The rest of the paper is organized as follows: the next section introduces the principle of SVM. The model and process of traffic accident prediction based on PSOM-SVM are described, respectively, in "The prediction model of traffic fatalities based on PSOM-SVM" and "The process of traffic fatalities prediction based on PSOM-SVM" sections; test results and error value comparison of different model are presented in "Numerical test" section; the conclusions and direction for future research are presented in the final section.

The principle of SVM

SVM is a supervised learning model based on VC dimension theory and structural risk minimization principle of statistical learning theory. This method is a learning method in small sample situation, and it is proposed by Vapnik[30] and has been fairly mature. SVM has a better generalization ability to solve machine learning problems in classification and induction. SVM has the advantage that it does not get trapped in a local optima.[31,32] Moreover, SVM has the global optimal characteristics, these characteristics make SVM do not need to perform complex nonlinear optimization and not fall into local optima. When solving the nonlinear operation, the corresponding kernel function[33] is defined to greatly simplify the calculation, SVM maps the data in the nonlinear low-dimensional space into linear high-dimensional space, and transfers the search for the optimal linear regression hyperplane algorithm into solving convex programming problem under convex constraint, so as to get the global optimal solution.[34,35]

When the input training sample is nonlinear, the sample is fitted by the following method to obtain a nonlinear function. And then through this function the nonlinear data are mapped into the high-dimensional feature space, thus a linear regression of these data in the high-dimensional space is got, and then it can be transformed into the nonlinear regression of the original space. The following equation represents the fitting function[36] approximately

$$y = \omega\phi(x) + \varepsilon \tag{1}$$

where:

ω is for the weight vector;
x is input vector;
ε represents the offset value.

To minimize the following two values through training

$$P(f) = c\sum_i L(y - f(x)) + \frac{1}{2}\omega^2 \tag{2}$$

$$L(y - f(x)) = \frac{1}{n}\begin{cases} \left|y - f(x)\right| - \varepsilon & \left|y - f(x)\right| \geq \varepsilon \\ 0 & \left|y - f(x)\right| < \varepsilon \end{cases} \tag{3}$$

where:

$c\sum_i L(y - f(x))$ is for the experienced error term;
$\frac{1}{2}\omega^2$ is a regular item;

$L(y - f(x))$ represents the loss function, balancing the weighting function of training error term and the complex term;
c is the penalty factor;
ε stands for loss function parameter, whose value affects the number of support vector.

Here introduces the slack variables ξ_i and ξ_i^*, and then the optimization problem can be converted into

$$\min \frac{1}{2}|\omega|^2 + c\sum_i (\xi_i + \xi_i^*) \tag{4}$$

$$s.t.\begin{cases} y_i - \omega\phi(x) - \varepsilon \le \varepsilon + \xi_i \\ \omega\phi(x) + \varepsilon - y_i \le \varepsilon + \xi_i^* \end{cases} \tag{5}$$

where, the Lagrange multiplier a_i and a_i^* are introduced, and the problem is transferred further into a simple optimization problem of the dual problem

$$\begin{aligned} \max \sum_i y_i(a_i - a_i^*) - \theta \sum_i (a_i + a_i^*) \\ -\frac{1}{2}\sum_i \sum_j (a_i - a_i^*)(a_j - a_j^*)k(x_i, x_j) \\ -\sum_i (a_i - a_i^*) = 0 \end{aligned} \tag{6}$$

$$s.t. \quad 0 \le a_i \le C, 0 \le a_i^* \le C \tag{7}$$

The final prediction function finished is as follows

$$y = \sum_i (a_i - a_i^*)k(x_i, x_j) + \varepsilon \tag{8}$$

$k(x_i, x_j)$ represents the kernel function, which is the inner product of the two vectors $\varphi(x_i)$ and $\varphi(x_i)$ in the feature spaces. The kernel function is set to avoid operating the complex operation caused by $\varphi(x_i)$ and $\varphi(x_j)$. It is the key to the nonlinear SVM problem, which can map low-dimensional data to higher dimensions, and the data can be linearly separable. The detailed derivation process can be derived from the studies of Cao LJ and Francis EH,Cao LJ and Francis EH.[37,38]

The prediction model of traffic fatalities based on PSOM-SVM

To achieve a comprehensive assessment of traffic accidents, the choice of relevant indicators should follow the following three principles: representation, testability, and comparability. Traffic system consisted of three basic factors: people, vehicle, and road. The occurrence of traffic accidents has a strong randomness due to a variety of quantitative factors and qualitative factors. In the literature on traffic accident prediction, highway mileage, vehicle number, lane width, average daily flow, and population are selected as impact factors.[39] The combination of many factors including person, vehicle and road, highway mileage, vehicle number, population size, led to the occurrence of traffic accidents. Traffic fatality is the most serious consequences of traffic accidents, traffic accidents involving the death are highly valued, traffic statistics have few omissions. Therefore, as the output variable, the traffic fatality can be compared with the known real data for accuracy.

The current widely used indicators of traffic accidents are the number of traffic fatalities, the number of injuries, the number of road traffic accidents and economic losses. Because there is no uniform statement about the definition of injury, the road traffic accident statistics about road accidents has not been completed yet. The most comparable traffic accident deaths are selected as predictor index. Therefore, we get the traffic accident prediction model shown in Figure 1.

The process of traffic fatalities prediction based on PSOM-SVM

SVM is a theory of machine learning law in small sample situation; it has unique advantages in the small sample and nonlinear problems, especially in terms of prediction. However, in the learning process of SVM, the selection of parameters has a strong subjectivity, which seriously restrains the accuracy and effect of SVM prediction. The value of penalty factor c and kernel parameter σ affects the prediction accuracy of SVM, and finding the optimal c and σ is the priority. At present, these parameters are usually defined artificially based on the specific issues, and the optimal parameter combination is determined by choosing the parameters for many times and comparing with each other. Parameters that are manually set are blind and of low efficiency, so it is needed to adopt swarm intelligence optimization[40] algorithm to improve the parameter

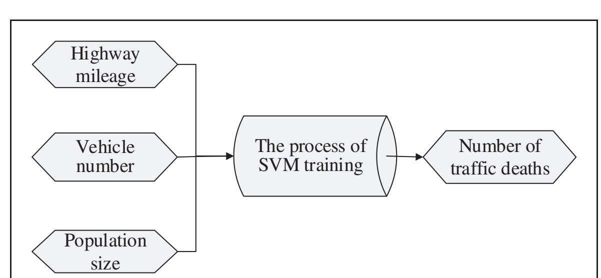

Figure 1. The structure diagram of traffic fatalities prediction model.

choosing of the SVM. At the same time, the design and implementation of PSO algorithm is relatively simple. Not only the convergence speed is fast, but the parameters required to be set are less.[37]

PSO algorithm is a kind of population intelligence algorithm proposed on the basis of studying the behavior of birds and fish by Kennedy and Eberhart.[36] The idea comes from the theory of artificial life and evolutionary computation; it imitates the foraging behavior and achieves the optimal group through the bird collective collaboration.

Compared with the evolutionary computation, the PSO algorithm is a global search strategy, which uses the simple operation of the v–s model. PSO has unique memory mechanism, thus it can adjust the search strategy by keeping track of the current search based on real time, which makes PSO an efficient parallel search algorithm. Due to the parameter setting requirements of PSO algorithm, although PSO has fast convergence speed, it exists some limitations of stagnation. Therefore, to further expand the solution space and improve the prediction accuracy of SVM, the mutation operation of GA is introduced into PSO for predicting SVM parameters.

In the process of using PSOM to solve the problem and optimize the parameters, each particle represents a solution to the problem. Through the preset fitness function, each particle has its corresponding fitness value. Particle velocity determines the direction and distance of particle movement. According to the particle itself and the surrounding particles of inertia, the particle velocity can be dynamically updated timely. In every optimization search process, the particle is updated by two values. One value is the optimal solution obtained by the particle itself, known as the individual extremum, and the other is global optimal solution, called the global extremum. The mutation operation is performed on the particle position according to the mutation probability, so that the particles evolve into new particles, thus updating the particle position. After performing a search optimization in PSO, mutation operation adds the nature of global optimization.

Particle swarm with mutation optimization (PSOM) optimizes the parameters of SVM as the following steps

Start and set parameter;
Repeat
Initialize parameters of PSOM, such as population size and the iteration number, set num = 1;
Randomly generate particle position and the particle velocity, set X_i = the ith position, V_i = the ith velocity;
Mean square error (MSE) is chosen as the fitness function;
 Repeat
Calculate the fitness value of each particle, find the optimal X_i and V_i;
$V_{i+1} = \omega V_i + c_1\lambda_1(P_i - X_i) + c_2\lambda_2(P_g - X_i)$;
$X_{i+1} = X_i + V_{i+1}$;
Find the optimum particle fitness;
Roulette selection is used cooperating with elite strategy;
Mutation operators are used to create a child population;
Set num = num + 1;
If fitness agrees **then**
 Output the best individual and optimal solution;
Else
 Run the operators of mutation;
End if
 Until the stopping criterion is met;
 SVM training and prediction;
Until the prediction accuracy is achieved.

After the algorithm iteration is complete, the optimal result of the memory in the population is the optimal parameter.

Numerical test

Data

Dates of highway mileage, vehicle quantity, population quantity, and traffic fatality are collected from the website of National Bureau of Statistics of China, and the related data are shown in Appendix 1.

Collect the sample data from 1981 to 2012 as the experimental data. Samples of 1981–2006 are training data, while 2007–2012 are test data. In the process of training samples, parameters of PSO are set as follows: The population scale is of 20, the iteration number is N=200. The initial values of accelerating factor c_1 and c_2 are 1.5 and 1.7, respectively. The mutation probability is P_m=1−n/200.

Data normalization

Due to the large difference in data units and magnitude of different variables, different variables must be normalized. If the original data are used directly for the model calculation, it is likely to generate potential data, resulting in a large error. After normalizing, the data will fit well which improves the precision of prediction.

The normalization is done using the following equation

$$A_i^l = \frac{A_i}{\|A\|_2} = \frac{A_i}{\sqrt{A_1^2 + A_2^2 + \cdots + A_i^2}} \tag{9}$$

where:

A_i is the ith original value of variables that needs to be normalized. In this paper, it refers to the ith original value of highway mileage, vehicle number, and population size;

A_i^l is the ith value of highway mileage, vehicle number, and population size after normalization.

The training of the model

The result analysis is evaluated with two evaluation criteria, including mean absolute percentage error (MSE) and the coefficient of determination. The expression of these criteria is as follows

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{10}$$

$$(R^2) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{11}$$

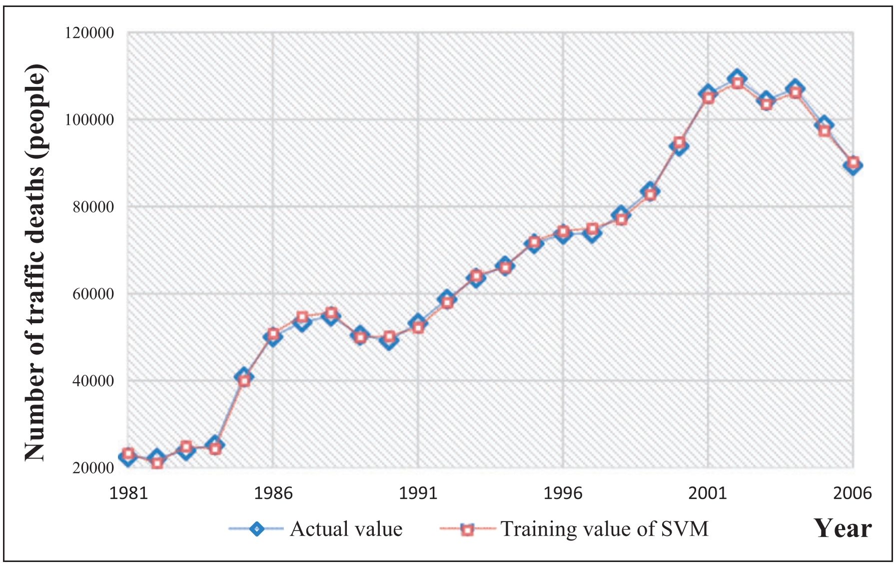

Figure 2. Training diagram of SVM prediction model.
SVM: support vector machine.

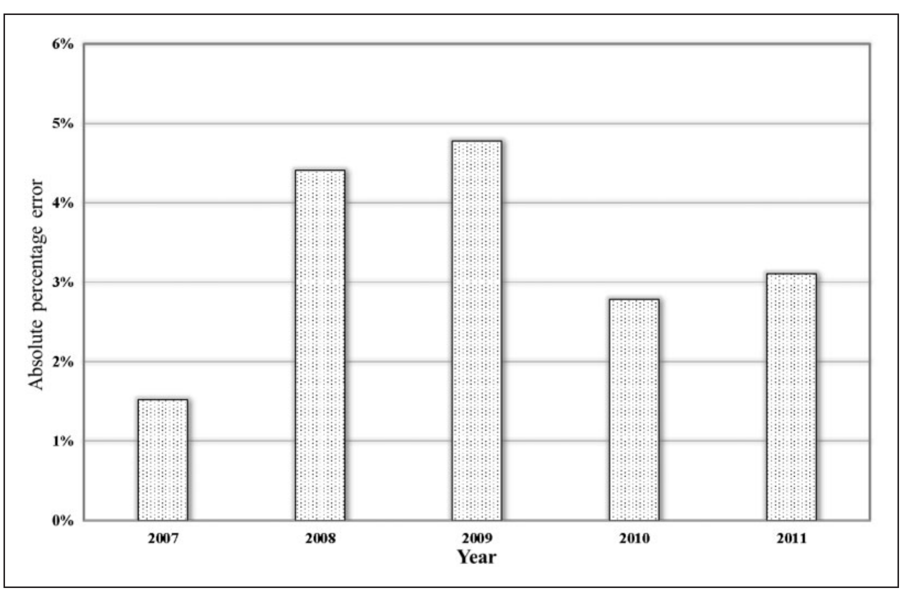

Figure 3. Absolute percentage error of the traffic fatalities prediction.

where n is the size of fitting or predicting sample, $\hat{y}_i$ stands for the estimated traffic fatalities at year i, y_i is the observed number of traffic fatalities, and $\bar{y}$ means the average value of traffic fatalities.

The model performance is better if the value of MAPE is smaller and R^2 is larger.[41–44] The traffic fatalities training curve based on SVM prediction model is shown in Figure 2. The black curve represents the actual output, while yellow one represents the predicted fitting output. From the picture we can see that the two curves fit well. The MSE is 3.79% and the measurement coefficient (R^2) is 0.973. The results show that the forecasting model of traffic accident based on PSOM-SVM has strong identification ability, and the fitting is stable, the error is small.

As seen in Figure 2, the blue curve represents the actual value, it links real traffic fatalities from 1981 to 2006 with a smooth curve, expressing the reality of traffic fatalities from 1981 to 2006. While the red curve represents the training curve, it predicts traffic fatalities between 1981 and 2006 through parameters by PSOM-SVM and represents the prediction results in the form of curve. As shown in Figure 2, the two curves are fitting well, so it is shown that the optimized SVM parameters in the training are accurate and can be used in tests.

The prediction of the model

Predict traffic accidents in 2007–2012 using the trained model. The absolute percentage error of traffic accident prediction is shown in Figure 3. The MSE is 3.63% and the measurement coefficient (R^2) is 0.973.

As seen in Table 1, this paper compares the predictions of PSOM-SVM, PSOM-BPNN, PSOSVM, SVM, BPNN, K-NN, Bayesian network, MSE, respectively, the MSE of these predictions are 3.63%, 4.29%, 6.429%, 7.388%, 7.997%, and 8.233%, obviously. R^2 are 0.973, 0.947, 0.8782, 0.8234, 0.792, and 0.7875. The predicted results of several methods are shown in Figure 4. The prediction model of SVM based on PSOM is better than PSOM-BPNN, PSO-SVM, SVM, BPNN, K-NN, Bayesian network. SVM model is better than neural network model. This is because that SVM algorithm has global optimality and can avoid the local optimal point in the prediction, which avoids the shortcomings of the neural network method, thus prediction accuracy improved. PSOM-SVM model can avoid the manual selection of parameters; it can intelligently search optimization and expand the search population to avoid falling into the local optimal. All of these advantages increase input variable accuracy and prediction.

Table 1. Predict results comparison of PSOM-SVM, PSOM-BPNN, PSO-SVM, SVM, BPNN, K-NN, and Bayesian network.

		PSOM-SVM		PSOM-BPNN		PSO-SVM		SVM		BPNN		K-NN		Bayesian network	
Year	Actual value (people)	Predicted value	APE (%)	Predicted value	APE (%)	Predicted value	APE (%)	Predicted value	APE (%)	Predicted value	APE (%)	Predicted value	APE (%)	Predicted value	APE (%)
2007	81,649	77,575	5.17	85,756	5.03	76,244	6.62	73,169	10.39	86,829	6.34	86,825	7.53	75,680	7.31
2008	73,484	74,586	1.52	76,607	4.25	74,616	1.54	75,205	2.35	76,755	4.45	77,570	5.56	77,974	6.11
2009	67,759	70,821	4.41	71,899	6.11	71,127	4.97	71,239	5.14	74,239	9.56	75,152	10.91	74,765	10.34
2010	65,225	68,427	4.78	61,038	6.42	69,465	6.50	71,981	10.36	58,117	10.90	58,918	9.67	71,924	10.27
2011	62,387	64,058	2.78	65,587	5.13	64,252	2.99	65,792	5.46	66,862	7.17	67,453	8.12	68,039	9.06
2012	59,997	58,183	3.11	62,199	3.67	61,875	3.13	62,935	4.90	63,535	5.90	63,711	6.19	63,783	6.31

APE: absolute Percentage Error; BPNN: back propagation neural network; K-NN K Nearest Neighbor; PSO: particle swarm optimization; PSOM: particle swarm with mutation optimization; SVM: support vector machine.

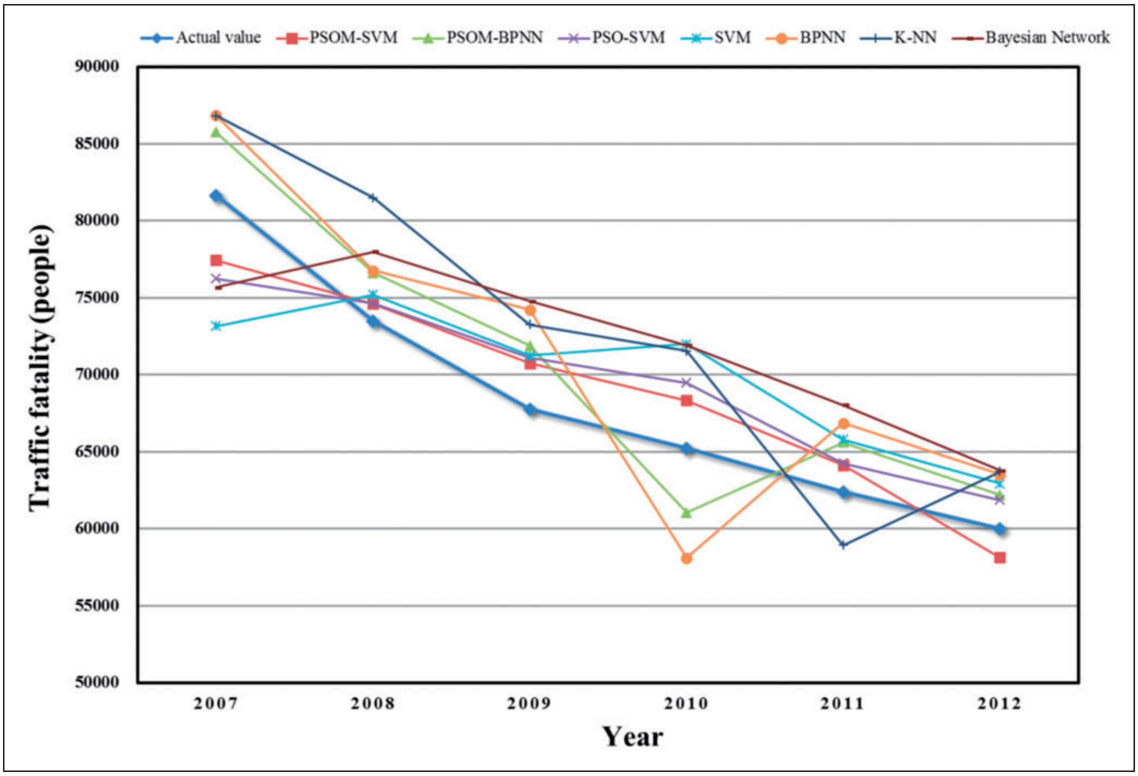

Figure 4. Predict results of different methods.

Conclusions

SVM model has the advantages of strong learning and good generalization ability when solving small sample problem. PSO is easy to fall into the local optimal, the introduced mutation operation can improve the defects above of PSO. PSOM model has less parameters, simple program and fast culate convergence. In this paper, the traffic fatalities prediction model based on PSOM-SVM (PSO with mutation operation) is established, and the parameters of SVM are optimized by this model. The results of example analysis show that the prediction method based on PSOM-SVM model is superior to the prediction method of neural network and BPNN method in terms of the same data, and it overcomes the problem of ''overlearning'' phenomenon in neural network training progress, avoids the local optimal solution, and has extremely good generalization ability. Therefore, the prediction model based on PSOM-SVM is better than the general forecasting model of traffic accident, and the prediction accuracy is better.

The traffic fatalities prediction using PSOM-SVM can reduce casualties to a certain extent. However, it is limited that only highway mileage, vehicle number, population size, and traffic fatalities are selected as parameters in the model, because the prediction model will produce some unknown factors that cannot be ignored. It will be better that the parameters are supplemented in future studies, and the impact factors are taken into account as much as possible in the prediction.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in National Natural Science Foundation of China 71571026 and 51578112, Liaoning Excellent Talents in University LR2015008, and the central universities DUT16YQ104.

References

1. Poch M and Mannering F. Negative binomial analysis of intersection-accident frequencies. *J Transp Eng* 1996; 122: 105–113.
2. Clarke DD, Forsyth R and Wright R. Machine learning in road accident research: decision trees describing road accidents during cross-flow turns. *Ergonomics* 1998; 41: 1060–1079.

3. Abdel-Aty M and Haleem K. Analyzing angle crashes at unsignalized intersections using machine learning techniques. *Accid Anal Prev* 2011; 43: 461–470.
4. Xu C, Wang W and Liu P. A genetic programming model for real-time crash prediction on freeways. *IEEE Trans Intell Transp Syst* 2013; 14: 574–586.
5. Ramani RG and Selvaraj SA. Pragmatic approach for refined feature selection for the prediction of road accident severity. *Stud Inform Control* 2014; 23: 41–52.
6. Tesema TB, Abraham A and Grosan C. Rule mining and classification of road traffic accidents using adaptive regression trees. *Int J Simul* 2005; 6: 80–94.
7. Lee Y and Wei CH. A computerized feature selection method using genetic algorithms to forecast freeway accident duration times. *Computer-Aided Civil Infrastruct Eng* 2010; 25: 132–148.
8. Zhang C, Yan X, Ma L, et al. Crash prediction and risk evaluation based on traffic analysis zones. *Math Probl Eng* 2014; 2014: 1–9.
9. Nassiri H, Najaf P and Amiri AM. Prediction of roadway accident frequencies: count regressions versus machine learning models. *Scientia Iranica* 2014; 21: 263–275.
10. Zong F, Xu HG and Zhang HY. Prediction for traffic accident severity: comparing the Bayesian network and regression models. *Math Probl Eng* 2013; 2013: 206–226. DOI: 10.1155/2013/475194.
11. Yao BZ, Hu P, Zhang MH, et al. A support vector machine with the Tabu search algorithm for freeway incident detection. *Int J Appl Math Comput Sci* 2014; 24: 397–404.
12. Yasdi R. Prediction of road traffic using a neural network approach. *Neural Comput Appl* 1991; 8: 135–142.
13. Quek C, Pasquier M and Lim BBS. POP-TRAFFIC: a novel fuzzy neural approach to road traffic analysis and prediction. *IEEE Trans Intell Transp Syst* 2006; 7: 133–146.
14. Xie Y, Lord D and Zhang Y. Predicting motor vehicle collisions using Bayesian neural network models: an empirical analysis. *Accid Anal Prev* 2007; 39: 922–933.
15. Kunt MM, Aghayan I and Noii N. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods. *Transport* 2011; 26: 353–366.
16. Deublein M, Schubert M, Adey BT, et al. A Bayesian network model to predict accidents on Swiss highways. *Infrastruct Asset Manag* 2015; 2: 145–158.
17. Kunt MM, Aghayan I and Noii N. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods. *Transport* 2011; 26: 353–366.
18. Yao BZ, Yao JB, Zhang MH, et al. Improved support vector machine regression in multi-step-ahead prediction for rock displacement surrounding a tunnel. *Scientia Iranica* 2014; 21: 1309–1316.
19. Li X, Lord D, Zhang Y, et al. Predicting motor vehicle crashes using support vector machine models. *Accid Anal Prev* 2008; 40: 1611–1618.
20. Li Z, Liu P, Wang W, et al. Using support vector machine models for crash injury severity analysis. *Accid Anal Prev* 2012; 45: 478–486.
21. Yang J and Zhao J. Road Traffic Safety Prediction Based on Improved SVM. In: *ICTE 2013: Safety, Speediness, Intelligence, Low-Carbon, Innovation*. 2013, pp.107–114.
22. Yu B, Yang ZZ and Yao JB. Genetic algorithm for bus frequency optimization. *J Transp Eng* 2010; 136: 576–583.
23. Yu B, Yang ZZ, Sun XS, et al. Parallel genetic algorithm in bus route headway optimization. *Appl Soft Comput* 2011; 11: 5081–5091.
24. Yu B, Zhu HB, Cai WJ, et al. Two-phase optimization approach to transit hub location – the case of Dalian. *J Transp Geography* 2013; 33: 62–71.
25. Li T, Yang Y, Wang Y, et al. Traffic fatalities prediction based on support vector machine. *Arch Transp* 2016; 39: 21–30.
26. Cai Z, Zhao H, Yang Z, et al. A modular spectrum sensing system based on pso-svm. *Sensors* 2012; 12: 15292.
27. Xiao Y, Kang N, Hong Y, et al. Misalignment fault diagnosis of DFWT based on IEMD energy entropy and PSO-SVM. *Entropy* 2017; 19: 6.
28. Ding R, Dong HB and Feng XB. Particle swarm optimization genetic algorithm applied in classification question. *Comput Eng* 2009; 35: 201–203.
29. Yang J, Zhang Z and Li Q. *Simultaneous optimization of start up scheme and pipe network for multi-source water injection system based on improved genetic particle swarm optimization algorithm*. Guilin, China: International Conference on Materials Engineering and Information Technology Applications 2015, vol. 8, 2015, pp.995–1024.
30. Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw* 1999; 10: 988–999.
31. Ahmadi MA, Galedarzadeh M and Shadizadeh SR. Low parameter model to monitor bottom hole pressure in vertical multiphase flow in oil production wells. *Petroleum* 2015; 2: 258–266.
32. Fazeli H, Soleimani R, Ahmadi MA, et al. Experimental study and modeling of ultrafiltration of refinery effluents using a hybrid intelligent approach. *Energy Fuels* 2013; 27: 3523–3537.
33. Dong B, Cao C and Lee SE. Applying support vector machines to predict building energy consumption in tropical region. *Energy Build* 2005; 37: 545–553.
34. Gan XS, Duanmu J and Cong W. Fatalness assessment of flight safety hidden danger based on support vector machine. *J Saf Sci Technol* 2010; 6: 206–210.
35. Guan Z, Song T and Qi L. An application of support vector machine in foundation settlement prediction. *Trans Shenyang Ligong Univ* 2008; 2: 024.
36. Kennedy J and Eberhart RC. Particle swarm optimization. In: *Proceedings of IEEE International Conference on Neural Networks,* Vol. 4, pp.1942–1948. IEEE Press.
37. Cao LJ and Tay FEH. Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on neural networks* 2003; 14: 1506–1518.
38. Chang LY. Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. *Saf Sci* 2005; 43: 541–557.
39. Cao C and Xu J. Short-Term Traffic Flow Predication Based on PSO-SVM. In: *International Conference on*

Transportation Engineering 2007, SiChun, China, 2015, pp.167–172. ASCE.
40. Ding S and Huang H. Research on parameters optimization of SVM based on swarm intelligence. *Int J Collaborative Intell* 2014; 1: 4.
41. Kunt MM, Aghayan I and Noii N. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods. *Transport* 2011; 6: 353–366.
42. Lin L, Wang Q and Sadek AW. A combined m5p tree and hazard-based duration model for predicting urban freeway traffic accident durations. *Accid Anal Prev* 2016; 91: 114–126.
43. Lee SB, Han DH and Lee YI. Development of freeway traffic incident clearance time prediction model by accident level. *J Korean Soc Transp* 2015; 33: 497–507.
44. Hong D, Kim J, Kim W, et al. Development of traffic accident prediction models by traffic and road characteristics in urban areas. In: *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 5, 2005, pp.2046–2061.

Appendix 1

	Traffic fatalities	Highway mileage	Vehicle number	Population size
Year	people	10,000 km	10,000	10,000
1981	22,499	89.75	234.73	100,072
1982	22,164	90.7	259.77	101,654
1983	23,944	91.51	283.88	103,008
1984	25,251	92.67	587.37	104,357
1985	40,906	94.24	655.65	105,851
1986	50,063	96.28	819.07	107,507
1987	53,439	98.22	1061.03	109,300
1988	54,814	99.96	1190.2	111,026
1989	50,441	101.43	1318.53	112,704
1990	49,243	102.83	1476.26	114,333
1991	53,204	104.11	1657.66	115,823
1992	58,723	105.67	1945.03	117,171
1993	63,551	108.35	2331.64	118,517
1994	66,362	111.78	2735.6	119,850
1995	71,494	115.7	3179.78	121,121
1996	73,655	118.58	3609.65	122,389
1997	73,861	122.64	4209.32	123,626
1998	78,067	127.85	4861.3	124,761
1999	83,529	135.17	5404.73	125,786
2000	93,853	167.98	6016.22	126,743
2001	105,930	169.8	6851.88	127,627
2002	109,381	176.52	7975.68	128,453
2003	104,372	180.98	9649.96	129,227
2004	107,077	187.07	10,783.44	129,988
2005	98,738	334.52	13,039.45	130,756
2006	89,455	345.7	14,528.90	131,448
2007	81,649	358.37	15,977.76	132,129
2008	73,484	373.02	16,988.77	132,802
2009	67,759	386.08	18,658.07	133,450
2010	65,225	400.82	20,706.13	134,091
2011	62,387	410.64	22,512.08	134,735
2012	59,997	423.75	24,102.03	135,404

Data source: http://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0S0D02&sj.