Research Journal of Applied Sciences, Engineering and Technology 6(6): 1065-1073, 2013
DOI:10.19026/rjaset.6.4014
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: October 31, 2012 Accepted: December 28, 2012 Published: June 30, 2013

Research Article

Traffic Accidents Forecasting Based on Neural Network and Principal Component Analysis

Yu-Rende, Zhang Qiang, Zhang-Xiaohong and Huo-Lianxiu
School of Transportation and Vehicle Engineering, Shandong University of Technology, China

Abstract: A number of factors may affect the occurrence of road traffic accidents and these factors may exist information overlap, which sometimes even obliterate the real traffic characteristics and the inherent laws. In order to improve the forecasting accuracy of traffic accident forecasting model, this study proposed a new traffic accidents forecasting method based on neural network and principal component analysis. Compared with other models, the results show the model baed on neural network and principal component analysis is more accuracy.

Keywords: BP neural network, forecasting, principal component analysis, road traffic accidents

INTRODUCTION

Road traffic accidents forecast methods mainly include gray forecasting method, time series method, regression analysis and BP neural network at present (Guo-Hong, 2006; Dong-Ping, 2007; Xiang-Yong, 2004). Domestic scholars have done a great quantity of work on road traffic accidents forecast. The forecasting model for fatalities established by the Beijing Transportation Research Institute, the forecasting model for traffic accident fatalities in Tianjin established by the Research Institute of Tianjin city; the time series decomposition prediction method by Jilin University, the traffic accidents time series models by Beijing University of Technology and the traffic accidents forecasting based on neural network by Shandong University of Technology are typical (Sayed, 2000; Ren-De *et al*., 2008; Xiang-Yong, 2003).

Road traffic accident caused by various factors, these factors may exist the information of overlap that sometimes effaces the really characteristics and inherent law about traffic accidents. So this study will bring principal component analysis into the road traffic accident forecast, eliminate some overlap informations, combined with BP neural network to forecast the road traffic accident (the BP neural network based on PCA) and compare the predicted results with the BP neural network prediction results that wasn't conducted of principal component analysis. And draw the conclusion: the BP neural network based on PCA have been significantly improved than BP neural network in the prediction precision. And draw the conclusion: the BP neural network based on PCA have been significantly improved than BP neural network in the prediction precision.

LITRETURE REVIEW

In the road traffic accident for empirical research, in order to more fully and accurately reflect the characteristics and laws of development of the traffic accidents, we tend to consider multiple indicators of impact the traffic accidents, these indicators are also known as variable in the multivariate statistics. This produces the following problems: on the one hand in order to avoid missing important information will be considered as much as possible indicators, while on the other hand, with the increase in consideration indicators increase the complexity of the research accident, at the same time because the indexes is the reflection of the traffic accident, inevitably causes a large number of overlapping information, this information overlap sometimes effacement the really characteristics and inherent law of the traffic accident. Therefore, we hope that fewer variables involved in a traffic accident research, while get more information. The principal component analysis is a multivariate statistical methods, the study of how to through the original variable of the few linear combination to explain the original variable most information.

Correlation between accidents involving many variables, there must be a co-factors of play a dominant role, According to this, the original variable correlation matrix or the covariance matrix of the internal structure of the relationship, use of a linear combination of the original variables to form several indicators (principal components), keep the original variable in the main information under the premise of dimensionality reduction and simplify the problem effect, makes it easier to grasp the principal contradiction in the study of a traffic accident problem. Generally speaking, the

Corresponding Author: Yu Ren-de, School of Transportation and Vehicle Engineering, Shandong University of Technology, China, Tel.: 13409012886

principal components and the original variables using principal component analysis follows the basic relationship between:

- The number of principal components is far less than the number of the original variables
- The principal components retain the vast majority information of variable
- Each principal component is the linear combination of the original variables
- Each principal component is irrelevant

Through the principal component analysis of road traffic accident influence factor, we could find some main compositions from complicated relationship between the variables, which can quantitative analyses effectively with lots of statistical data, reveal the inner relationship between variables and get on deep inspiration between traffic accident characteristics and the law of development, lead the research work further.

PRINCIPAL COMPONENT ANALYSIS OF ROAD TRAFFIC ACCIDENTS

The export of sample principal components: In the study of traffic accidents, overall covariance matrix $\sum$ and correlation matrix R is usually unknown, so need through the sample data to estimate. Set with *m* samples and each sample has *n* indicators, so get a total of *mn* data and the original data matrix as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{1}$$

In which:

$$\begin{aligned} &\bar{x}_i = \frac{1}{m}\sum_{k=1}^{m} x_{ki}, i = 1,2,\ldots, p \\ &S = \frac{1}{m-1}\sum_{k=1}^{m}(x_{ki} - \bar{x}_i)(x_{ki} - \bar{x}_i)^T \end{aligned} \tag{2}$$

$$R = (r_{ij})_{n\times n}, r_{ij} = \frac{S_{ij}}{\sqrt{S_{ii}S_{jj}}} \tag{3}$$

S is the sample covariance matrix, unbiased estimate of the overall covariance matrix $\sum$, *R* is the correlation matrix of the sample, the estimates for the overall correlation matrix.

Known from the foregoing discussion, if the original data array *X* is standardized processes. The covariance matrix obtained from the matrix *X* is the correlation matrix, namely *S* and *R* exactly the same. Because the covariance matrix solving principal component process with a correlation matrix solution based on principal component process is consistent, below we introduced only by correlation matrix *R* is based on principal component.

Principal component y covariance for:

$$\mathrm{cov}(Y) = u^T \mathrm{cov}(X)u = u\sum u^T = \wedge \tag{4}$$

Which ^ is a diagonal matrix:

$$\wedge = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Assume that material matrix X for already after standardization of the data matrix, while the correlation matrix instead of the covariance matrix and the type and can be expressed as: $uRu^T = \wedge$, use u^T left multiplied type, get $Ru^T = u^T\wedge$, then

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \tag{5}$$

Expand all of the above equation to get the n^2 equations, here only consider the n equation derived from the first column in the matrix multiplication:

$$\begin{cases} r_{11}u_{11} + r_{12}u_{12} + \cdots + r_{1n}u_{1n} = u_{11}\lambda_1 \\ r_{21}u_{11} + r_{22}u_{12} + \cdots + r_{2n}u_{1n} = u_{12}\lambda_1 \\ \cdots\cdots\cdots\cdots \\ r_{n1}u_{11} + r_{n2}u_{12} + \cdots + r_{nn}u_{1n} = u_{1n}\lambda_1 \end{cases} \tag{6}$$

Finishing been:

$$\begin{cases} (r_{11} - \lambda_1)u_{11} + r_{12}u_{12} + \cdots + r_{1n}u_{1n} = 0 \\ r_{21}u_{11} + (r_{22} - \lambda_1)u_{12} + \cdots + r_{2n}u_{1n} = 0 \\ \cdots\cdots\cdots\cdots \\ r_{n1}u_{11} + r_{n2}u_{12} + \cdots + (r_{nn} - \lambda_1)u_{1n} = 0 \end{cases} \tag{7}$$

In order to get the top homogeneous equation of non-zero solution, according to the theory of linear equations is known, requires the coefficient matrix determinant is 0, that is:

Table 1: Each index's original statistical data on city A

Year	GDP(one hundred million yuan)	The per capita income (yuan)	The total retail sales of social consumer (one hundred million yuan)	The resident population (Ten thousand people)	The amount of vehicle ownership (ten thousandcars)	Not seized vehicles (ten thousand cars)	Total passenger traffic (million man-time)
1995	149.12	4685	59.229	305.57	4.8973	1.711	418.1371
1996	171.19	4866	72.9769	318.85	6.8019	3.0127	541.3785
1997	197.51	5343	82.7542	321.26	7.1463	3.5045	562.2507
1998	219.55	5369	89.5491	325.87	8.2687	3.6846	143.0193
1999	237.59	6082	97.0716	331.57	9.2887	3.9182	157.03
2000	264.81	6453	108.5266	337.45	9.5739	4.2566	164.58
2001	302.75	6909	121.6556	341.29	11.2707	4.5085	173.34
2002	336.37	7306	136.5816	346.27	12.3872	5.152	183.0499
2003	380.92	7985	153.6707	348.7	18.6141	6.8782	185.108
2004	443.62	8989	175.5226	350.85	21.4609	11.509	192.3833

Year	Total freight(One million tons)	Light controlled intersections(number)	The number of traffic police (number)	Urban road length (km)	Road area(Ten thousand square meter)	The number of the driver (ten thousand people)	The number of accident(number)
1995	28. 3694	46	659	963	745	11.9556	191
1996	42.28	46	700	775	566	14.183	277
1997	39.9629	46	717	775	566	16.8791	247
1998	37.1104	46	737	777	572.4	19.6753	207
1999	41.3968	46	765	779	579.6	21.9909	238
2000	46.43	46	749	781	585	25.0338	253
2001	48.85	46	743	800	631.54	29.0066	309
2002	51.3148	55	729	781	609.3	34.1664	305
2003	53.178	56	787	787	621.3	39.6696	309
2004	57.3421	64	714	953	1040.3	48.5918	285

$$\begin{vmatrix} r_{11}-\lambda_1 & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22}-\lambda_1 & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn}-\lambda_1 \end{vmatrix} = 0 \quad (8)$$

That is /R- λ_1I/ = 0

For λ_2, …, λ_n can get completely similar equation ,thus, the variance λ_i(i = 1, 2, …n) of new variables (main component) we asked is the n roots of |R- λ_1I| = 0, λ is the characteristic value of the correlation matrix, the respective u_{ij} is a component of the feature vector.

R is a positive definite matrix, its characteristic roots are non-negative real numbers , arranged by size order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$, its corresponding feature vector referred to as γ_1, γ_2 , … γ_n, The relative to Y_1 variance for: var(Y_1) = var(γ_1^T X) = λ_1

Similarly have

var (Y_i) = var (γ_i^T X) = λ_i

That is for Y_1 have maximum variance, Y_2 the second largest varianc, …, and covariance for:

$$\begin{aligned} \operatorname{cov}(Y_i, Y_j) &= \operatorname{cov}(\gamma_i^T X, \gamma_j^T X) = \gamma_i^T R \gamma_j^T \\ &= \gamma_i^T \Big(\sum_{\alpha=1}^{n} \lambda_\alpha \gamma_\alpha \gamma_\alpha^T\Big)\gamma_j \\ &= \sum_{\alpha=1}^{n} \lambda_\alpha (\gamma_i^T \gamma_\alpha)(\gamma_\alpha^T \gamma_j) = 0 \quad (i \neq j) \end{aligned}$$

This can have, new variables (principal component) Y_1, Y_2, … Y_n not related to each other and λ_i is the variance of Y_i, then $Y_1 = \gamma_1^T$ X, $Y_2 = \gamma^T_2$ X, …, $Y_n = \gamma^T_n$ X respectively referred to as a first, second,…, *n* main ingredient. By the process of seeking principal components that, The direction of the Principal component in the geometry is actually *R* direction of the eigenvectors; The variance contribution of main components is the corresponding characteristic value of *R*. In this way, the process of solving the principal component of the sample data is actually converted into eigenvalues and eigenvectors process of seeking the correlation matrix or the covariance matrix (Xue, 2004; Shu and Jian-She, 2008; Jian-Xi and De-Yan, 2011).

Index selection of principal component analysis and principal component determination: Using the index statistics data of city A from 1995 to 2004 (Table 1) predicts the road traffic accident. Data are selected the 13 factors which affect traffic accident. Namely: GDP (X1), the per capita income (X2), the total retail sales of social consumer goods (X3) , the resident population (X4), the amount of vehicle ownership (X5), not seized vehicles (X6), total passenger traffic (X7), total freight (X8), light controlled intersections (X9), the number of traffic police (X10), urban road length (X11), road area (X12), the number of the driver (X13).

Thirteen indicators of the impact of road traffic accidents are subjected to principal component analysis by using SPSS13.0 and then put out the results as shown in Table 2-5. From Table 2 we can see that there is extremely significant relationship between GDP, the per capita income , the total retail sales of social consumer goods , the resident population, the amount of vehicle ownership, not seized vehicles, total freight, light controlled intersections and the number of the

Table 2: The matrix of correlation coefficient

	Index						
Index	GDP	The per capita income	The total retail sales of social consumer goods	The resident population	The amount of vehicle ownership	Not seized vehicles	Total passenger traffic
GDP	1.000	0.996	0.999	0.944	0.972	0.923	-0.610
The per capita income	0.996	1.000	0.994	0.940	0.967	0.919	-0.605
The total retail sales of social consumer goods	0.999	0.994	1.000	0.953	0.970	0.918	-0.603
The resident population	0.944	0.940	0.953	1.000	0.872	0.793	-0.692
The amount of vehicle ownership	0.972	0.967	0.970	0.872	1.000	0.953	-0.514
Not seized vehicles	0.923	0.919	0.918	0.793	0.953	1.000	-0.422
Total passenger traffic	-0.610	-0.605	-0.603	-0.692	-0.514	-0.422	1.000
Total freight	0.923	0.921	0.935	0.952	0.882	0.837	-0.465
Light controlled intersections	0.882	0.871	0.876	0.712	0.914	0.923	-0.323
The number of traffic police	0.360	0.365	0.380	0.619	0.235	0.171	-0.674
Urban road length	0.162	0.183	0.125	-0.141	0.229	0.339	0.070
Road area	0.564	0.577	0.537	0.295	0.620	0.754	-0.141
The number of the driver	0.999	0.995	0.997	0.928	0.978	0.936	-0.593

	Index					
Index	Total freight	Light controlled intersections	The number of traffic police	Urban road length	Road area	The number of the driver
GDP	0.923	0.882	0.360	0.162	0.564	0.999
The per capita income	0.921	0.871	0.365	0.183	0.577	0.995
The total retail sales of social consumer goods	0.935	0.876	0.380	0.125	0.537	0.997
The resident population	0.952	0.712	0.619	-0.141	0.295	0.928
The amount of vehicle ownership	0.882	0.914	0.235	0.229	0.620	0.978
Not seized vehicles	0.837	0.923	0.171	0.339	0.754	0.936
Total passenger traffic	-0.456	-0323	-0.674	0.070	-0.141	-0.593
Total freight	1.000	0.761	0.469	-0.097	0.364	0.913
Light controlled intersections	0.761	1.000	-0.031	0.424	0.748	0.903
The number of traffic police	0.469	-0.031	1.000	-0.671	-0.360	0.322
Urban road length	-0.097	0.424	-0.671	1.000	0.860	0.200
Road area	0.364	0.748	-0.360	0.860	1.000	0.598
The number of the driver	0.913	0.903	0.322	0.200	0.598	1.000

Table 3: Analysis of principal component extraction of variance decomposition

	Initial Eigenvalues			Extraction sums of squared loadings		
Component	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	9.218	70.910	70.910	9.218	70.910	70.910
2	2.694	20.723	91.633	2.694	20.723	91.633
3	0.711	5.470	97.103			
4	0.183	1.409	98.512			
5	0.100	0.770	99.283			
6	0.050	0.388	99.670			
7	0.033	0.251	99.921			
8	0.009	0.068	99.989			
9	0.001	0.011	100.000			
10	0.000	0.000	100.000			
11	0.000	0.000	100.000			
12	0.000	0.000	100.000			
13	0.000	0.000	100.000			

Table 4: Public factor variance change table

Index	Initial	Extraction
GDP (X1)	1.000	0.993
The per capita income (X2)	1.000	0.988
The total retail sales of social consumer goods (X3)	1.000	0.993
The resident population (X4)	1.000	0.990
The amount of vehicle ownership (X5)	1.000	0.960
Not seized vehicles (X6)	1.000	0.941
Total passenger traffic (X7)	1.000	0.552
Total freight (X8)	1.000	0.891
Light controlled intersections (X9)	1.000	0.919
The number of traffic police (X10)	1.000	0.897
Urban road length (X11)	1.000	0.874
Road area (X12)	1.000	0.919
The number of the driver (X13)	1.000	0.996

driver in city A. The direct correlation between many variable is stronger. Prove that they exist on the information of the overlap.

The number of principal components is former m, which the extraction principle is the corresponding characteristic value greater than 1. Characteristic value to some extent can be regarded as a size of principal component influence strength index. If the character values are less than 1, show that the explanation of the principal components is smaller than the average explanation of a directly into original variable, so we can use characteristic values greater than 1 as the inclusion criteria. Through the Table 3, extraction two

Table 5: The matrix of original factor's load

	Component	
Index	1	2
GDP (X1)	1.000	0.993
The per capita income (X2)	1.000	0.988
The total retail sales of social consumer goods (X3)	1.000	0.993
The resident population (X4)	1.000	0.990
The amount of vehicle ownership (X5)	1.000	0.960
Not seized vehicles (X6)	1.000	0.941
Total passenger traffic (X7)	1.000	0.552
Total freight (X8)	1.000	0.891
Light controlled intersections (X9)	1.000	0.919
The number of traffic police (X10)	1.000	0.897
Urban road length (X11)	1.000	0.874
Road area (X12)	1.000	0.919
The number of the driver (X13)	1.000	0.996

Table 6: Principal component F over the years

Year	F1	F2	Year	F1	F2
1995	1925.4724	489.6046	2000	2596.831	157.4357
1996	1938.5395	306.0319	2001	2774.8622	183.878
1997	2105.3221	293.7045	2002	2919.4712	161.993
1998	2212.279	175.0506	2003	3168.4449	166.3259
1999	2459.9827	154.2773	2004	3625.4685	441.3045

principal components, m = 2. From Table 5, we can see that the load of GDP, the per capita income, the total retail sales of social consumer goods, the resident population, the amount of vehicle ownership, not seized vehicles, total freight, light controlled intersections and the number of the driver in the first principal component is higher. The first principal component basic reflects the index information. The load of urban road length, road area in the second principal component is higher, which explains that the second principal component basic reflects the information of two indexes. So we should adopt two new variable to instead of the original thirteen variable.

Then 13 factors will be into two principal components, with F1 and F2 to say. With the data in Table 5 divided by principal component corresponding eigenvalue and then open square root. We get the corresponding coefficient of every index in the principal component, finally get the expression of principal component analysis is as follows:

$$F1 = 0.328X1+0.327X2+0.327X3+0.306X4+0.322X5+0.312X60.199X7+0.302X8+0.295X9+0.114X10+0.065X11+0.200X12+0.329X13 \quad (9)$$

$$F2 = -0.021X1-0.016X2-0.040X3-0.214X4+0.049X5+0.128X6+0.263X7-0.134X8+0.207X9-0.537X10+0.557X11+0.452X12+0.005X13 \quad (10)$$

Bring the various indexes over the years into principal component expression and then get the principal component value over the years (Table 6).

THE FORECAST OF ROAD TRAFFIC ACCIDENT BASED ON NEURAL NETWORK

The multilayer perceptron of BP algorithm is the most widely used neural network so far. Multilayer perceptron includes the input layer, hidden layer and output layer (Jian-Xi and De-Yan, 2011; Li-Qun, 2007; Hecht-Nielsen, 1989; Xiu, 2007).

The forecast of road traffic accident based on BP neural network:

- **The selection of sample:** In the course of modeling, the sample should put into two parts : the training sample and the test sample, test sample is mainly to check out and test the network model, this study selects the related data of A city from1995 to 2004 (Table 1), the data of 1995-2002 as the training sample, 2003 and 2004 as the test sample.

Among them: each year's gross national product (X1), the per capita income (X2), the total retail sales of social consumer goods (X3), permanent population (X4), motor vehicle ownership (X5), not inspection vehicle (X6), the passengers amount (X7), freight amount (X8), light controlled intersection (X9), the number of traffic police (X10), the length of urban roads (X11), roads area (X12), the number of drivers (X13)as the input samples.

- **Accident frequency as the output sample:** The formula $y = (x- x_{min})/(x_{max} - x_{min})$ should be normalized processing, the sample data of normalized processing is in Table 2.
  The determination of the number of hidden layers and hidden layer neurons of the BP neural network: Using an input layer, a middle layer (hidden layer) and an output layer of the BP neural network, the number of hidden layer neurons can be determined according to the formula $n_1 = \sqrt{n+m}+a$, among them: n is the input neuron number, m is an output neuron number, a is a constant between 0 and 10. Here n =13, m = 1, the value range of n_1 is [4, 14], network training times choose 1000, precision choose 0.001.
  After training, when hidden layer neurons is 9, the effect of network prediction is the best (Fig. 1 and 2):

The accident forecast of BP neural network based on PCA:

- **The selection of sample:** Also choose the data of A city from 1995 to 2002 as the training sample, the data of 2003 and 2004 as the testing sample. Select the each year's principal component value F1 and F2 as the input samples, accident frequency as the output sample.

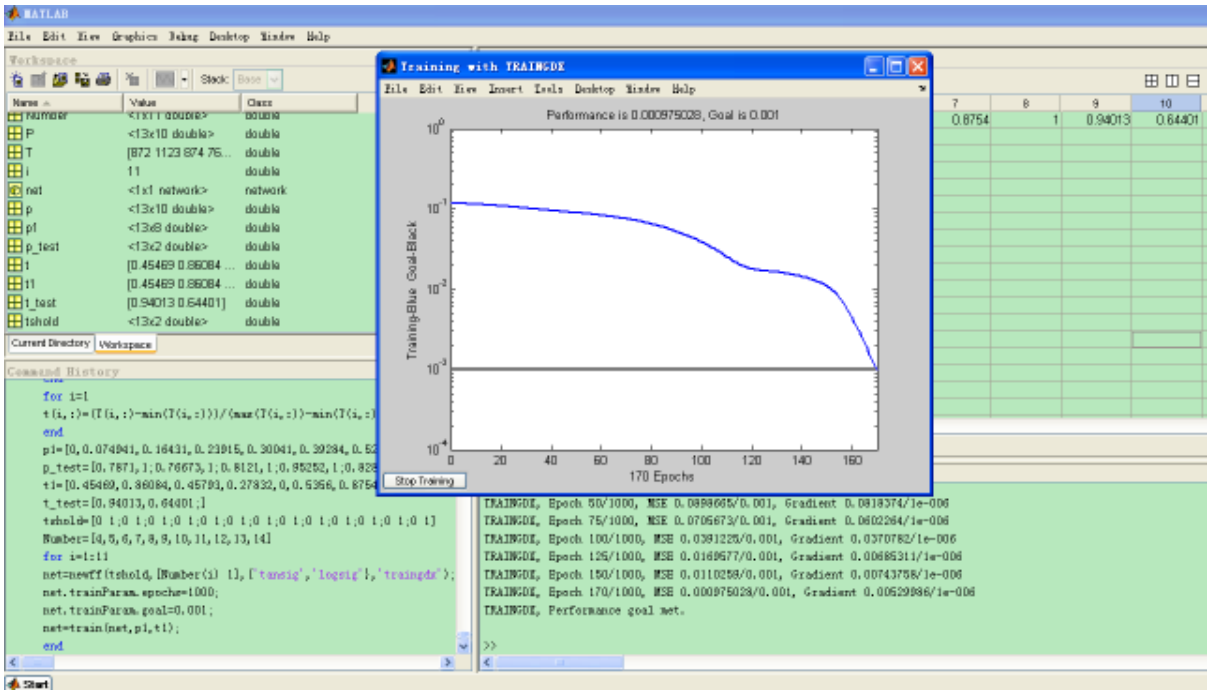

Fig. 1: The train diagram of BP neural network prediction accident

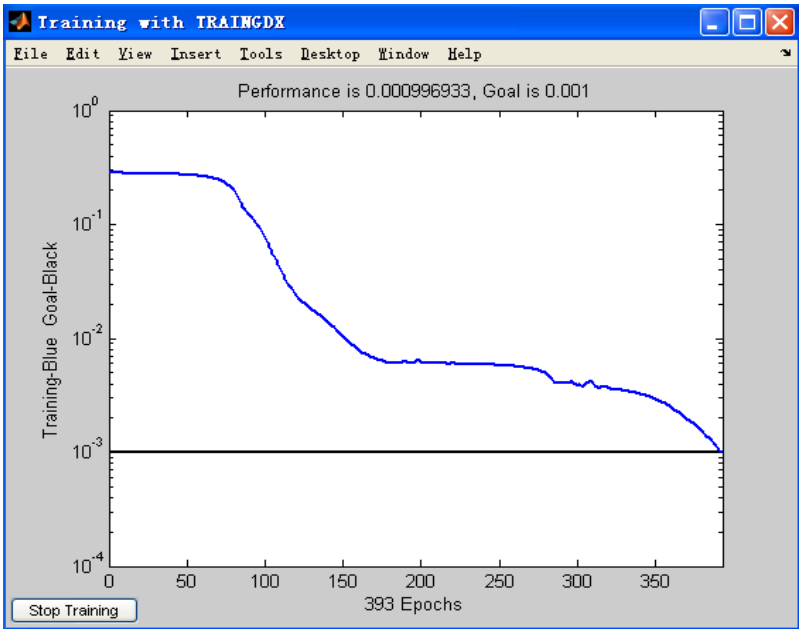

Fig. 2: Training error curve (hidden layer neurons: 9)

- **The pretreatment of learning sample:** The formula $y = (x- x_{min})/(x_{max} - x_{min})$ should be normalized, the sample data of normalized processing is in Table 7.
- **The determination of the number of hidden layers and hidden layer neurons of the BP neural network:** The determination of the number of hidden layer neurons also based on the formula $n_1 = \sqrt{n+m}+a$, among them: n is the input neuron number, m is an output neuron number, α is a constant between 0 and 10. Here n = 2, m = 1, the value range of n_1 is [2, 12], network training times choose 1000, precision choose 0.001.

After training, when hidden layer neurons is 12, the network prediction effect is the best (Fig. 3 and 4).

Table 7: The sample data sheet after normalize

Year	Index: Gross national product (one hundred million yuan)	Per capita income (yuan)	The total retail sales of social consumer goods) one hundred million yuan)	Permanent population (Ten thousand people)	Motor vehicle ownership (The thousand car)	Not inspection vehicle (The thousand car)	The passengers amount (Millions of people)
1995	0	0	0	0	0	0	0.65624
1996	0.074941	0.042054	0.11821	0.29329	0.11499	0.13285	0.95021
1997	0.16431	0.15288	0.20229	0.34651	0.13578	0.18304	1
1998	0.23915	0.15892	0.26071	0.44832	0.20354	0.20142	0
1999	0.30041	0.32458	0.3254	0.5742	0.26512	0.22526	0.03342
2000	0.39284	0.41078	0.4239	0.70406	0.28234	0.25979	0.051429
2001	0.52166	0.51673	0.5368	0.78887	0.38478	0.2855	0.072324
2002	0.63582	0.60897	0.66515	0.89885	0.45219	0.35117	0.095486
2003	0.7871	0.76673	0.8121	0.95252	0.82813	0.52734	0.10039
2004	1	1	1	1	1	1	0.11775

Year	Index: Freight amount (One million tons)	Light controlled intersection (pcs)	The number of traffic police (pcs)	The length of urban roads (km)	Roads area (ten thousand square meters)	The number of drivers (Ten thousand people)	Accident frequency (pcs)
1995	0	0	0	1	0.3774	0	0.52535
1996	0.4813	0	0.38679	0	0	0.060798	0.87887
1997	0.40015	0	0.54717	0	0	0.13439	0.37183
1998	0.3017	0	0.73585	0.010638	0.013494	0.21071	0.37183
1999	0.44964	0	1	0.021277	0.028674	0.27392	0.12958
2000	0.62337	0	0.88679	0.031915	0.040059	0.35697	0.59577
2001	0.70689	0	0.84906	0.13298	0.13818	0.46541	0.89155
2002	0.79197	0.5	0.79245	0.031915	0.091292	0.60625	1
2003	0.85631	0.55556	0.66038	0.06383	0.11659	0.75646	0.94013
2004	1	1	0.51887	0.94681	1	1	0.64401

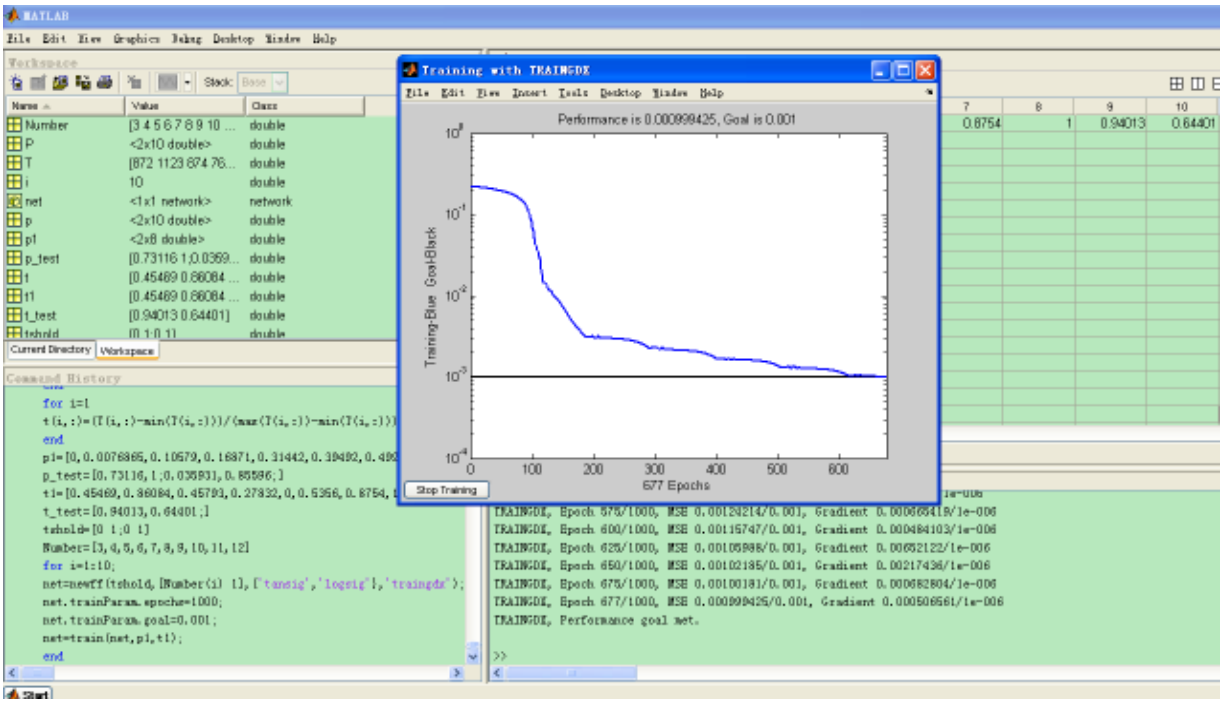

Fig. 3: The train diagram of BP prediction accident numbers based on principal component analysis

- **The output of prediction result:** From the third step, it is known that when the number of hidden layer neurons is 12, the training effect is the best, so we select the hidden layer neurons is 12 for accident forecast.
- **The reverse normalized processing of the prediction result:** Reverse normalized processing can according to the formula $x = x_{min} + y(x_{max} - x_{min})$ and the normalized data can react to the original data. Based on the reverse normalized processing we can get practical predicted value shown in Table 8.
- **The contrast of predicted results:** The predicted result of accident numbers before and after the principal component analysis is in Table 9 and error graph (Fig. 5).

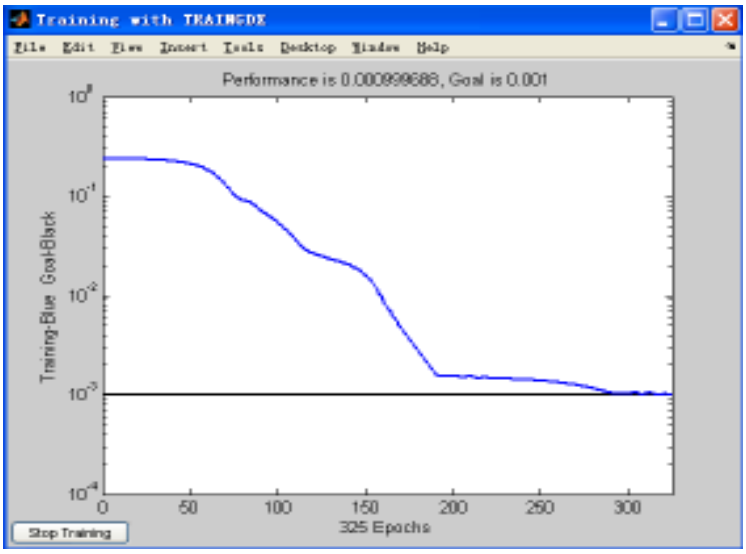

Fig. 4: training error curve (hidden:layer neurons: 12)

Table 8: The sample list after reverse normalized

	1995	1996	1997	1998	1999
F1	0	0.00769	0.10579	0.16871	0.31442
F2	1	0.45256	0.41579	0.06195	0
Accident frequency	0.52535	0.87887	0.52817	0.37183	0.12958
	2000	2001	2002	2003	2004
F1	0.39492	0.49964	0.58471	0.73116	1
F2	0.00942	0.08828	0.02301	0.03593	0.85596
Accident frequency	0.59577	0.89155	1	0.94013	0.64401

Table 9: The prediction results of BP neural network and the BP neural

Year	Actual value	BP prediction results	RPE (%)	BP prediction based on PCA	RPE (%)
2003	1172	1202	2.56	1173	0.10
2004	899	919	2.22	911	1.33

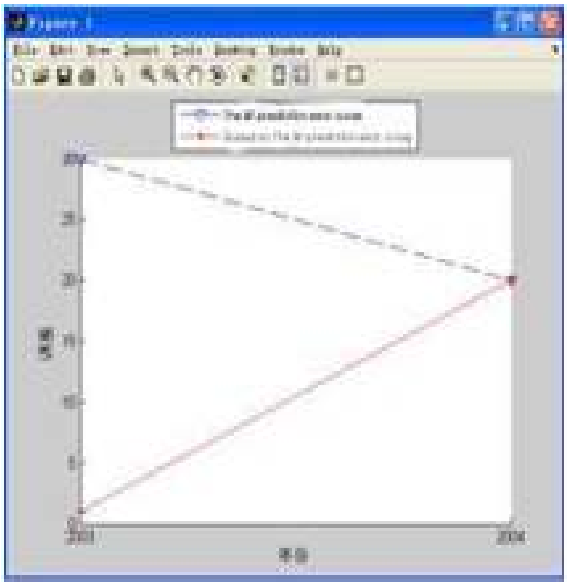

Fig. 5: The error curve of BP neural network and the BP neural network based on PCA

From the prediction results of road traffic accidents number, we can see the prediction precision is obviously better than it in a BP neural network because we eliminate some overlap information between variables of BP neural network based on PCA at the beginning of the prediction.

CONCLUSION

Road traffic accident caused by various factors, these factors may exist the information of overlap that sometimes effaces the really characteristics and inherent law about traffic accidents. So this study will bring principal component analysis into the road traffic accident forecast, eliminate some overlap informations, combined with BP neural network to forecast the road traffic accident (the BP neural network based on PCA) and compare the predicted results with the BP neural network prediction results that wasn't conducted of principal component analysis. And draw the conclusion: the BP neural network based on PCA have

been significantly improved than BP neural network in the prediction precision.

REFERENCES

Dong-Ping, W., 2007. Research and Application of Road Accidents Forecasting Method. Chong Qing Jiaotong University.

Guo-Hong, N., 2006. A Method of Forecasting the Traffic Accidents Based on ANN. Chang'an University.

Hecht-Nielsen, R., 1989. Theory of the back-propagation neural networks. Proceeding of the International Joint Conference on Neural Networks, pp: 596-611.

Jian-Xi, Z. and Z. De-Yan, 2011. The dam deformation forecasting of BP neural networkand principal component analysis. J. East China Inst. Technol. Nat. Sci., 34(3): 288-292.

Li-Qun, H., 2007. The Theory of BP Neural Network and Design and Application. Chemical Industry Press, Beijing.

Ren-De, Y., L. Fang and S. Peng, 2008. The prediction of road traffic based on neural network. Math. Prac. Theory, 38(6): 119-126.

Sayed, T., 2000. Applications of accident prediction models. Annual Conference Abstracts Canadian Society for Civil Engineering, pp: 112-114.

Shu, W. and C. Jian-She, 2008. Accident prediction model for tailings reservoir based on regression analysis of SPSS. China Saf. Sci. J., 18(12): 34-40.

Xiang-Yong, L., 2003. JIANG Ge-fu. Grey-markov model for forecasting road accidents. J. Highway Transport. Res. Dev., 20(4): 98-104.

Xiang-Yong, L., 2004. Research on Forecasting Methods of Road Accidents. Southwest Jiaotong University.

Xiu, W., 2007. Forecast of Road Accidents based on Grey Theory and Neural Networks. Shan Dong University of Science and Technology.

Xue, W., 2004. Satistics Analysis Method and Application by SPSS. Publishing House of Electronics Industry, China.