2011 Third International Conference on Measuring Technology and Mechatronics Automation

A New Model for Residual Value Prediction of the Used Car Based on BP Neural Network and Nonlinear Curve Fit

Shen Gongqi, Wang Yansong, Zhu Qiang

College of Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China
dieatiom21@163.com

***Abstract*—A new model for predicting the residual value of the private used car with various conditions, such as manufacturer, mileage, time of life, etc., was developed in this paper. A comprehensive method combined by the BP neural network and nonlinear curve fit was introduced for optimizing the model due to its flexible nonlinearity. Firstly, some distribution curves of residual value of the used cars were analyzed in time domain. Then, the BP neural network (NN) was established and used to extract the feature of the distribution curves in various conditions. A set of schemed data was used to train the NN and reached the training goal. Finally, the schemed data as inputs and the NN outputs were organized for nonlinear curve fit. Conclusion was drawn that the newly proposed model is feasible and accurate for residual value prediction of the used cars with various conditions.**

Keywords-Used Car; BP Neural Network; Nonlinear Curve Fit

I. INTRODUCTION

Owing to the rapid economy development in China, the national quantity of possessed cars has been increasing continuously. The gross national products of new cars have reached 10 million in 2009. With the time being, these cars are getting older and after several years the majority of them will flow into the used car trade market, which will become a potential and prosperous market. Now the national trade volume of the used car is increasing at the speed of 30% each year. To confront this extending market, it is very useful to develop a computation model (or an algorithm) for predicting residual value of the used cars with various conditions. Till now, however, researches on characteristics of the used cars and their residual value have just begun and there is no in-situ model is widely accepted and completely suitable for the marketing prediction. As a commodity, cars have their special characteristics, including high unit value, long life span, complex structure, advanced engineering technologies involving in mechanism, electronics, material science, etc., and the complicated conditions of usage such as the different owner habits, the maintenance level, and the high purchase tax [1]. When one estimate the residual value of a used car, the above characteristics should be carefully considered and determined whether and/or how they affect the residual price of the used cars.

From the above discussions, a new method based on the BP neural network and nonlinear curve fit was developed in this paper. This model combined the principals of the economic theory and the datasets of the trade records of the used cars can estimate the residual value of the used car directly and can be used in relative used car trade markets, companies and individuals.

II. BACKGROUD THEORY OF USED CAR

Some modeling assumptions were set as following: (1) the used car we mentioned here was private car only, not including the car used as commercial car such as taxi or as the chauffeur-driven car in government or company, which is different from the private car in appraisal conditions. (2) The data we acquired were from the used car trade market in Shanghai which may be a little different from the other place.

A. Factors analysis

According to the analysis of the recorded data, several factors which were considered effect on the used car residual value substantially were chosen as following: the car manufacturer, model, mileage, age, maintenance record, physical condition, market occupancy, after-sale service, and the driving habit of owner.

The above-mentioned factors were the original factors which were interconnected with each other, affecting the residual value coherently. If the relationship among these factors was neglected and the function of factors to residual value was calculated respectively, it will lower the accuracy of the prediction model. To solve this issue, according to the literature [2], four strengthened factors were adopted to replace the original factors, the strengthened factors were shown in Table 1.

B. Distribution curve of residual value

In the common sense, a car starts to lose its value since the day it was bought, at the beginning of the first several years, the slope of distribution curve of residual value over time is sharp, which then slowly turn to smooth. It is widely accepted in the field of used car appraisal that the lifetime of the car ,usually 15 years, is divided into 5 parts sequentially and averagely, each part constitutes the 5/15, 4/15, 3/15, 2/15, 1/15 of the whole new car value respectively [3]. It's usually used as an assistant standard for appraisal agent to assess the used car approximately. According to discussion above-mentioned, an assumption was drawn that the gradient of the price declination is proportional to the residual value [4]. It's shown as

$$d_P / d_t = -kP(t) \qquad (k > 0)$$

978-0-7695-4296-6/11 $26.00 © 2011 IEEE
DOI 10.1109/ICMTMA.2011.455

Table 1. The details of the strengthened factors

Factor name	description
Physical factor	The physical factor involved with age and mileage mainly which cannot be avoided during usage. With the age and mileage increasing, the car's physicochemical condition is inevitably worsening and gradually breaking down.
Reputation factor	The reputation factor involved with the original factors related to manufacturer. A car made by a famous car manufacturer which enjoys the reputation of high product quality and good after-sale service usually has a higher residual value than a car of the same physical factor, but made by inferior manufacturer.
Economic factor	The economic environmental factor involved with the policies drawn by government which affect the residual value, such as purchase tax, new environmental protection policy.
Wear factor	The wear factor involved with the original factors related to the pre-owner habit which affect the residual value such as the driving habits, the running condition and maintenance level of the car.

Solving the above differential equation, it can be rewritten as:

$$\int_{p(t)}^{a} \frac{dp}{p} = \int_{t}^{0} -k dt$$

$$\ln(\frac{a}{p(t)}) = -k(0-t)$$

$$p(t) = a * e^{-kt}$$

Where, $p(t)$ is the residual value of used car at the age of t, a is the replacement value of the car, k is the characteristic coefficient of the value declination which will be considered in the next section. e^{-kt} is the newness rate of the used car.

The above equation suggests that the distribution curve of the residual value is subjecting to the exponential declination. Here, the extreme condition over time should be discussed. When the age of the used car t is close to zero which means this used car is almost a new car. The newness rate of the used car e^{-kt} will be close to 1, but in fact it won't happen due to the traditional philosophy of people to the second hand commodity. In general, the maximum residual value of the used car is 90% of the new car value. When the age of the used car t is over 15 years which means the newness rate of the used car e^{-kt} will be close to 0, but in fact this used car still could be sold at least at the price of 5% or more of the new car value due to its metal and other materials which could be recycled. We set the up and low limit threshold of the newness rate e^{-kt} as 0.90 and 0.05 respectively in the upcoming model.

III. BP NEURAL NETWORK

A. Theory of BP neural network

Neural networks are composed of simple and densely elements operating in parallel which inspired by the biological neurons. Adjusting the weights to the appropriate value guided by the training algorithm and the dataset, the trained neural network acquired the ability to predict the outcomes based on the input data. The one of the most important advantages of neural network is relationship between inputs and outputs needn't be completely known prior to the modeling, so it is suitable for extracting the feature between the residual value and the factors which are complexly interconnected.

B. Establishment of BP neural network

1) Input data pre-processing

The input data should be carefully selected and pre-processed before being fed into the network. Owing to the limited numbers of the trade record dataset, here a new classification was developed that the used cars were classified into 15 types which is a combination of the quality classification with 'domestic car', 'joint venture company (JVC) made car' and 'imported car' and price classification with 'less than 80 k CNY', '80-150 k', '150-250 k', '250-350 k', 'more than 350 k'. Note that there are seldom imported cars of which the price is less than 80k CNY, so this case could be omitted. The new classification with 14 types of car was used to define the used car. The new classification was shown in Table. 2. The other factors with large numerical value such as age, mileage were normalized which can increase the velocity of convergence during the training phrase. The rest factors described in words such as the maintenance level, for computational convenience, were transformed into numbers [5] [6].

Table 2. The new classification for used car

Quality classification	Price classification(CNY)	New classification number
Imported car	More than 350 k	1
	250~350 k	2
	150~250 k	3
	80~150 k	4
Joint venture company made car (JVC car)	More than 350 k	5
	250~350 k	6
	150~250 k	7
	80~150 k	8
	Less than 80 k	9
Domestic car	More than 350 k	10
	250~350 k	11
	150~250 k	12
	80~150 k	13
	Less than 80 k	14

2) Structure of the neural network

The BP neural network adopted in this paper had one hidden layer, the number of the input, hidden, output neurons is 5, 9, 1. The transfer function between input and hidden layer was 'tansig' for its flexible nonlinearity. The 'logsig' transfer function was introduced between hidden and output layer Because of its perfect output range (0-1) to the newness rate. The initial weights and biases were selected randomly firstly and then set to the most suitable values which were obtained through the training phrase. The details were shown in Table. 3

Table 3. Structure of the neural network

Structure of the neural network	Details
Neural network type	Back – propagation neural network
Number of the layers	3
Number of the neurons	Input layer:5 hidden layer:9 output layer:1
Weights and biases	Selected randomly firstly and set to the suitable value obtained through training phrase
Training function	Levenberg-Marquardt algorithm ('TRAINLM')
Learning rate	0.001
Sum-square error	0.003

C. *Training phrase*

During the training phrase, Levenberg -Marquardt algorithm was introduced as the training function. The Levenberg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix which is complex for computation and time consuming. A scalar μ was introduced in this algorithm, which was variable and decreased after each successful step which the performance function was reduced and increased only when a tentative step would increase the performance function. So this algorithm appears to be the fastest method for training moderate-sized neural networks.

IV. NONLINEAR CURVE FIT

A. *Inputs selection strategy*

After training, the BP neural network extracted the feature of the complex relationship between residual value and the factors. Then a set of schemed data which presents the typical trade record with complicated conditions in the used car market was feed into the trained BP neural network. Its outputs and the schemed data as the inputs were used for nonlinear curve fit in this section. As mentioned before, the residual value reduced following the exponential curve over time, but yet not including the affect of the other strengthened factors. Here the first priority was determining how the time of used car age combined with mileage affects the residual value of the used car. Here five driving conditions were introduced which were critical underused-, underused-, normal used-, overused-, critical overused-condition with the mileage of 14.4k, 19.2k, 24k, 28.8k, 33.6k km each year respectively [7].

B. *Structure of the nonlinear curve fit*

As discussed in the section 1, four strengthened factors were introduced to present a large number of original factors. In application, several preconditions were drawn based on the data we acquired as following: (1) the economic factor was included in the replacement price. (2) The physical factor was affected by the time (age of the used car) and the mileage cohesively with which the feature had been determined by the BP neural network. (3) The wear factor was used to reflect the condition of the specific used car, which it doesn't affect the whole distribution curve of the residual value of the used car.

Based on the above preconditions and the equations in section II, the final function was established as following:

$$P(t) = a * e^{(-f(x_2)*k*t)} * f(x_1)$$

Where $f(x_1)$ and $f(x_2)$ denote the coefficient of the reputation and the wear factors, respectively. $p(t)$, a, k, t were described before.

The nonlinear curve fit was processed by using the 'nlinfit' function in MATLAB Toolbox. The initial weights of k, $f(x_1)$, $f(x_2)$ were set to 0, 0.5, 1 respectively by experimental test and experience. The schemed data and the output from BP neural network were used as the inputs and responses.

V. EXPERIMENTAL RESULT AND CONCLUSION

After the establishment of the model, the result was obtained as the form of matrix at the size of 14 X 3 which presents the coefficients of k, $f(x_1)$, $f(x_2)$ of 14 types of used car. The distribution curves of residual value based on the result from the new established model were shown as following:

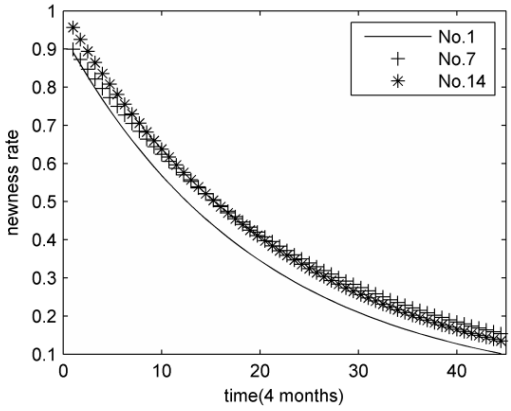

Figure 1. The distribution curves of newness rate of the 'No.1', 'No.7', 'No.14'

The used car types of 'No.1', 'No.7' ,'No.14' represent the 'imported car with its price more than 350k CNY', 'JVC car with its price ranged from 150-250k CNY', 'domestic car with its price less than 80k CNY'. These three types of cars are the most common and popular cars in both the new-, used-car trade market. According to the result shown in the figure.1, the No.1 car' value was reducing fastest with the most sharpest slope, which according to the investigation in the trade market were suffered significantly by the high replacement price and the traditional philosophy to the used car that a new CVT car is more acceptable than a used imported car of the same price. The No.14 car's value was reducing slower than the No.7 car at the first five years (15*4months) due to its low price but reducing faster than the No.7 car which had a better reputation with better quality that may lessen the fee for maintenance .

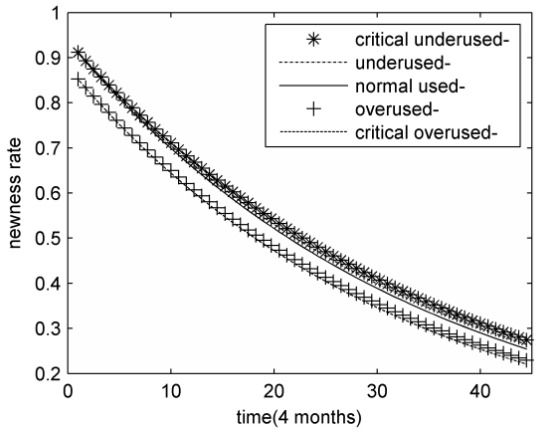

Figure 2. The distribution curves of newness rate of "No.6" with five driving conditions

As shown in figure.2, the No.6 car which presents the 'JVC car with its price ranged from 150k-250k CNY' also enjoyed the popularity. The result suggested that the distribution curves of newness rate with critical underused-, underused condition were close to the one with normal used condition. It indicated that underused condition can't improve the ability of keeping value. It may be interpreted by the phenomenon that the number value of the milometer was modified for a high price, so usually the price of a used car with the condition of underused is equivalent or a bit more to the normal used car. However, the newness rate of overused-, and critical overused car is much lower compared with the car with the condition above-mentioned. Overused condition would worsen the physiochemical condition of the car, which cause the car breaking down earlier.

Based on the results shown above, the trend of the distribution curves of the newness rate of classified used car which were obtained from the new established model matched with the features of the used car trade. To validate the prediction ability of the specific used car, an example of the used car from the [8] which is the most famous website involved in used car was introduced, the detail was shown in Table.4

Table.4 example for validation

Model	A4-2.0T-CVT
Manufacturer	FAW-Volkswagen
Date of Check-in	Sep,2007
Date of appraisal	Jul, 2010
Replacement price (including purchase tax)	370k CNY
Mileage	57k kilometers
Car condition	normal

This car manufactured by FAW-Volkswagen is a JVC car with its price more than 350k CNY, belonging to the car No.5. Calculation was practiced by using the coefficients of k , $f(x_1)$, $f(x_2)$ of the car No.5. The result of the new model was 275.524k CNY, the price of real trade record was 270k CNY, of which the relative error was 2.0049e-006, much less than we required (<5%).

According to the validation of the new established model, we concluded that this new model can predict the residual value of used car with various conditions accurately and conveniently, which can be use in estimation of residual value of the used cars in the second-hand trade markets, companies and individuals.

ACKNOWLEDGMENT

This work was supported by the Program for Professor of Special Appointment (Eastern Scholar), and partly supported by the 085 Knowledge Innovation Engineering (Grant no. JZ0901) at Shanghai Institutions of Higher Learning

REFERENCES

[1] Y. Z. Chen, "Appraisal valuation of used car based on replacement cost approach," Road and vehicle, 2002(4).(in Chinese)

[2] X. R. Feng, "Research on model for used car appraisal valuation," 2008(1). (in Chinese)

[3] G.Y. Zhang, "research on used car appraisal," Commercial Accountant, 2006(6): 37~38 (in Chinese)

[4] X.M. Jin, Working technical economics of vehicle. Beijing: China machine press, 2002. (in Chinese)

[5] Matthias Seeger, "Gaussian Processes for Machine Learning," International Journal of Neural Systems. 2004, 14: 69-106.

[6] Y. S. Wang. Vehicle appraisal. Beijing: China machine press. Mar 2006. (in Chinese)

[7] Z. H. Chen J. X. Chu, "New method for computing the newness rate of the used car based on the equivalent accumulating mileage," Journal of Shanghai University of engineering science. 2003, 17(11) 167-170 (in Chinese)

[8] X.W. Ma, The first auto,[2010,7,30] http://www.iautos.cn/usedcar/usedcar_shop_detail.asp?uploadid=865366