Sensors & Transducers, Vol. 171, Issue 5, May 2014, pp. 162-168

Sensors & Transducers

A Prediction Method of Airport Noise Based on Hybrid Ensemble Learning

[1, 2] Tao XU, [1] Qichuan YANG, [1, 2] Zonglei LV
[1] Civil Aviation University of China, Tianjin, 300300, China
[2] Information Technology Research Base, Civil Aviation Administration of China, 300300, China
[1] E-mail: qichuan171@163.com

Received: 23 April 2014 /Accepted: 27 May 2014 /Published: 31 May 2014

Abstract: Using monitoring history data to build and to train a prediction model for airport noise is a normal method in recent years. However, the single model built in different ways has various performances in the storage, efficiency and accuracy. In order to predict the noise accurately in some complex environment around airport, this paper presents a prediction method based on hybrid ensemble learning. The proposed method ensembles three algorithms: artificial neural network as an active learner, nearest neighbor as a passive leaner and nonlinear regression as a synthesized learner. The experimental results show that the three learners can meet forecast demands respectively in on-line, near-line and off-line. And the accuracy of prediction is improved by integrating these three learners' results. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Airport noise, Hybrid ensemble, Artificial neural network, Nearest neighbor, Nonlinear regression.

1. Introduction

With the rapid development of China's civil aviation, the airplane becomes gradually the main transport and travel way in people's life. The number of airport and flight are growing steadily in China. Only the Beijing Capital International Airport in first quarter of 2013, the volume of total transport passenger reaches 19.95 million, an increase of 4.81 %, and the aircraft movement reaches 135 thousands vehicles, an increase of 0.75 %. However, the air transport brings not only convenience and prosperity, but also environmental issues at the same time. Among them, the most obvious impact on the residents is undoubtedly the aircraft noise pollution which has become the most complained problem.

The production of airport aircraft noise often involves a number of factors: the aircraft type, atmospheric environment, the runways of landing and departing, tracks of flying and position of monitors [1, 2]. Analyzing the factors of airport noise and distribution around the airport, and predicting noise accurately have great significance in controlling the airport noise effectively and optimizing the construction of the airport.

The methods commonly used to predict airport noise can be broadly divided into two categories: one class of model is based on engine and airport environment parameters, which it calculates the value of each interesting noise points [3]. Such as the airport noise prediction model INM developed by U.S. Federal Aviation Administration (FAA); the NOISEMAP prediction model developed by the U.S. Department of Defense and the U.S. National Aeronautics and Space Administration (NASA) for the U.S. military, etc. Another class of model measures the ambient noise of the airport, combines with flight parameters and environment variables, then uses machine learning or data mining methods to build prediction models. Such as the neural network prediction model developed by Information Technology Research Base of Civil Aviation Administration of China, etc. However, the existing prediction models have different emphases on the storage space, the execution time and the prediction accuracy. For instance, the NOISEMAP has a longer running interval; the neural network model needs a lot of training time; the clustering learning demands a mass of data.

In order to solve these problems, this paper introduces a prediction method that integrates three state-of-the-art algorithms in hybrid ensemble way. It involves the BP neural network model as an active learner to meet online forecast demands and selects Nearest Neighbor as a passive learner to complete offline prediction tasks. Meanwhile, nonlinear regression is applied as a half positive and half negative learner to construct a near-line prediction model. In addition, these three methods can also be integrated together to get a better forecast accuracy.

2. Related Methodologies

2.1. Active Learning and Passive Learning

In accordance with the periods in which models are constructed, the supervised machine learning methods can be divided into passive learning and active learning.

When training data set are given, the active learning will process the data and construct the prediction model before receiving the test instance. Such as decision trees, neural networks, Bayesian network and so on [4, 5]. Active learning consumes a lot of time to establish prediction models in the training phase. When the instance needs to be predicted, the parameters are entered simply in the model. After a few calculations, prediction results can be obtained. The prediction time is small. But the generalization ability of active learning will vary with the training samples [6, 7]. In most cases, the result is poor when the properties of training sets are quite different form them of the test sets.

Corresponding to the active learning, the passive learning just simply stores training data without constructing a model. And the prediction work does not begin until a test tuple is given. Such as nearest neighbors, locally weighted regression and case-based reasoning. Passive learning only do a small amount of work with training data, and the prediction procedures fill with much calculating work. Passive learning does not provide an explanation or an insight into the structure of the training data. However, passive learning naturally supports incremental learning and is capable of modeling ultra-complex decision spatial of polygonal. They do not estimate the objective function in the entire instance space, but make partial and disparate estimates for each instance that needs to be classified. These characteristics are not easily seen in other learning algorithm.

2.2. Back Propagation Artificial Neural Network

Back propagation artificial neural network (BPANN) is currently one of the most widely used models. Fausett et al [8] introduces many useful application samples in his writing. Xie et al [9] propose several methods to improve the performance of network. Many new algometric is also proposed, such as Xiaoyuan’s articles [10]. Generally, BPANN contains an input layer, a hidden layer and an output layer. The network constantly adjusts the weights through error reverse propagation, the network output, therefore, would be closer to the true value. The structure of BPANN is just shown in Fig. 1.

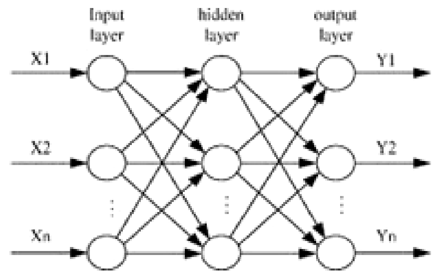

Fig. 1. Structure of BPANN.

The main training steps of BPANN:

Step 1: Processing the input propagation: the input layers accept the data and the output layers get the result through the intermediate layer by propagation calculations;

Step2: Processing the error back propagation: the output errors from the output layer pass through the middle layer to the input layer;

Step3: The input propagation model and the error back propagation process are repeated alternately;

Step4: Determining whether the global error of network tends to be a minimum.

For a BP network with a single hidden layer and d inputs, its output can be expressed as:

$$N(x) = \sum_{i=1}^{m} c_i \sigma\left(\sum_{j=1}^{d} w_{ij} x_j + \theta_i\right), x \in R^d, d \geq 1, \quad (1)$$

where $1 \leq i \leq m, \theta_i \in R$ is the threshold, $w_i = \left(w_{i1}, w_{i2}, \ldots, w_{idj}\right)^T \in R^d$ is the i-th neuron connection weights between the input layer and the hidden layer. The c_i is the connection weights between hidden layer and output layer. The σ is the node's activation function of hidden layer. Typically, the activation function σ is sigmoid function which satisfies $\sigma(t) \to 1(t \to \infty), \sigma(t) \to 0(t \to -\infty)$.

As BPANN has good approximation capability with the data structure and the nonlinear relationship among attributes, it is considered as an active learning algorithm.

2.3 Regression Analysis

Regression analysis is a mathematical and statistical analysis model that can be used to build model between one or more predictor variables and a response variable. Lefever et al [11] use it to predict strength in his research. The work of Chen et al [12] reaches an excellent fit in prediction of atmospheric environment.

In actual science researches, scholars often encounter in the nonlinear regression problems which cannot be processed in linear way. To this end, stepwise regression using the least squares method is put forward to establish equation based on multiple regression in the process. This method calculates partial regression sum of squares for each independent variable. Then it selects variables which have the minimum partial regression sum of squares to do statistical tests. At the same time, the independent variables with significant effect are introduced gradually to the equation. When the equation does not introduce any new variables, the stepwise regression procedure is end.

In general, for the dependent variable Y and the independent variables $x_1, x_2, x_3, \ldots, x_m$, the multivariate regression can be expressed as

$$Y = a_0 + a_1 x_1^{b1} + a_2 x_2^{b2} + \cdots a_m x_m^{bm} + \varepsilon, \quad (2)$$

where $a_0, a_1, \ldots, a_m$ are the undetermined coefficients of the equation; $b_0, b_1, \ldots, b_m$ are the factor coefficient for the non-linear relationship, and ε is random error in normally distributed.

2.4. Nearest Neighbor

Nearest neighbors (NN) algorithm is an analogy-based learning method, namely, to learn by comparing a given test tuple with similar training tuple. Since the algorithm describes training tuples with N dimensional vectors, each tuple represents a point in N-dimensional space. All training tuples are stored in the N-dimensional space. When a tuple position is given, the nearest neighbor algorithm searches for the space, finds the closest unknown tuple in those training tuples. This process is described as follows:

Input: Training tuple set D, Test tuple set Z
Output: Prediction set N of Z
For each $z_i \in Z$
For each $d_j \in D$
Compute distance (z_i, d_j)
If distance$(z_i, d_j) < Max$
Set z_i in neighbor for d_j
End

$$n_i = \frac{1}{length(neighbor)} \sum_{j=1}^{length(neighbor)} d_j$$

Set n_i to z_i
End

The search speed is determined by calculating the distance between tuples. In order to improve the speed, the technologies of increment is proposed by Wu et al [13] to calculate the distance between the geometric data; Tao et al [14] use the space filling curve technical to process K-nearest neighbor queries; Zhang et al [15] propose a similar query technology based on learning to construct similarity matrix of database and to train classifier and calculate the distance.

Euclidean vector is used here to measure the distance between tuples, which is calculated as follows:

$$\text{distance}(z, d) = sqrt\left(\sum_{i=1}^{length(z)} \left(z_i - d_i\right)^2\right), \quad (3)$$

where d represents the training set and z represents the test set. Length (z) indicates that number of elements in the data set is z.

3. The Hybrid Ensemble Method

To meet the requirements of different kinds of predictions, the airport noise prediction method based on hybrid ensemble consists of three parts: P_{BP} (BP neural network), P_{NR} (nonlinear regression) and P_{NN} (nearest neighbor). Linear combination is used to integrate these three predictions model. Namely:

$$P_{final} = w_1 P_{BP} + w_2 P_{NR} + w_3 P_{KNN}, \quad (4)$$

In formula (4), $w_1, w_2, w_3 \in [0,1]$ means the weights of models. Different weights correspond to the various prediction properties, shown in Table 1.

Table 1. Function of weights.

(w_1, w_2, w_3)	On-Line	Near-Line	Off-Line	Hybrid
(1,0,0)	√	×	×	×
(0,1,0)	×	√	×	×
(0,0,1)	×	×	√	×
$w_{1,2,3} \in (0,1)$	×	×	×	√

When one of the w_i is set to 1, the corresponding model are considered to predict alone. And the hybrid ensemble prediction is active when it is between 0 and 1.

These three prediction models are trained respectively according to the airport historical noise data and the parameters of them are constructed through the training procedure.

3.1. Description of Models

3.1.1. BPANN

The BPANN is given as an online prediction model for its active learning properties. In order to establish a BPANN prediction model for a specific airport, the history noise monitoring data of the airport is used to train the network. During the training process, the structure of the network is altered several times to get a better fitting performance.

BP neural network with at least one S-type hidden layer and a linear output layer can approximate any rational function. Increasing the number of neurons or layers in the network can increase its accuracy. In specific airport, the aircraft noise propagation model is a continuous nonlinear rational function. BP neural network with three layers are used in the propagation fitting.

The transfer function is a continuously differentiable non-decreasing function. The tangent S-type transfer function, therefore, is used in the input and hidden layer. A linear transfer function, therefore, is used in the output layer. Neuron number of hidden layer currently lacks theoretical guidance. Generally, it is determined through experience and experiments. Related experiments show that when the input layer and hidden layer neurons are 6 and 15, the network can quickly achieves the desired accuracy. The number of neurons in the output layer is determined by the specific issues. Since the noise value is the only output value of the network, the output layer neuron number is 1. In order to speed up network convergence and to avoid falling into local minimum, additional momentum and adaptive learning rate are applied in the network.

3.1.2. Nonlinear Regression

The nonlinear regression model is involved as a near-line prediction model. The model combines both active learning and passive learning process.

Active learning part works before receiving the test set. Based on historical monitoring data, the binary attributes are divided into a plurality of subsets. Multiple linear regression models are built based on each subset.

When a test set is given, binary parameters are used to classify the set and to choose the corresponding constructed model. And non-binary parameters are input in the model to obtain the prediction values.

3.1.3. Nearest Neighbor

The nearest neighbor is introduced as offline prediction models. Although there is no training process before forecasting the test sets. Some preprocessing is executed in order to accelerate the searching process.

The parameters of historical data are ranged according to the size of its value field. Then a list of parameters is given for the next search process. The binary parameters are located at front part of the list apparently. After obtaining the test parameters, the binary parameters are calculated at first in order to accelerate the process. These continuous parameters are calculated by Equation 3 after the binary parameters.

3.2. Evaluation Criteria of Airport Noise

There are many noise evaluation criteria currently. Such as the maximum A sound level (LAmax), equivalent continuous A sound level (LEQ), sound exposure level (SEL), perceived noise level (PNL), effective perceived noise level (EPNL) and weighted equivalent continuous perceived noise level (WECPNL) and so on.

Among them, the LAmax, EPNL and SEL reflect the evaluation criteria of a single noise event. LAmax can reflect subjective evaluation of human better and can be measured easily. But it lacks the duration expression of the noise. EPNL is able to measure the intensity of a single noise event. The disadvantage is that its calculation and measurements are complex. SEL is the A sound level converted to one second of full noise in a flight event. It takes into account both the duration and intensity of sound. SEL expression is:

$$L_{AE} = 10\lg\left[\frac{1}{T_0}\int_{t_1}^{t_2}\frac{P_A^2(t)}{P_0^2}dt\right], \tag{5}$$

or

$$L_{AE} = 10\lg\left(\frac{1}{T_0}\int_{t_1}^{t_2}10^{L_A(t)/10}\right), \tag{6}$$

where T_0 is the standard time 1 second; t_1 and t_2 stand for the start time and end time of event respectively; PA (t) is for the A-weighted instantaneous sound pressure at time t; P_0 is the reference sound pressure; LA(t) is for the instantaneous A sound level during the sound interval. As SEL is normalized to one second intervals, it is easy to compare the energy value in different noise events. This article, therefore, uses SEL as the noise evaluation criteria of a single flight event.

4. Experiments and Results

4.1. Experimental Data and Data Preprocessing

Monitoring noise data sets of 2011 in a large hub airport of China are used in the experiment. The data sets consist of 23 monitoring points around the airport. Four types of aircraft B738, A321, B733, and A320 are selected in the experiment since they have the largest number of flying event in the airports. Statistical data are shown in Table 2.

Table 2. Statistical Table of Experiment Data

No	Aircraft	Total Number	Training Set	Test Set
1	B738	7584	5688	1896
2	A321	2608	1956	652
3	B733	1700	1275	425
4	A332	1668	1251	417

All the flying event records of each model during a month consist the training data. The format of data is shown in Table 3.

Table 3. Record format of monitoring data

Flight	Operation	Aircraft Type	Runway	Spatial-Distance (m)	Horizontal-Distance (m)	Height (m)	Wind-Velocity (m/s)
CCA975	DEP	B738	18L	2148	2060	610	0.29
CCA975	DEP	B738	18L	1070	885	602	0.29
CCA1136	ARR	B733	18L	1131	1057	403	0.29
…	…	…	…	…			

Meanwhile, the standardization is used in this paper to prevent that the attributes with larger initial range have larger weights than other attributes and to prevent the high computational complexity caused by different accuracy. The deviation standardized processing is used in process the raw data:

$$Di=(di-MIN(D))/(MAX(D)-MIN(D)) \quad (7)$$

The output results are recovered by against standardization:

$$di=Di*(MAX(D)-MIN(D))+ MIN(D) \quad (8)$$

The MIN (D) and MAX (D) in Equation (7) and (8) denote the minimum and maximum values sample respectively in data set D.

4.2. Experiment Description

The experiment consists of two parts: the first part is to test performance of three prediction models in airport noise environment; the second part is the test of hybrid ensemble that integrates there models together.

In the regression model experiment, the data residuals are calculated at first, then the data points with larger residuals from the original data are removed. Part of results is shown in Fig. 2, the red points' residual is too large and relatively unusual. These data points are excluded.

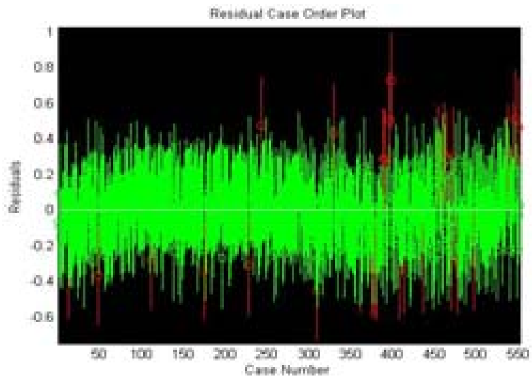

Fig. 2. Data residuals statistics.

Stepwise regression is constructed after the above cleaning process. The parameters of stepwise in each step are shown in Table 4. After observing the constant value (Intercept), the residual standard deviation (RMSE), the multiple correlation coefficient of fitting R2, the corrected multiple correlation coefficient Adj-R2, F value and the corresponding probability P, optimal equation form can be determined as y= f(b, x1, x4, x6, x2, x3).

Data cleansing and stepwise processes are repeated until there are no outstanding data points. The regression equation expression is constructed at the end of the process. And each type of flight will have its own expression.

The mean absolute forecast error and average time of regression prediction model is shown in Table 5 compared with other two models.

Table 4. Parameters of regression.

Step	Introduced Variables	Parameters					
		Intercept	RMSE	R2	Adj-R2	P	F
1	f(b, x1)	0.856545	0.147619	0.570902	0.570122	4.03459e-103	731.758
2	f(b, x1,x4)	0.846088	0.146113	0.580372	0.578843	2.99407e-104	379.651
3	f(b, x1, x4, x6)	0.828471	0.144923	0.587934	0.585678	4.55126e-105	260.628
4	f(b, x1, x4, x6, x2)	0.726348	0.142669	0.601383	0.598468	9.34866e-108	206.311
5	f(b, x1, x4, x6, x2, x3)	0.651367	0.142218	0.604622	0.601001	1.55569e-107	166.991

Table 5. Errors and time.

Aircraft	BPANN		Regression		Nearest Neighbor		Hybrid Ensemble	
	Error(dB)	Time(s)	Error(dB)	Time(s)	Error(dB)	Time(s)	Error(dB)	Time(s)
B737	3.7	3.1	2.8	6.5	2.9	18	2.7	23
A321	3.1	3.0	3.5	5.5	1.77	14	1.76	21
B733	2.9	2.9	3.0	8.7	2.2	16	2.1	25
A332	2.1	3.0	2.4	8.1	4.1	15	1.9	23

As for the BP neural network model in the experiment, model parameters are set in accordance with the section 3.1. Here the iterative training process of type B738 is shown in Fig. 3 as an example. The network achieves a relatively stable state in the 4th iteration. The training process of other three models is similar.

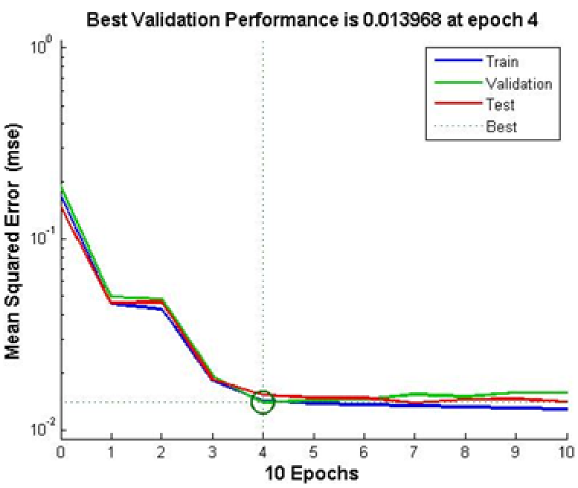

Fig. 3. BPANN training process of B738.

After training the BPANN model, the test data sets are input to get the prediction values. Part of the comparison between the true value and prediction value is shown in Fig. 4. Meanwhile each forecast time is recorded with mean absolute error in Table 5.

There is no training process for the nearest neighbor as a passive learning model. In the experiment, the training sets of the four types of aircraft are treated as a comparative database to find the nearest tuples. The results are recorded and mean absolute errors are calculated with time, shown in Table 5.

Finally, according to the Equation (4), three prediction models are heterogeneously integrated by weights. The weights are calculated in accordance with the mean absolute error of each model given in Table 5. The greater the error is, the smaller its weight will be. The results are also recorded in Table 5.

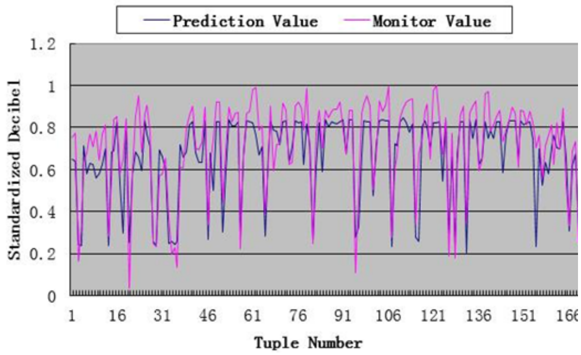

Fig. 4. Comparison of true and prediction value by BPANN.

4.3 Analysis of Experimental Results

From the experimental results of Table 5, it can be seen that the three prediction models have significant difference in time performance. The BPANN is the fastest model and the Nearest Neighbor is the slowest one. As the requirement for the online prediction is generally defined within 5 seconds, BPANN is competent in this job. The interaction operation with the database in Nearest Neighbor is involved in its prediction process. The finding and comparing process in the database consumes a lot of time.

Thus it is arranged as the off line prediction model. Meanwhile, the regression has a moderate performance.

Secondly, the prediction accuracy of the three models are dissimilar in different data sets. After analyzing the experiment results and training data together, some conclusions can be made. When the training data and test sets contain similar data items, nearest neighbor model gives more accurate predictions than the other two models just like the experiment of A321 shown in Table 5. When training data is lack of similar flight events, BPANN can get more accurate predictions due to its outstanding generalization ability than others just as the A332 in Table 5. In addition, the results of regression prediction model can reach highly accurate when the binary parameters are consistent, which, for example, can be seen in the result of B737 in regressing since it has good fit character in continuous data items.

When prediction models are heterogeneously integrated, the prediction time are extended. But accuracy is significantly improved by this ensemble method. The proposed model combines the advantages of these prediction models and can adapt to different training data.

5. Conclusions

Based on the idea of hybrid ensemble, this paper proposes an integrated method that contains BPANN, nearest neighbor and non-linear regression. The three prediction models in proposed method can be applied respectively for online, near-line, off-line prediction requirements. Meanwhile, when the three prediction models are endowed with weights, the generalization and prediction accuracy have been significantly improved contrasted with any of them.

Acknowledgments

This work is partially supported by the key project of National Natural Science Foundation of China (No. 61139002), the National High-tech R&D Program of China (863 Program) (No. 20012AA063301), the Science and Technology Projects of Civil Aviation Administration of China (No. MHRD201006, MHRD201101), the Fundamental Research Funds for the Central Universities (No. 3122013P013).

References

[1]. W. Schuster, M. Porretta, High-performance trajectory prediction for civil aircraft, in *Proceedings of the 29th International Conference on Digital Avionics Systems*, 2010, pp. 1-9.

[2]. B. Jandl, Prediction of noise exposure levels using simulated flight trajectories, in *Proceedings of the 30th International Conference on Digital Avionics Systems*, 2011, pp. 1-13.

[3]. L. P. Sanchez, Aircraft classification and acoustic impact estimation based on real-time take-off noise measurements, *Neural Processing Letters,* Vol. 38, Issue 2, 2012, pp. 239-259.

[4]. T. N. Le, A semi-supervised ensemble learning method for finding discriminative motifs and its application. *Journal of Universal Computer Science*, Vol. 19, Issue 4, 2013, pp. 563-580.

[5]. I. King, H. Yang, Ensemble learning for imbalanced e-commerce transaction anomaly classification, *Neural Information Processing Springer*, Vol. 58, No. 63, 2009, pp. 866-874.

[6]. L. Rokach, Ensemble-based classifiers, *Artificial Intelligence Review*, Vol. 33, Issue 2, pp. 1-39, 2010.

[7]. E. Lughofer, P. Kazienko, Hybrid and ensemble methods in machine learning, *Journal of Universal Computer Science*, Vol. 19, Issue 4, 2013, pp. 457-461.

[8]. L. Fausett, Fundamentals of neural networks: architectures, algorithms, and applications, *Pearson Education*, India, 2006, pp. 23-25.

[9]. R. Xie, X. Wang, Y. Li, et al, Research and application on improved BP neural network algorithm, in *Proceedings of the 5th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2010, pp. 1462-1466.

[10]. X. Y., Li, B. Qi, L. Wu, A new improved BP neural network algorithm, in *Proceedings of the 2nd International Conference on Intelligent Computation Technology and Automation*, 2009, pp. 19-22.

[11]. C. Lefever, J. H. Anning, B. T. Jensen, Regression models for strength assessment method selection and performance predictions, in *Proceedings of the International Journal of Exercise Science*, 2013, pp. 43-45.

[12]. Y. Chen, R. Shi, S. Shu, et al, Ensemble and enhanced PM10 concentration forecast model based on stepwise regression and wavelet analysis, *Atmospheric Environment*, Vol. 74, 2013, pp. 346-359.

[13]. Z. Wu, Y. Chen, Predicting protein subcellular localization using the algorithm of increment of diversity combined with weighted k-nearest neighbor, in *Proceedings of the 2nd International Conference on System Engineering and Modeling*, 2013, pp. 590-594.

[14]. Y. Tao, K. Yi, Quality and efficiency in high dimensional nearest neighbor search, in *Proceedings of the ACM SIGMOD Conference on Management of Data*, 2009, pp. 563-576.

[15]. D. Zhang, J. Wang, Self-taught hashing for fast similarity search, in *Proceedings of the 33rd ACM Conference on Research and Development in Information Retrieval*, 2010, pp. 18-25.