

PAPER • OPEN ACCESS

## Deep Belief Network based Coal Mine Methane Sensor Data Classification

To cite this article: Xuefeng Wu *et al* 2019 *J. Phys.: Conf. Ser.* **1302** 032013

View the [article online](#) for updates and enhancements.

### You may also like

- [Effect of Ni on the kinetics of intermetallic compounds evolution on the Sn-0.7Cu-10Bi-xNi/Co interface during various reflow](#)  
He Gao, Fuxiang Wei, Caixia Lin et al.
- [Research on dynamic characteristics of robot joint system considering gear crack failure and bearing local failure](#)  
Jiaxu Zhao, Yanbin Lu, Guo Ye et al.
- [Floristic Composition and Characteristics of Xuzhou, China](#)  
Shuming Ju, Peijiao Bu, Yuting Zhang et al.



The banner features a large white circle on the left containing the number '250' in red, blue, and green, with a blue ribbon below it that says 'ECS MEETING CELEBRATION'. To the right of the circle, the ECS logo is displayed above the text 'The Electrochemical Society' and 'Advancing solid state & electrochemical science & technology'. The background is a collage of confetti and people's hands.

**250th ECS Meeting**  
**October 25–29, 2026**  
**Calgary, Canada**  
*BMO Center*

**Step into the  
Spotlight**

**SUBMIT YOUR  
ABSTRACT**

**Submission deadline:  
March 27, 2026**

# Deep Belief Network based Coal Mine Methane Sensor Data Classification

Xuefeng Wu<sup>1</sup>, Zhikai Zhao<sup>2,3</sup> and Li Wang<sup>4</sup>

<sup>1</sup> Xuhai College, China University of Ming and Technology, No. 1 Jinshan East Road, Xuzhou City, Jiangsu Province, China.

<sup>2</sup> IoT Perception Mine Research Center, China University of Ming and Technology, No. 1 Jinshan East Road, Xuzhou City, Jiangsu Province, China.

<sup>3</sup> School of Information and Control Engineering, China University of Ming and Technology, No. 1 Daxue Road, Xuzhou City, Jiangsu Province, China.

<sup>4</sup> School of Mechanics and Civil Engineering, China University of Ming and Technology, No. 1 Daxue Road, Xuzhou City, Jiangsu Province, China.  
Email: cumtaust@126.com

**Abstract.** Gas explosion is the main hazard which affects the safety of coal mine production. One way to solve the problem is to predict the dangerous levels of methane concentration using the sensor time series from hazard monitoring systems. In this paper we proposed our method based on DBN to classify the dangerous levels of methane concentration. A multilayer RBM network is built to reconstruct the methane sensor data and depth characteristics of methane disaster classification are extracted. Then a BP network is used for classification learning. We test our method on a real coal mine methane sensor data, and contrast with two classical methods, SVM and KNN. The experiments results showed that our proposed deep feature learning methods could achieve better performance than the two classical shallow methods.

## 1. Introduction

Coal mine gas accident is a major accident, according to the China National Coal Supervision Bureau data, in 2016, a total of 13 coal mine gas accidents occurred in China, resulting in the death of 170 people. How to effectively prevent gas disaster has been a subject worthy of further study. In recent years, with the development of the coal mine information, we can easily get associated with coal mine gas disaster sensor data, the sensor data provides the basis for the analysis and prediction of coal mine gas disaster, dig out the guidance mode and the gas disaster prevention from these data is a good idea[1-5].

At present, the research on the prediction of coal mine gas disaster is mainly focused on the prediction of gas emission and the evaluation and prediction of coal and gas outburst risk. In this paper, we propose a method of time series prediction based on wavelet radial basis function, taking[6] as the research object. Cheng Jian et al [7,8] analyzes the factors of coal and gas outburst mechanism and the influence of the establishment of Fisher (Fisher) discriminant analysis model, confirming the chaotic characteristics of gas concentration signal, and the associated space in the gas concentration on the prediction model of least squares support vector machine, the coal mine gas concentration in the medium and long term forecast. Yin Hongsheng [9] based on the KPCA/KICA multivariate time series reduction and feature extraction theory, multi-dimensional time series of dimensionality reduction and feature extraction, the use of LS-SVM algorithm, the classification of gas time series. In this paper, the characteristics of gas disaster information are analyzed by using independent



component analysis ([10]) method, and the feature extraction model is established based on maximum entropy and support vector machine. Kozielski et al. [11] proposed a rule based regression method to predict the gas concentration. There are many applications of in-depth learning model[12-14].

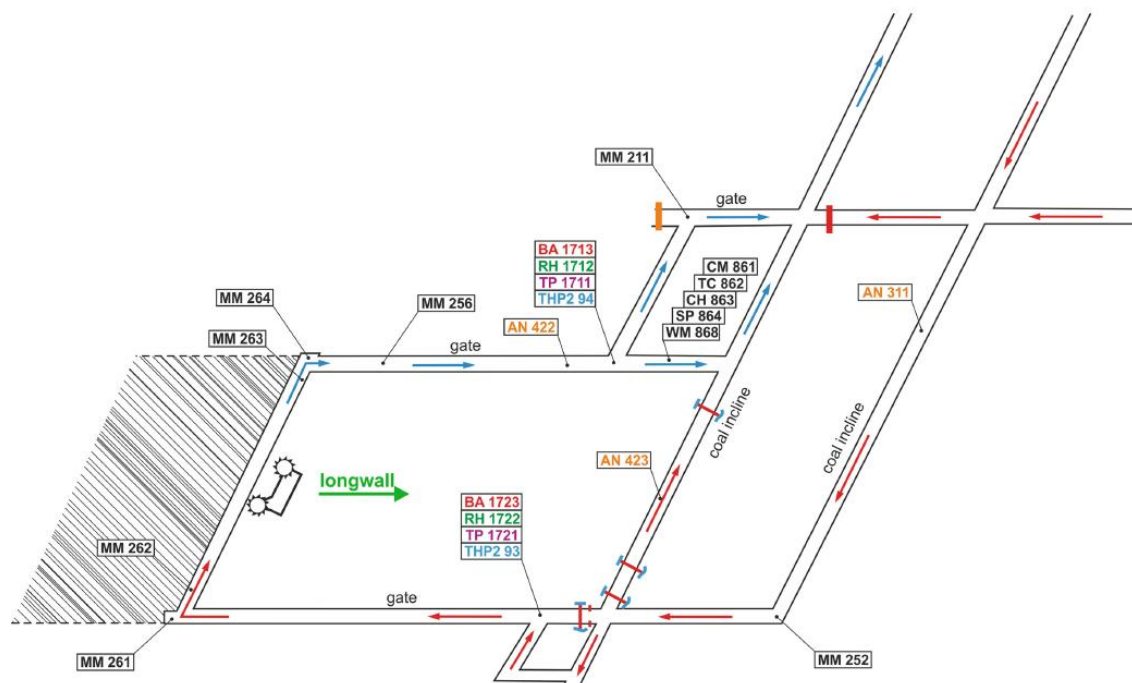
Based on the depth characteristics of coal mine gas data as a starting point, to predict the coal mine gas disaster as the goal, to deep learning technology, coal mine gas data extraction algorithm based on depth characteristics of deep learning, construction of coal mine gas disaster prediction model based on deep learning and application research, combined with the real data of mine, provide a warning new ideas and new model for coal mine gas disaster prediction.

## 2. Materials and Methods

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation as well as the experimental conclusions that can be drawn.

### 2.1. Methane Sensor Data

In this paper, the data given by [15] (IJCRS 15 Data) in coal mine data mining is used as the experimental data. This data set comes from a real mine in Poland, the data collected from March 2, 2014 to June 16, 2014. The number of samples in training set is 51700, and the number of samples is 5076. All samples were divided into two categories, namely "warning" and "normal". The class label is based on the MM263, MM264, and MM256 data of the three methane sensors. Division of the standard refers to the training sample period of 3 to 6 minutes after the alarm threshold is exceeded, the label is more than the sample is marked "warning", otherwise marked as "normal". As shown in Figure 1, the layout of the space of the mining 28 sensors, including gas, wind speed, temperature and humidity, collected once every 1 second to 10 minutes will be calculated, obtained 16800 data to model the 10 minutes of the state coal mine safety with this data, the need for the operation of the 16800 dimensional vector space. In this paper, the data of three methane sensors are selected.



**Figure 1.** Methane Sensor.

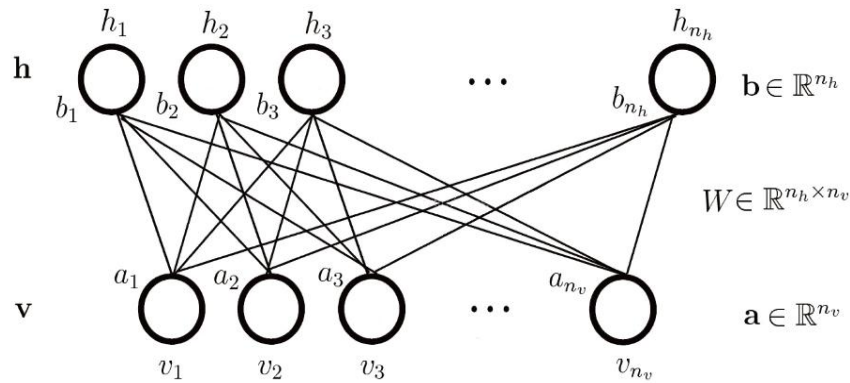
### 2.2. Data Reconstruction of Methane Sensor Data based on RBM

This paper uses RBM based network to reconstruct the methane sensor data, and an encoder of the data is achieved. Then we construct a multi-layer RBM network to achieve multi-layer representation of methane sensor data. Suppose a sample of methane sensor data is  $x_i = (x_1, x_2, \dots, x_m)$ , the label

$y_i = 1$  represents a "warning" sample, while the label  $y_i = 0$  represents a "normal" sample. One RBM contains two layers, a visible layer and a hidden layer. The connections between neurons have the following characteristics: there is no connections in inner layers, and there are all connections between layers. So RBM is a bipartite graph. The architecture of RBM network is showed in figure 2.  $x_i = (x_1, x_2, \dots, x_m)$  as a input,  $n_v = m$ , the input vector is each sample vector,  $v_i = x_i$ , the output vector  $h_i$  represent a encoder of sample  $x_i$ . The methane sensor data is used as visible unit of RBM, and the hidden layer unit characteristic detector for coded methane sensor data. There is an energy function between the visible and hidden layers:

$$E(\mathbf{v}, \mathbf{h}) = - \sum_{i \in \text{methanes}} b_i v_i - \sum_{j \in \text{features}} b_j h_j - \sum_{i,j} v_i h_j w_{ij} \quad (1)$$

- $n_v, n_h$  respectively represent the number of neurons in the visible and hidden layers.
- $\mathbf{v} = (v_1, v_2, \dots, v_n)^T$  is state vector of visible layer,  $v_i$  is the state of the  $i$ th neuron of visible layer.
- $\mathbf{h} = (h_1, h_2, \dots, h_{n_h})^T$  is state vector of hidden layer,  $h_j$  is the state of the  $j$ th neuron of hidden layer.
- $\mathbf{a} = (a_1, a_2, \dots, a_{n_v})^T$  is bias vector of visible layer,  $a_i$  is the bias of the  $i$ th neuron of visible layer.
- $\mathbf{b} = (b_1, b_2, \dots, b_{n_h})^T$  is bias vector of hidden layer,  $b_j$  is the bias of the  $j$ th neuron of hidden layer.
- $\mathbf{W} = (w_{ij} \in \mathbb{R}^n)$  is matrix of weight between hidden and visible layers,  $w_{ij}$  is connection weight between the  $i$ th neuron of hidden layer and the  $j$ th neuron of visible layer.



**Figure 2.** RBM.

For a given methane sensor data, the binary state  $h_j$  is set 1 according to  $\sigma(b_j + \sum_i v_i w_{ij})$ ,  $\sigma(x)$  is logistic function  $(1 + \exp(-x))^{-1}$ ,  $b_j$  is bias of  $j$ ,  $v_i$  is state of  $i$ th methane sensor data,  $w_{ij}$  is weight between  $i$  and  $j$ . Once hidden layer unit choose a set of binary states, each  $v_i$  is set 1 according to  $\sigma(b_i + \sum_j v_j w_{ij})$ ,  $b_i$  is bias of  $i$ , hidden layer unit will be updated again according to the reconstruction error.

For a sample  $x$ , a Contrastive Divergence algorithm is used to train:

- Assign  $x$  to  $v_1$ , then compute the activation probability  $P(h_1|v_1)$  of each neuron in hidden layer.
- Taking a sample from the calculated probability distribution by Gibbs sampling:

$$h_1 \sim P(h_1|v_1) \quad (2)$$

- Reconstruct visible layer using  $h_1$ , and inverse calculation visible layer by hidden layer, then compute the activation probability  $P(v_2|h_1)$  of each neuron in visible layer.

- In the same way, taking a sample from the calculated probability distribution by Gibbs sampling:

$$v_2 \sim P(v_2 | h_1) \quad (3)$$

- Using  $v_2$  to compute the activation probability of each neuron in hidden layer, and get the probability distribution  $P(h_2 | v_2)$ .
- update weight

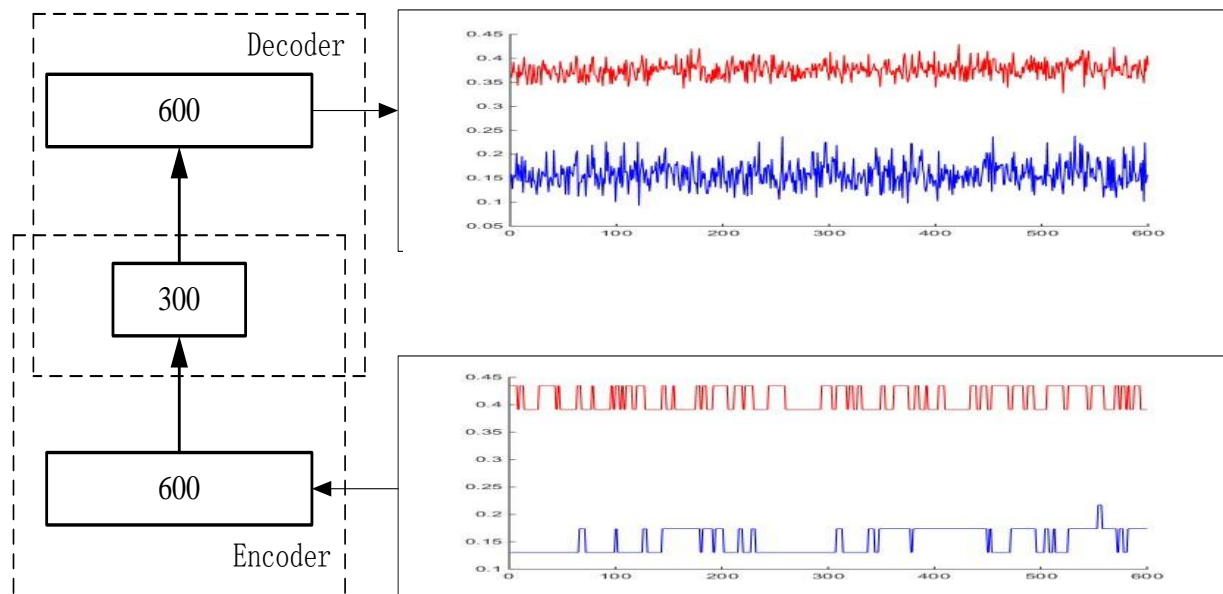
$$W \leftarrow W + \lambda(p(h_1 | v_1)v_1 - P((h_2 | v_2)v_2)) \quad (4)$$

$$b \leftarrow b + \lambda(v_1 - v_2) \quad (5)$$

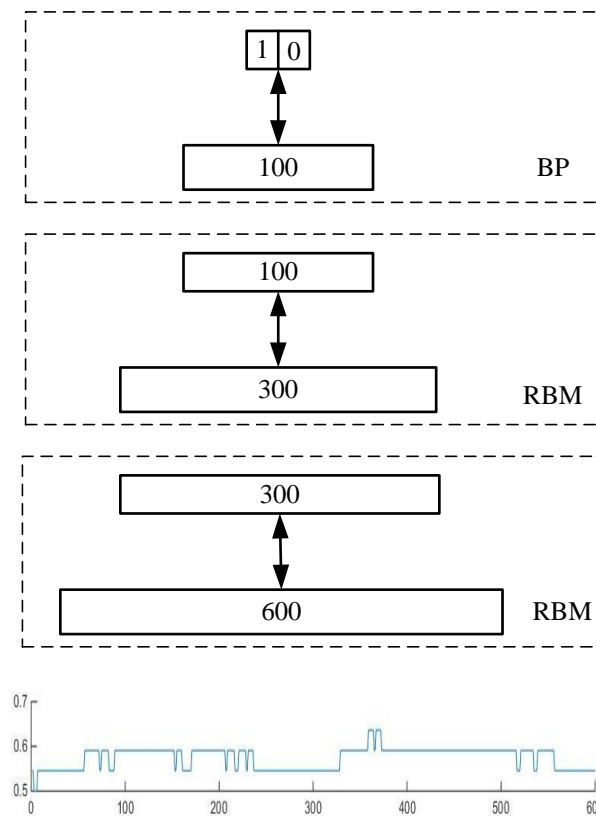
$$c \leftarrow c + \lambda(h_1 - h_2) \quad (6)$$

### 2.3. Methane Sensor Data Classification

The effective training process of RBM makes it suitable as a component of DBN. The hidden units of each layer of RBM express the characteristics of a higher degree of input data through learning. DBN can be considered as a neural network with trained initial weights. When not using the category label and back propagation in DBN structure, DBN can be used for dimensionality reduction. DBN can be used to classify categories when they are associated with a feature vector. On the basis of multi tier RBM, add a final layer representing the desired output variable. The output of the last layer RBM network is used as the input of BP neural network. The supervised BP neural network is used to propagate the error and adjust the whole model from top to bottom. The RBM neural network is optimized as the input of BP neural network, which solves the problem that BP neural network is easy to fall into local minimum and slow convergence speed.



**Figure 3.** Feature encoder.



**Figure 4.** DBN.

In the first part, the feature extraction algorithm based on multi level RBM can get the input data, and the DBN has better classification effect. Therefore, this paper is based on the DBN algorithm to identify the classification of coal mine gas sensor data:

- Select the appropriate training set and test set.
- Data preprocessing.
- Unsupervised training of each layer separately RBM.
- The training set is input into DBN, and then the feature vector is obtained, and then the parameters of the DBN are adjusted to achieve a satisfactory feature vector.
- The training set, the feature set of DBN, classification, test the classification results, modify the parameters of DBN. Finally, the classification results are analyzed.

### 3. Experiments and Result

#### 3.1. Data Preprocessing

First, the training data set of abnormal data processing, using the mean instead. Next, in order to avoid the impact of the dimension, each sensor data need to be normalized. The linear function converts the original data linearization method to the range of [0-1]:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (7)$$

This method can realize the scaling of the original data,  $x_{norm}$  is normalized data,  $x$  is original data,  $x_{max}$  and  $x_{min}$  is respectively the maximum and minimum values of the original data set.

One problem is that the number of positive and negative samples in the training set and the test set is large, and table 1 shows the percentage of positive samples that represent the "warning".



**Table 1.** Positive sample proportion for "warning".

Data\sensor ID	MM263	MM264	MM256
Training	0.009	0.026	0.025
Testing	0.007	0.025	0.027
Upsample Training	0.146	0.161	0.160

By repeating the positive samples, the total number of samples is 60000, and the proportion of positive samples is shown in the table

### 3.2. Classification Results

The quality criterion chosen for the evaluation of submissions was based on a well known Area Under ROC Curve(AUC) measure. This measure was selected due to the sparsity of positive examples from the "warning" decision class for each of the considered target sensors. Table 1 shows proportions of the cases from the "warning" class for the methane meters MM263, MM264 and MM256 in the training and test parts of the data.

In order to compare the advantages and disadvantages between the deep learning algorithm and the common shallow learning algorithm, DBN algorithm proposed in this paper, KNN algorithm and SVM algorithm are adopted in the experiment. Table 2-4 respectively show the classification results of data sets from MM263, MM264 and MM256. Figure 5 shows the ROC curves of three methods classification result on MM263, MM264 and MM256. As shown in Table 2-4, KNN obtains slightly better result than DBN just in AUC and Recall on MM264, but DBN obtains the optimal results in Accuracy, Specificity, Precision and Fmeasure, and it performs best in almost all indicators for the other two data sets. The deep network learning model can extract the essential features closer to the classification target, which is better than the method of manually constructing features. This method can achieve better results for the classification tasks.

**Table 2.** Classification results of data sets from the methane meters MM263

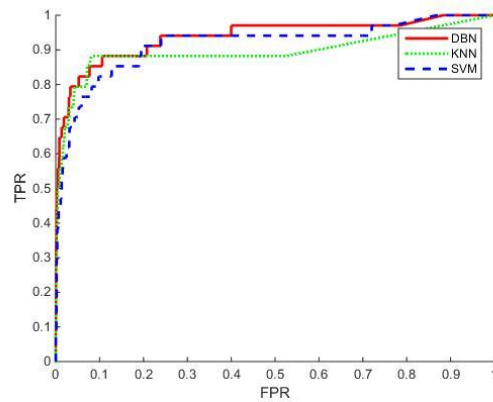
Alg	AUC	Accuracy	Specificity	Precision	Recall	Fmeasure
KNN	<b>0.956</b>	0.972	0.980	0.445	<b>0.653</b>	0.529
SVM	0.941	0.919	0.921	0.212	0.855	0.340
DBN	0.951	<b>0.976</b>	<b>0.982</b>	<b>0.500</b>	0.702	<b>0.584</b>

**Table 3.** Classification results of data sets from the methane meters MM264

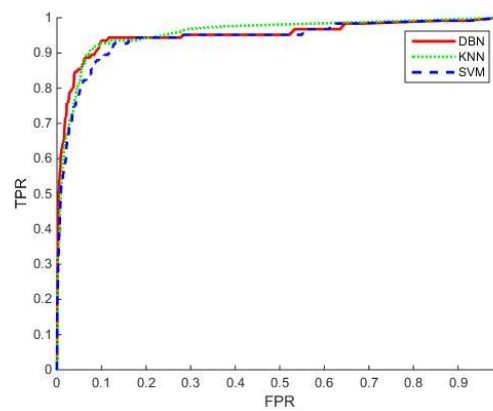
Alg	AUC	Accuracy	Specificity	Precision	Recall	Fmeasure
KNN	0.896	0.989	0.992	0.300	0.529	0.383
SVM	0.917	0.951	0.953	0.091	0.706	0.162
DBN	<b>0.939</b>	<b>0.992</b>	<b>0.995</b>	<b>0.425</b>	<b>0.500</b>	<b>0.460</b>

**Table 4.** Classification results of data sets from the methane meters MM256

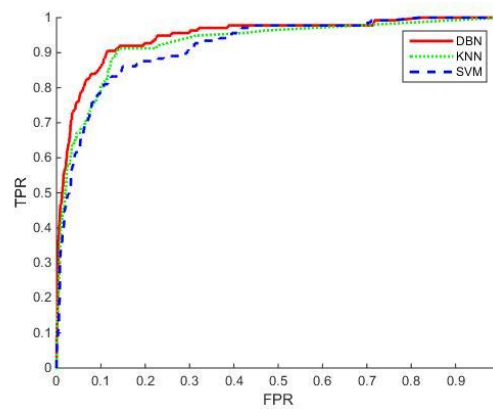
Alg	AUC	Accuracy	Specificity	Precision	Recall	Fmeasure
KNN	0.926	0.960	0.969	0.353	<b>0.613</b>	0.448
SVM	0.918	0.897	0.900	0.179	0.788	0.292
DBN	<b>0.946</b>	<b>0.963</b>	<b>0.972</b>	<b>0.390</b>	0.657	<b>0.489</b>



(a) MM263



(b) MM264



(c) MM256

**Figure 5.** The ROC curves of three methods classification result

#### 4. Conclusions

In this paper we presented our method based on DBN to classify the dangerous levels of methane concentration. Compared with the shallow classification learning method, deep learning method extracts the features of methane sensor through multi-layer network model, and obtains better features for classification. The experimental results show that better results can be obtained in classification effect. Therefore, more deep network models can be used to classify methane sensor data.

#### 5. Acknowledgments

This work was supported by the Natural Science Foundation of Jiangsu Province (No. BK20140216)



## 6. References

- [1] Wan X Y, Yang H, Sun L H, et al. Research of Coal Mine Gas Monitoring System Based on Technology Fusion of Wireless Sensor Network and 3DGIS[J]. *Advanced Materials Research*, 2011, 314-316:2263-2267.
- [2] Hui L I, Yun H U, Cun-Hua L I. The technique of gas disaster information feature extraction based on rough set theory[J]. *Journal of Shandong University*, 2013, 8(4).
- [3] Kang-Wu F , Qiu-Lin L I . Dynamic Analysis Method of Comprehensive Information on Gas Disaster Prevention and Control of Coal Mining Face[J]. *Coal Science & Technology*, 2013, 43(3):527-536.
- [4] Zhang S , Wang B , Li X , et al. Research and Application of Improved Gas Concentration Prediction Model Based on Grey Theory and BP Neural Network in Digital Mine[J]. *Procedia CIRP*, 2016, 56(Complete):471-475.
- [5] Zhang G, Liu Z J, Yang Y W, et al. Prediction of Coal and Gas Outburst by BP Neural Network[J]. *Applied Mechanics & Materials*, 2014, 539:664-668.
- [6] Gao, L. Application research of information in sensing and control system of coal mine. PH.D. Thesis, University of Mining and Technology, Xuzhou, China, 2006.
- [7] Hou Y.; Cheng J.; Li S. Forecasting coalmine gas concentration based on RBF neural network. *Proceedings of the International Conference on Information Acquisition*. 2007, 192-194.
- [8] Cheng, J.; Bai, J.; Qian J.; Li, S. short time prediction. *CUMT Journal*. 2008, 37, 231-235.
- [9] Yin, H. Coal mine gas time series analysis method and warning. PH.D. Thesis, University of Mining and Technology, Xuzhou, China, 2010.
- [10] Peng, H. Gas disaster prediction method research. PH.D. Thesis, University of Mining and Technology(Beijing), Beijing, China, 2013.
- [11] Kozielski M.; Skowron A.; Łukasz W. Regression Rule Learning for Methane Forecasting in Coal Mines. *Beyond Databases, Architectures and Structures*. Springer International Publishing, 2015: 495-504.
- [12] Hinton G E, Osindero S, Teh Y W. A Fast Learning Algorithm for Deep Belief Nets[J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [13] Pawłowski K, Kurach K. Detecting Methane Outbreaks from Time Series Data with Deep Neural Networks [M]. *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*. 2015:475-484.
- [14] Prasad S C, Prasad P. Deep Recurrent Neural Networks for Time Series Prediction[J]. *Computer Science*, 2014:1-19.
- [15] Janusz, A.; Sikora, M.; Wróbel, Ł.; Stawicki, S.; Grzegorowski, M.; Wojtas, P.; Ślęzak, D. Mining Data from Coal Mines: IJCRS'15 Data Challenge. *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*. 2015, 10, 429-438.