Gas Data Prediction Based on LSTM Neural Network

Danhui Li¹, Yingjun Zhang¹, Dali Gong², Lihu Pan¹, Yajuan Zhao²

¹Taiyuan University of Science and Technology, Institute of Computer Science and Technology, China, 030024
²Elite Digital Technology Co., Ltd, China, 030024

*Corresponding author e-mail: 461643989@qq.com

Abstract. To improve the prediction accuracy of the coal seam gas content in the unmined areas, based on the analysis and study of the main geological factors affecting coal seam gas content, fuzzy mathematics is used as a means of expressing and processing inaccurate data and fuzzy information and neural network as a way to solve the problem to combine fuzzy mathematics with neural network organically and establish a coal seam gas content prediction model based on LSTM neural network. The results of this study show that the LSTM neural network model can not only solve the problems of difficulty to express fuzzy information quantitatively and determine the learning samples, etc., but also extract the non-linear relationship between the coal seam gas content and its various influencing factors accurately. According to the instance operation verification, the prediction accuracy is improved by 4.84% ~ 25.79% compared with that of the neural network model. The effect of application to the prediction of coal seam gas content is more ideal. It has a good application prospect and can provide a theoretical basis for implementing scientific mine gas management and preventing various gas accidents.

Keywords: Coal seam Gas Content, Prediction, Non-linear, Fuzzy Mathematics, LSTM Neural Network

1. Introduction

China is endowed with abundant coal resources. Among the proven primary energy resource reserves, coal accounts for 94%, oil and natural gas account for 5.4% and 0.6%, respectively; while 95% of coal resources in China are mined underground, and the conditions for the formation of coal seams are complex[1-2]. With the deepening of coal seams, the gas content of coal seams has increased significantly.
The prediction of gas content in coal seams can not only guide the scientific application of outburst prevention measures, reduce the amount of engineering work for outburst prevention measures, but also ensure the personal safety of workers in outburst coal seams due to the constant inspection of outburst working surface. Hence, the prediction of coal seam gas content is of great practical significance\textsuperscript{[4-5]}. 

Given the complexity, dynamics, non-linearity, and random uncertainty of the coal seam gas content system, the prediction model should be further developed in a non-linear, multi-parameter direction. Many scholars have applied the theories of mathematical statistics, calculus, and artificial intelligence. Different methods have been used to establish different prediction models, with the common ones as follows: regression analysis, gray theory, artificial neural networks, etc. However, no effective solution has been found in the quantitative expression of fuzzy information. Fuzzy mathematics can address inaccurate problems in processing and expression, with its own unique advantages. The salient feature of fuzzy logic is that it can more naturally and directly express the logical meanings used by humans. According to some fuzzy rules, the quantitative expression of fuzzy information can be achieved. On this basis, the author takes the second coal seam of a coal mine as the research object. Through the analysis of various geological factors and quantitative and qualitative factors that affect the gas content of the coal seam, an LSTM neural network model is established. Combined with the fuzzy theory with the neural network, fuzzy mathematics is used as a means to express and process the inaccurate and fuzzy information, which can not only process fuzzy information and accomplish the fuzzy reasoning function but also accurately capture the nonlinear relationship between seam gas content and various influencing factors with the characteristics of the neural network\textsuperscript{[7]}, to achieve the purpose of predicting the coal seam gas content in unmined areas of a minefield.

2. Design of LSTM neural network model

There are various factors affecting the gas content of coal seams. In addition to the amount of gas generated, it mainly depends on the conditions of gas emission and migration after coal generation and the capacity of coal to retain gas, all of which ultimately depend on coalfield geological conditions and coal seam occurrence conditions. The main factors affecting the gas content are the degree of coal metamorphosis, the coal seam burial depth, the coal seam surrounding rock properties, the geological structure, the hydrogeological conditions, etc.

The deeper the burial depth of the coal seam, the higher the thickness of the overlying bedrock, the longer the distance that the gas in the coal seam moves to the surface, and the more difficult it is to escape. Meanwhile, the increase in depth also reduces the coal seam under the effect of overburden pressure. Air permeability, which is conducive to gas preservation

The distribution of gas is related to the thickness of the coal seam. In general, with the increase of the thickness of the same coal seam, the amount of gas generated is large, and the content is increased, that is, it reflects the amount of gas associated with the gas during the coal formation process. The gas generated in coal seams in the long geological ages presents different degrees of escape, which is closely related to the surrounding rock of the coal seam. When the lithology of the top and bottom of the coal seam is a dense and complete rock, such as mudstone and oil shale, the gas reservation is relatively easy; the roof is a rock with multi-porous or brittle fractures, such as conglomerate and sandstone, and gas is prone to escape.
After the above analysis, the author selected the buried depth of coal seam, the thickness of coal seam, the roof lithology, the thickness of the overlying bedrock, and the sandstone ratio 30 m above the roof as the influencing variables for predicting the gas content in the coal seam.

The basic principle of fuzzification is as follows: the input data space is divided according to the corresponding membership function, and the corresponding fuzzy rules are obtained. According to the type and number of membership functions, the input data can be divided into different groups. In general, the number of measured gas content data is small, and the data is simple, so the number of membership functions should be appropriately increased. The mixed membership function of triangle and trapezoid is used by the author, and the functional relationship is shown in Figure 1.

![Figure 1. Membership function of fuzzy variables](image)

The fuzzy calculation formula is as follows:

\[
\mu_{A_1}(x) = \begin{cases} 
1 & x < a \\
\frac{b-x}{b-a} & a \leq x \leq b \\
0 & x > b 
\end{cases} 
\]

(1)

\[
\mu_{A_2}(x) = \begin{cases} 
0 & x < a \text{ or } x > c \\
\frac{x-a}{b-a} & a \leq x \leq b \\
\frac{c-x}{c-b} & b < x \leq c 
\end{cases} 
\]

(2)

\[
\mu_{A_3}(x) = \begin{cases} 
0 & x \leq b \\
\frac{x-b}{c-b} & b < x \leq c \\
1 & x > c 
\end{cases} 
\]

(3)

In the formula: a, b, and c represent the critical values on which the variable interval is divided into fuzzy subsets; \( \mu_{A_1}(x), \mu_{A_2}(x), \mu_{A_3}(x) \) indicate the defuzzification processing of the data in the membership function. The maximum membership function method is used, that is, the element with the highest
degree of membership in the fuzzy set of all fuzzy rule reasoning results is used as the output value. The significant advantage of this method is simple calculation.

LSTM neural networks can be divided into 5 types (Table 1), and the type used in this article is type III, i.e., the input data of the neural network is first fuzzed and preprocessed. The input information contains qualitative knowledge and quantitative data to achieve input. The purpose of information fuzzification, and then the BP neural network is used to extract the fuzzy rules from known fuzzy data to infer the prediction results. The network model is shown in Figure 2.

| type   | type I (HNN) | type II (FNN1) | type III (FNN2) | type IV (FNN3) | type V (HFNN) |
|--------|--------------|----------------|-----------------|----------------|---------------|
| structure | Replace sigmoid function with "and or" operation | Weight value | The input value is the amount of blur | Input value and weight are both fuzzy | On the basis of FNN, use "and or" operation instead of S-shaped function |

| Burial depth | Coal seam thickness | Roof lithology | Bedrock thickness | Sandstone ratio of roof 30m |
|--------------|---------------------|----------------|------------------|-----------------------------|

![Figure 2. LSTM neural network prediction model diagram](image)

It is verified through the practice that the BP network based on the error back-propagation algorithm has a strong mapping capacity, which can complete the non-linear mapping between various influencing factors and the gas content, and implement the prediction of the gas content. Hence, the author established the neural network part of the model using the traditional BP neural network based on the backward transfer of errors. The burial depth of the coal seam, the thickness of the coal seam, the roof lithology, the thickness of the overlying bedrock, and the sandstone ratio 30 m above the roof are selected as the factors affecting the gas content of coal seam. Hence, the number of neurons in the fuzzy layer is 5, and the output value of the defuzzification layer is the gas content value. The number of neurons in the defuzzification layer is 1, and the membership function is determined to be 3. The number of corresponding input layer neurons is 15, and the number of neurons in the output layer is 3. The number of neurons in the hidden layer depends on the training effect of the network. Its empirical formula is as follows
\[ n_1 = \sqrt{n + m + d} \] (4)

Where \( n_1 \) is the number of hidden layer units; \( n \) is the number of input units; \( m \) is the number of output units; \( d \) is a constant between 0 and 10

3. Application of LSTM neural network in gas content prediction

3.1. Data preparation

The author takes a coal mine in Hemei as an example to collect and consolidate relevant data affecting the gas content of coal seams. After the data reliability analysis, a total of 19 geological exploration borehole gas content was obtained; 9 gas content measured during the production period. The 26 sets of data are used as training samples to complete the network fitting training. The latter two sets of data do not participate in learning training but are used as test samples to evaluate the prediction accuracy of the model.

3.2. Network training

The minimum mean square error of the network is set to 0.001, the learning rate is 0.3, the momentum coefficient is 0.8, and the maximum number of training is 25 000. The weights and thresholds of the network are randomly assigned to initial values between 0 and 1. After multiple training comparisons, the number of nodes in the selected hidden layer is 7. After the network training is completed, the training results are compared with those of the BP neural network training. The network error change trend is shown in Figure 3 and Figure 4.

![Figure 3. Error trend of BP neural network](image-url)
Figure 4. LSTM neural network error change trend

Figure 3 shows that the BP neural network reaches the fitting accuracy after 25 steps of training, and the network training stops. Figure 4 shows that after the fuzzy neural network reaches a certain number of training times, and the fitting accuracy no longer decreases. When the maximum training times is reached, the training stops. The final fitting accuracy of the network is 100.

4. Conclusions

1) According to the gas geological theory, based on the analysis and study of various geological factors affecting the gas content of coal seams, five main factors that determine the establishment of the LSTM neural network model are finally determined as follows: the burial depth of the coal seam, the thickness of the coal seam, the roof lithology, the thickness of the overlying bedrock, and the sandstone ratio 30m above the roof.

2) The fuzzy mathematics is combined with the multilayer BP neural network to establish an LSTM neural network model. The mixed membership function of triangle and trapezoid is used to implement the fuzzification of qualitative knowledge (roof lithology) and the quantitative data based on the established membership function.

3) The comparative analysis of the BP neural network model and the LSTM neural network model shows that the LSTM neural network model has a higher prediction accuracy and stronger applicability.

4) The gas content of coal seam is affected by multiple uncertain factors. The key to establishing an LSTM neural network model for predicting the gas content of coal seams lies in the selection of the main influencing factors, the establishment of a reasonable membership function, the sufficient measured gas content data, and appropriate network structure parameters. If the above aspects can be reasonably resolved, the model established will be more in line with the actual situation at the production site.

References

[1] Tian, Z. B., Zhang, Y. B., Wang, J. Q., Yang, X. D., & Xue, G. Q. (2016). Study on the pre-stack seismic inversion prediction method for rich coal-bed-gas reservoirs: a case in
southeastern shanxi province. Chinese Journal of Geophysics- Chinese Edition, 59(12), 4494-4504.

[2] Peng, Y., Wang, Y. , Zi, Y. Y. , Tsui, K. L. , & Zhang, C. . (2017). Dynamic reliability assessment and prediction for repairable systems with interval-censored data. RELIABILITY ENGINEERING & SYSTEM SAFETY, 159, 301-309.

[3] Zhang, L. , Hua, C. , Li, J. , & Guan, X. . (2016). Operation status prediction based on top gas system analysis for blast furnace. IEEE Transactions on Control Systems Technology, 25(1), 1-8.

[4] Ajani, A. , Kelkar, M. , Sarica, C. , & Pereyra, E. . (2016). Foam flow in vertical gas wells under liquid loading: critical velocity and pressure drop prediction. International Journal of Multiphase Flow, 87, 124-135.

[5] Islam, A. S. M. A. , Adoo, N. A. , & Bergstrom, D. J. . (2016). Prediction of mono-disperse gas-liquid turbulent flow in a vertical pipe. International Journal of Multiphase Flow, 85, 236-244.