Research on the Probabilistic Neural Network Model of Select Well-Layer in Coal Seam Refracturing

Xuesong Wan

College of Science, Southwest Petroleum University, Xindu District, Sichuan,610500, China.

wxsong19@163.com

Abstract

In coal seam gas mining, with the continuous precipitation of gas, the production is gradually reduced, which seriously affects the production. In view of this problem, the refracturing of the coal seam is one of the main measures at present. However, it is very important to select potential well-layer for refracturing, which has a significant impact on the production of after fracturing. Although with the deepening of the theoretical research and information collection of gradually complete, and the problem of accuracy of well-layer selection has been solved step by step, but for the reasons of the coal reservoir properties, requires a detailed analysis of parameters, in order to form the key method of selecting well-layer, so research the model of select the well-layer of refracturing has the extremely vital significance. In this paper, we propose to use the probabilistic neural network(PNN) model which has more advantages in classification and recognition to predict. First of all, theoretical research is carried out, then the prediction mathematical model is established, and finally, the model test is carried out by using the preprocessing data. The error of the prediction results of the test model is quite small, so the model is good for forecasting.

Keywords

Fuzzy recognition, Probabilistic neural network, Bayesian decision, Clustering analysis.

1. Introduction

At present, coal-bed methane has become an important part of daily life energy, however, due to the complexity of coal seam physical and chemical properties, almost all the gas wells can realize large-scale production after fracturing, even the part of the high yield of gas wells in the early, with the continuous precipitation of coalbed methane, the production will reduce gradually, therefore, improve coalbed methane production become the urgent matter. To solve this problem, the refracturing of the coal seam is the main measure at present. However, before the refracturing, it is very important to optimize the well-layer with mining potential for refracturing for production of fracturing. Therefore, it is necessary to research the technology of well-layer optimization[1-3].

At present, fuzzy mathematics[4], cluster analysis[5], and neural network[6,7] are common methods of well-layer optimization, and some effects have been achieved. However, for the uncertainty of the well-layer with few data and poor information, due to there is often a large error, so the prediction effect of the above technologies is not ideal.

About this situation of technical error during select well-layer, in this paper, a PNN model is proposed to predict production when selecting well-layer [8], and the established mathematical model is used to predict the yield by using the preprocessing data. The result shows that the prediction error is quite small, so this model is suitable for well-layer optimization.

2. Probabilistic neural network mathematical model.

2.1 Bayesian decision rule.

The main idea of bayesian decision making is that it can be described in the form of probability in the decision-making problem. When using Bayesian decision for pattern recognition requires two

conditions: one is the number of categories that need to be decided is certain; the other one is the probability distribution of the population of each category is known[9,10].

To this end, the basic Bayesian classifier is established:

$$\text{if} \quad p(\omega_i \mid x) = \min_{j \in n} p(\omega_i \mid x) \text{, for the } x \in \omega_i \tag{1}$$

Where $p(\omega_i \mid x)$ is the posterior probability;x is the sample vector of the d-dimensional eigenvector space;n is the total number of categories; ω_i is ith classification.

Theoretically, the Bayesian classifier can guarantee the minimum error rate in general mode classification. If the risk value is modified, it can become the classifier of the minimum risk decision. Therefore, the loss function is introduced in this paper $l(\omega_i, \omega_j)$, represents the actual is ω_i and classified as ω_j , so, the expected risk at a given observed value is:

$$R(\omega_i \mid x) = \sum_{j=1}^{n} l(\omega_i, \omega_j) p(\omega_i \mid x) \tag{2}$$

To sum up, establish the Bayesian decision rule with minimum risk:

$$R(\omega_i \mid x) = \sum_{j=1}^{n} l(\omega_i, \omega_j) p(\omega_i \mid x) \text{, for the } x \in \omega_i \tag{3}$$

As for the above Bayesian decision rules, distribution functions of prior probability and class conditional probability cannot be given by theory. Therefore, by means of statistical methods, real values and real distributions can be approximated infinitely according to the training of known classified samples. However, to achieve the minimum error rate or minimum risk, the number of samples is required to approach an infinite number of preconditions, so that the preconditions become a weakness of the Bayesian classifier[11].

2.2 Parzen window method.

Generally, in pattern recognition, this paper cannot get all the information about the probabilistic structure of the problem, but only some fuzzy and general information. In the face of the above problems, the Parzen window method is usually used for parameter estimation[12-14].

In the research of the probability density, Gauss developed Parzen's conclusion and proposed a special case of the multivariate gaussian kernel function.

From what has been discussed above, the estimation of probability density function can be expressed as:

$$f_\omega(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d N} \sum_{i=1}^{M} \exp[-\frac{(x - x_i)^T (x - x_i)}{2\sigma^2}] \tag{4}$$

Where N is the total number of training samples; M is the number of categories ω_i in the training sample; d is the dimension of the training sample; x_i is the ith vector of class W in the training sample; σ is the smoothing factor; $f_\omega(x)$ is the sum of multiple Gaussian distributions over each sample.

For the above equation(4), remove the common elements, then the discriminant function can be simplified as:

$$g(x) = \frac{p(\omega_i)}{N} \sum_{i=1}^{M} \exp[-\frac{(x - x_i)^T (x - x_i)}{2\sigma^2}] \tag{5}$$

2.3 Structural model of probabilistic neural network.

PNN model[15-17] based on Bayesian decision algorithm to determine the input vector of the category of the state identification model, due to its absorption of the RBF neural network and the advantages of the classic probability density estimation principle and the nature of strong nonlinear

mapping, compared with the traditional feedforward neural network, PNN model has significant advantage in classification and pattern recognition. Therefore it is well suited to do work related to the classification and pattern recognition[18-20].

PNN, in essence, uses the structure of the neural network to implement the Parzen window method, and its network structure is shown in figure 1. As we can be seen from the figure, the PNN structure is composed of four layers, namely, input layer, sample layer, summation layer and output layer (the sample layer and summation layer can also be called hidden layer). The equation (5) can be effectively calculated when PNN is used for parallel processing [21,22].

Algorithm steps of PNN:

In the first step, the input vector $\vec{x}$ is input into the input layer, and the difference $\vec{x}-\vec{x}_{ik}$ between the input vector and the training sample vector are calculated by the network, where the size $||\vec{x}-\vec{x}_{ik}||$ represents the distance between $\vec{x}$ and $\vec{x}_{ik}$, and the resulting vector is output by the input layer.

In the second step, the output vector $\vec{x}-\vec{x}_{ik}$ of the input layer will be sent to the sample layer, where the number of sample layer nodes are equal to the sum of the number of training samples, that is, $N=\sum_{i=1}^{M} N_i$, where M is the total number of classes.

In the third step, the output value of the sample layer is sent to the summation layer, where the number of nodes in the summation layer are M, and each node corresponds to a class, and then the decision is made by the competitive transfer function of the summation layer.

Finally, the result of the decision is output by the competitive layer, and the output result of the class with the largest probability value is 1.

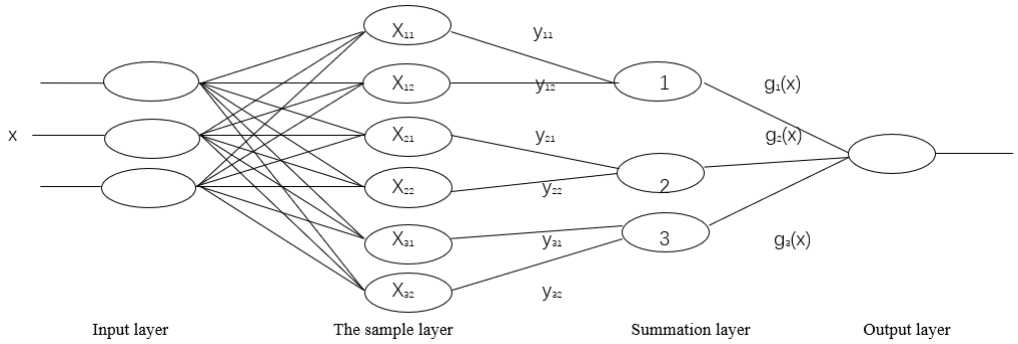

Fig.1 Probabilistic neural network structure

2.4 Establishment of probabilistic Neural Network Model.

The PNN model is a classical forward network, which is very suitable for classification and pattern recognition. With the in-depth theoretical research, the data collection is more complete, and the problem of large error in the well-layer selection are gradually solved. However, for the reasons of coal reservoir properties, different parameters need to be deeply analyzed to form the most critical well-layer selection methods, which usually include the yield statistics method and fuzzy identification model. At present, the method of well-layer selection has developed from a simple non-analytic statistical method with low data requirements to a detailed and multi-parameter comprehensive identification analysis method. Therefore, using a PNN model to predict production is of great significance for the study of well-layer selection methods in refracturing.

According to the parameter analysis results that affect the comprehensive evaluation of coal seam, the input layer is composed of 30 parameters, including:effective thickness, perforation thickness, distance between water layer and fracturing layer, preflush percentage, the total fluid volume of construction, total sand amount of construction, average sand ratio, fracturing fluid type, the highest construction pressure, the minimum construction pressure, compensated density, compensated neutron, hole diameter *X*, hole diameter *Y*, deep lateral electrical resistance ,shallow lateral electrical

resistance, natural potential, natural gamma, interval transit time, bottom hole pressure before fracturing, dynamic fluid level before fracturing, accumulated water production before fracturing, depth of moving liquid surface from coal seam, number of days for start interval fracturing, initial bottom-hole pressure, current bottom hole pressure, initial dynamic level depth, current level of dynamic liquid level, high rank time.

The output layer is composed of one parameter, that is, predicting the classification results of samples.

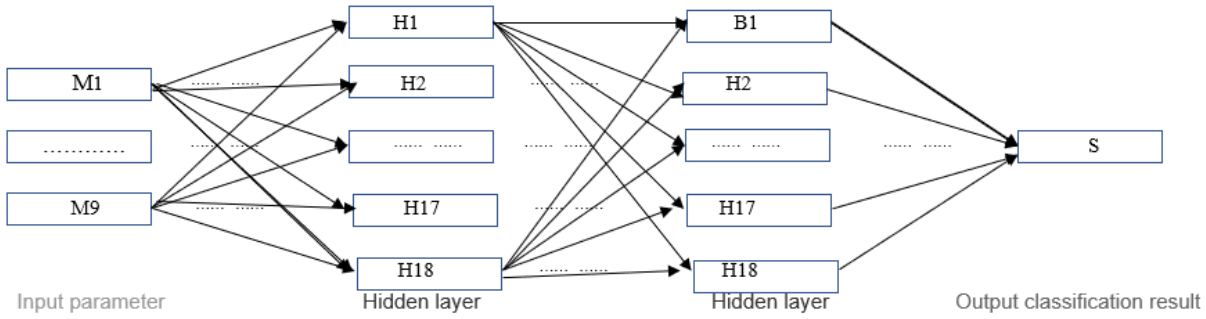

Fig. 2 Probabilistic neural network prediction model

In terms of time, the extraction process is divided, and the effects of initial and end pressure of well bottom, and the dynamic liquid face value of different time periods on production are taken into account, respectively. Neural network training is carried out for different cumulative discharge and production stages of the same well-layer.

2.5 PNN algorithm

Since PNN can use the linear learning algorithm to complete the work of the nonlinear learning algorithm, it has a good advantage in classification and pattern recognition.

In pattern classification, Bayes classification mainly deals with many sample classification problems, and the main classification rule is the optimization decision of minimum expected risk. Generally, when applied to Bayes classification rules, the probability density function is unknown and only many pattern samples are known. Therefore, this paper uses these large numbers of known training samples to train the PNN.

The first step: in order to reduce the error and effectively avoid the phenomenon of decimals being eaten by large numbers, it is necessary to normalize the matrix.

Let the original input matrix be:

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots\cdots & X_{1m} \\ X_{21} & X_{22} & \cdots\cdots & X_{2m} \\ & & \cdots\cdots & \\ X_{n1} & X_{n2} & \cdots\cdots & X_{nm} \end{pmatrix} \tag{6}$$

It can be seen from the matrix (6) of the sample that there are m learning samples, and each sample has n characteristic attributes. Before the normalization factor is solved, the matrix B^T must be calculated:

$$B^T = \left(\frac{1}{\sqrt{\sum_{k=1}^{n} x_{1k}^2}}, \frac{1}{\sqrt{\sum_{k=1}^{n} x_{2k}^2}} \cdots\cdots \frac{1}{\sqrt{\sum_{k=1}^{n} x_{nk}^2}} \right) \tag{7}$$

Further calculation:

$$C_{m\times n} = B_{n\times 1}[1\ \ 1\ \ \cdots\ 1] \bullet \mathrm{X}_{m\times n} = \begin{pmatrix} \frac{x_{11}}{\sqrt{M_1}} \cdots\cdots \frac{x_{1n}}{\sqrt{M_1}} \\ \cdots\cdots \\ \frac{x_{m1}}{\sqrt{M_m}} \cdots\cdots \frac{x_{mn}}{\sqrt{M_m}} \end{pmatrix} = \begin{pmatrix} C_{11} \cdots\cdots C_{1n} \\ \cdots\cdots \\ C_{m1} \cdots\cdots C_{mn} \end{pmatrix} \tag{8}$$

Where $M_1 = \sum_{k=1}^{n} X_{1k}^2$, $M_2 = \sum_{k=1}^{n} X_{2k}^2 \cdots\cdots M_m = \sum_{k=1}^{n} X_{mk}^2$, and the normalized learning matrix is *C*.

The second step: classify the matrix in the sample layer. if the total number of samples is *m*, and divide it into *c* categories when the number of samples is the same. Let the number of samples in each category be *k*, and then *m=k*c*.

The third step: distance calculation of pattern. Assuming that the sample matrix to be identified is composed of *P n*-dimensional vectors, after normalization, the input sample matrix to be identified is:

$$D = \begin{pmatrix} d_{11} & \mathrm{d}_{12} & \cdots\cdots & \mathrm{d}_{1n} \\ & \cdots & \cdots & \\ d_{p1} & \mathrm{d}_{p2} & \cdots\cdots & \mathrm{d}_{pn} \end{pmatrix} \tag{9}$$

Calculate the distance between vectors: the distance between the center vector of each network node in the sample layer and the sample vector to be identified:

$$E = \begin{pmatrix} \sqrt{\sum_{k=1}^{n} | d_{1k} - C_{_{1k}} |^2} \cdots\cdots \sqrt{\sum_{k=1}^{n} | d_{1k} - C_{_{mk}} |^2} \\ \cdots \qquad\qquad \cdots\cdots \qquad\qquad \cdots \\ \sqrt{\sum_{k=1}^{n} | d_{pk} - C_{_{1k}} |^2} \cdots\cdots \sqrt{\sum_{k=1}^{n} | d_{pk} - C_{_{mk}} |^2} \end{pmatrix} = \begin{pmatrix} E_{11} \cdots\cdots E_{1m} \\ \cdots\ \cdots\ \cdots \\ E_{p1} \cdots\cdots E_{pm} \end{pmatrix} \tag{10}$$

The fourth step: activate the neurons of the radial basis function in the sample layer. Normally, the gaussian function with standard deviation σ =0.1, and the initial probability matrix after activation is obtained as follows:

$$P = \begin{pmatrix} e^{-\frac{E_{11}}{2\sigma^2}} & e^{-\frac{E_{12}}{2\sigma^2}} & \cdots\cdots e^{-\frac{E_{1m}}{2\sigma^2}} \\ & \cdots \quad \cdots \quad \cdots & \\ e^{-\frac{E_{p1}}{2\sigma^2}} & e^{-\frac{E_{p2}}{2\sigma^2}} & \cdots\cdots e^{-\frac{E_{pm}}{2\sigma^2}} \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} \cdots\cdots P_{1m} \\ & \cdots\cdots\cdots & \\ P_{p1} & P_{p2} \cdots\cdots P_{pm} \end{pmatrix} \tag{11}$$

The fifth step: obtain the initial probability sum of each sample belonging to each type in the summation layer:

$$S = \begin{pmatrix} \sum_{l=1}^{k} P_{1l} & \sum_{l=k+1}^{2k} P_{1l} & \cdots & \sum_{l=m-k+1}^{m} P_{1l} \\ & \cdots\ \cdots & \vdots & \\ \sum_{l=1}^{k} P_{pl} & \sum_{l=k+1}^{2k} P_{pl} & \cdots & \sum_{l=m-k+1}^{m} P_{pl} \end{pmatrix} = \begin{pmatrix} S_{11} & \mathrm{S}_{12} & \cdots\cdots & \mathrm{S}_{1c} \\ & \cdots & \cdots & \\ S_{p1} & \mathrm{S}_{p2} & \cdots\cdots & \mathrm{S}_{pc} \end{pmatrix} \tag{12}$$

On the type, S_{ij} represents the initial probability sum of the *i*th sample belonging to the *j* class in the sample to be identified

Step six: calculate the probability *Prob* b_{ij}, that is, the probability that the *i*th sample belongs to the *j* class.

$$probb_{ij} = \frac{S_{ij}}{\sum_{l=1}^{c} S_{il}} \tag{13}$$

2.6 Model Testing.

According to the algorithm of the PNN model, the corresponding program was written by Matlab. Due to the limited available data, 30 wells were selected as model prediction, 25 wells as training samples, and 5 wells as prediction samples. Table 1 and Table 2 are the basic data of classification samples, table 3 is the data after PNN training, and table 4 is the error after PNN training.

Table1 Basic data of classified samples (part)

well number	The depth of coal seam	the thickness of coal seam	Average discharge	prepaid fluid	The total fluid volume	Amount of sand	Percentage of preflush	Sand percentage	The intensity of sand	GR
H1	849.3	16.3	8.5	360	762	33	0.4724	0.0821	2.0245	69.54
H2	776	12.8	7.7	432.4	1040	36.8	0.4156	0.0605	2.8750	105.6
H3	1114	14.5	8.6	480	1150	45	0.4174	0.0672	3.1034	94.04
H4	771.6	14.1	7.7	357.1	961.3	45.1	0.3715	0.0746	3.1986	59.41
H5	623.1	23.75	8.7	500	1110	55	0.4505	0.0902	2.3158	55.18

Table 2 Classified sample basic data(part)

Well number	SP	CALX	CALY	MSFL	LLS	LLD	DT	RHOB	CNL
H1	-16.732	22.294	22.432	1230.53	2111.98	2551.32	407.77	1.40	45.014
H2	-43.664	25.237	25.406	1212.87	7450.69	7914.34	375.05	1.51	44.661
H3	163.698	25.256	21.864	389.217	9619.64	9071.58	347.05	1.68	71.771
H4	-39.543	22.341	22.304	166.517	1089.36	898.324	424.20	1.36	39.11
H5	96.404	23.282	21.539	1803.01	2010.29	2333.09	382.66	1.47	74.317

Table3 Data after probabilistic neural network training(part)

Well number	data	Well number	data	Well number	data	Well number	data	Well number	data
H1	3041.32	H2	3203.73	H3	179.71	H4	4128.32	H5	677.06

Table 4 Rate the error after neural network training(part)

Well number	error	Well number	error	Well number	error	Well number	error	Well number	error
H1	-0.32	H2	0.27	H3	0.29	H4	-0.32	H5	-0.06

3. Conclusion

In this paper, we carried out theoretical research on PNN first of all, and based on the theory, established the corresponding mathematical model, and finally used pre-processing data to test the established mathematical model. Table 4 is the data predicted by this model. It can be seen from Table 4 that, after the training of the PNN model, although certain errors are generated, the errors are between [-0.5, 0.5], which is quite small. Therefore, this model is suitable for well-layer optimization in the refracturing of the coal seam.

References

[1] Ge Niu, D.Y, Li, etc. Simple analysis on the principle of well selection and layer selection for self-flowing water separation in carbonate reservoir [J], China petroleum and chemical industry standards and quality, 2008,38(20):9-10+12. (In Chinese)

[2]L.Y.Song.J.W.Wang, etc, Optimization of repeated fracturing technology for horizontal Wells in tight reservoirs [J], Evaluation and development of oil and gas reservoirs,2018,8(04):63-67. (In Chinese)
[3]P.F.Zhang, Research and application of fracturing well selection and layer selection [J]. Logging technology abroad,2017,38(03):42-45. (In Chinese)
[4]Y.Q.Fu,J.Z.Zhao et al., Application of multi-level fuzzy clustering in fracturing and acidizing well selection, Gas industries, 2001;21(5):58～ 60. (In Chinese)
[5] Wei Wang, Application of cluster analysis in the selection of horizontal wells by repeated fracturing. Chemical Engineering & Equipment 2018. 3. (In Chinese)
[6]Hong Liu,J.Z.Zhao, Application of Fuzzy Neural Network system in optimal selection of fracturing well layer, Drilling and Production process,2002. (In Chinese)
[7]Z.K.Yi,Hai Jiang, Application of Artificial Neural Network in well selection and layer selection by compound fracturing, Western prospecting project,2007.9. (In Chinese)
[8]J.S.Fan,Y.Z.Su, New Development of structural damage Identification based on probabilistic Neural Network[J], The world of science and technology ,2016, 4. (In Chinese)
[9] H.G.Fan,J.J.Pei,Yang Zhao. An optimized probabilistic neural network with unit hyperspherical crown mapping and adaptive kernel coverage[J]. Neurocomputing,2019.
[10]GERBEC D, GASPERIC S, SMON I, et al. Determining the load profiles of consumers based on fuzzy logic and probability neural networks [J]. IET Proceedings - Generation Transmission and Distribution, 2004, 151(3): 395-400.
[11]SPECHT D F. Probabilistic neural networks and the polynomial Adaline as complementary techniques for classification [J]. IEEE Transactions on Neural Networks, 1990, 1(1): 111.
[12]BISHOP C M. Bayesian methods for neural networks [J]. Neural Computing Research Group, 1995,
[13]BERGER J O. Robust Bayesian analysis: sensitivity to the prior[J]. Journal of Statistical Planning & Inference, 1990, 25(3): 303-28.
[14] S.M.M. Fatemi Bushehri,Mohsen Sardari Zarchi. An expert model for self-care problems classification using probabilistic neural network and feature selection approach[J]. Applied Soft Computing Journal,2019,82.. (In Chinese)
[15]W.J.Liu,Qi Li, Study on Fault Diagnosis of PEMFC Water Management Based on Probabilistic Neural Network and Linear Discriminant Analysis, Proceedings of the CSEE Vol.39 No.12 Jun. 20, 2019. (In Chinese)
[16]T.X.Wen,F.E.Yu, Prediction Model of Coal and Gas outburst probability Neural Network based on Grey Relational Entropy, Application Research of Computers Vol. 35 No. 11 Nov. 2018. (In Chinese)
[17] j.j.Xu, Min Jiang Particle Swarm Optimization Based Algorithm for Conditional Probability Neural Network Learning Artificial Intelligence and Robotics Research, 2016, 5(1), 13-22
[18]Na Liu, Fault diagnosis of hydropower units based on probabilistic neural networks [J]. Wind of science and technology,2019(29):163. (In Chinese)
[19]B.S.Wang,Y.Q.Ni, A probabilistic neural network is used to identify the location of structural damage[J], Journal of vibration engineering, 2001, 14(1): 60-4. (In Chinese)
[20]D.H.Li,Hao Liu, Fault diagnosis method and application based on probabilistic neural network[J], Systems engineering and electronics, 2004, 26(7): 997-9. (In Chinese)
[21]PARZEN E. On Estimation of a Probability Density Function and Mode [J]. Annals of Mathematical Statistics, 1962, 33(3): 1065-76.
[22]DUIN R P W. On the Choice of Smoothing Parameters for Parzen Estimators of Probability Density Functions [J]. IEEE Transactions on Computers, 1976, C-25(11): 1175-9.