ISSN: 2414-1895 DOI: 10.6919/ICJE.202208_8(8).0085

Automatic Identification and Application of Logging Deposition Microphasic

Jingyue Zhang

School of computer Science, Yangtze University, Jingzhou 434023, China

Abstract

Sedimentary studies Have a direct impact on the development of oil and gas in China. Reservoir sedimentary phase research can better describe the reservoir, ensure the smooth implementation of exploitation, and use logging data to quickly and accurately determine the sedimentary microface is a problem that needs to be studied and solved in oilfield exploration and development. In the process of determining non-structural hidden oil and gas reservoirs, it can be achieved through the study of sedimentary facies and sedimentary environments, but the research on these two aspects is still relatively lacking, and the effect is not very good.In this paper, Bayesian network and bp neural network methods are used to achieve the identification of deposited microphases. The main steps include: logging data extraction and processing, sedimentary phase and logging data mapping establishment, and sedimentary phase identification based on the data.

Keywords

Sedimentary Phase Classification; Fisher; BP Neural Network Model; Bayesian Networks.

1. Introduction

As far as logging data is concerned, it is an indirect data for the study of geological conditions, and sedimentology is to transform these data into various geological models and models, and then use these models and models to explain the underground geological conditions, that is, including both forward and reverse. Positive problem: to establish a corresponding geological model and model for various geological phenomena that need to be studied in nature, this model and model can be divided into two categories, namely mathematical models and geometric models, mathematical models quantitatively describe underground geological phenomena, and geometric models are qualitative descriptions of geological phenomena. Inversion problem: Establish a relationship with various geological models and models with various logging parameters and curve forms in order to correctly reflect underground geological phenomena. The problem of inversion includes two factors, one is the objective factor, that is, the accuracy of the logging data, and the introduction of new logging technology related to logging sedimentology; The other is the subjective factor, that is, the addition of human thought in the process of inference and hypothesis, which is also the key to the problem of inversion, and human intelligence analysis and expert knowledge should be comprehensively utilized.

2. Deposition Phase Recognition based on BP Neural Network

2.1 What is the BP Neural Network Model

BP network is a multi-layer feed-forward network trained by error reverse propagation algorithm, which is composed of two processes: forward propagation of information and backpropagation of error. Each neuron in the input layer is responsible for receiving input information from the outside world and passing it to the neurons in the middle layer; the middle layer is the internal information

ISSN: 2414-1895 DOI: 10.6919/ICJE.202208_8(8).0085

processing layer, responsible for information transformation, which can be designed as a single hidden layer or a multi-hidden layer structure according to the needs of information change capabilities; the neurons in the output layer are responsible for the processing results of the output information to the outside world. When the actual output does not match the desired output, the backpropagation phase of the error is entered. The error corrects the weights of each layer by the output layer (or otherwise) by way of the error gradient descent and reverses the hidden layer and the input layer layer by layer. The repeated process of forward propagation and error backpropagation of information is the process of continuous adjustment of the weights of each layer, and it is also the process of neural network learning and training. This process continues until the error in the network output is reduced to an acceptable level, or until the number of pre-set studies is set [1].

2.2 Extraction of Log Curve Feature Parameters

Through the study of the core and its corresponding sedimentary facies, the mapping conversion relationship from each logging phase to the sedimentary phase is determined by mathematical methods and knowledge reasoning, and the sedimentary phase library is established by using this relationship. After layering the natural gamma log curve, the curve characteristic parameters within each layer segment are extracted[2].

2.3 The Process of Deposition Phase Recognition of BP Neural Network Models

From a set of logging responses that can reflect the characteristics of the formation, the variation characteristics of the logging curve are extracted, including amplitude characteristics, morphological characteristics, etc. and other logging interpretation conclusions (such as sedimentary structure, paleowater flow direction, etc.), the formation profile is divided into a limited logging phase, these logging phases are scaled by geological data such as core analysis, and the mapping conversion relationship between each logging phase to geological phase is determined by mathematical methods and knowledge reasoning, and finally the sedimentary phase of the formation is described and studied by logging data.

Well logging phase analysis has roughly the following basic steps:

a) Depth correction of a selected set of logging curves;

b) Environmental impact correction of logging curves and standardization of whole well field range;

c) Automatic stratification of logging curves;

d) Extraction of log curve feature parameters;

e) Establish the corresponding relationship between the sedimentary phase and logging, that is, the sedimentary phase bank;

f) The deposited phase is identified using a BP neural network.

3. Bayesian Network Sedimentary Phase Identification Method

3.1 Bayesian Network Classification Principles

Bayesian networks are directed acyclic graphs that apply probabilities to represent dependencies between variables, determined by the network structure and network parameters. The network structure consists of nodes representing variables and directed edges connecting these nodes, where the directional edges represent the causal relationship between the variables. The network parameters consist of a series of conditional probability tables about causality between variables, and these conditional probability tables list all the conditional probabilities of each node relative to its parent node.

3.2 A Process based on Bayesian Network Classification

(1) Define variables. Define variables based on the attribute parameters used for classification in your study, including variables that represent the individual attribute parameters that participate in the classification and categorical variables that represent the final classification results. For example,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202208_8(8).0085

variables defined in the Sedimentary Phase Classification of Bayesian Networks include sedimentary phase categorical variables and variables representing the sedimentary properties involved in the classification. Due to the different numerical sizes and ranges of each attribute parameter, and considering that bayesian networks based on discrete variables are easier to implement algorithmically, each attribute parameter can be discretized and standardized according to certain rules.

(2) Network training. This includes network structure training and network parameter training. According to the different Bayesian network training algorithms, the network training algorithm can only define the variables before network training, but not the relationship between the variables (directed edges), and the network structure and network parameters are completely obtained by the network training algorithm; the initial network model can also be constructed according to the relationship between the variables in the prior information to improve the accuracy of network training.

(3) Classification of data. Based on the trained Bayesian network classification model, the network parameters of the corresponding variables are queried, and the posterior probability value $P(C|X)$ corresponding to each category of the sample to be classified can be calculated through the above formula, that is, when the indicator parameter is valued as $X= (X_1, X_2, X_3,...,X_n)$, the probability of the deposited phase category $C= c_i$. The category with the largest probability value is considered to be the same qualitative classification result as the traditional method classification result.

3.3 Sedimentary Phase Classification Ideas and Processes

The basic idea of sedimentary phase classification based on Bayesian network is to effectively integrate the prior information about sedimentary phases and the distribution characteristics of typical sedimentary phase training samples, construct a reasonable sedimentary phase classification model, obtain the probability values of sedimentary phases in the form of probabilistic reasoning, and estimate the reliability of sedimentary phase classification results according to the probability distribution. The data required for sedimentary phase classification includes prior information, sample data for training (sedimentary properties) and instance data to be classified (sedimentary properties). The goal of classification is to divide the deposited facies into different categories, the specific steps of which are as follows[5]:

a)Extraction of sedimentary properties;

b) Pre-treatment of sedimentary attribute data;

c) Establish the initial Bayesian network model and set the prior information;

d) Enter learning samples to train the network;

e) The trained classification model is used for the classification of sedimentary phases.

3.4 Construction and Training of Classification Models

Before the network structure training, only the network nodes are defined as the initial network structure, and before the network parameters are learned, the conditional probability distribution table of each node parameter is set to a uniform distribution as the initial network parameter of bayesian network parameter learning.

4. Sedimentary Phase Classification Ideas and Processes

4.1 Sedimentary Phase Identification Step

The goal of classification is to divide the deposited facies into different categories, the specific steps of which are as follows:

(1) Extraction of sedimentary properties;

(2) Pre-treatment of sedimentary attribute data;

(3) Establish an initial Bayesian network model or bp neural network model;

ISSN: 2414-1895 DOI: 10.6919/ICJE.202208_8(8).0085

(4) Enter learning samples to train the network;

(5) The trained classification model is used for the classification of sedimentary phases.

4.2 Features are Preferred

Fisher criterion is one of the effective methods of feature selection, the main idea is to identify the characteristics with strong performance as small as possible and the distance between classes as large as possible. The features are sorted and those with strong identification performance are selected to achieve the purpose of dimensionality reduction.

Define that there are n samples in the dataset that belong to class C ω1, omega2,..., ωC, and each class contains samples, representing the mean of sample x class i samples, and the mean of all samples takes the value on the k-dimensionality. The Fisher criterion for a single feature is expressed as:

$$J_{Fisher}(k) = S_B^{(k)} / S_W^{(k)} \tag{1}$$

where and represent the inter-class variance and intra-class variance of the dimension feature on the training sample set, respectively.

$$S_B^{(k)} = \sum_{i=1}^{c} \frac{n_i}{n} (m_i^{(k)} - m^{(k)})^2 \tag{2}$$

$$S_W^{(k)} = \frac{1}{n} \sum_{i=1}^{c} \sum_{x \in \omega_j} (x^{(k)} - m_i^{(k)})^2 \tag{3}$$

Equations (2) and (3) are expressions for the intra-class variance and interclass variance of the k-dimensional feature, respectively.

Known as the Fisher ratio or Fisher criterion of a feature, the larger the Fisher ratio of a certain dimension feature on the training sample set indicates that the better the class differentiation of the dimension feature, that is, it contains more discriminating information, while the noise feature tends to be 0.

4.3 Correlation Analysis

Correlation analysis refers to the analysis of two or more variable elements with correlation to measure the closeness of the factors of two variables. There needs to be a certain relationship or probability between the elements of correlation before correlation analysis The method of correlation analysis is mainly used to analyze the size of the correlation between two things.

5. Conclusion

Analysis of neural network recognition results:

a) The mapping effect of identification combined with logging characteristic parameters is relatively obvious, and the identification accuracy of the sedimentary phase is relatively high. However, there will still be misjudgments in recognition, especially for thin layer segments, indicating that the neural network responds to thin formations.

b) Most of the current logging sedimentary phase research is single-well research, in further research work, multiple wells and a variety of logging curves can be comprehensively studied, and gradually developed into a computer three-dimensional sedimentary phase automatic display technology.

Bayesian network identification results analysis:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202208_8(8).0085

The transcendental information can be introduced into the process of classification, and the a priori information and the distribution characteristics of the training samples can be effectively used to guide the classification, and the probability values of each sedimentary phase class can be obtained in the form of probabilistic reasoning, so as to estimate the reliability of the classification results;

Identify the reasons for the low accuracy:

a) Near the layer boundary point, the data changes dramatically, and the properties of each layer microphasic are not obvious.

b) Sedimentary phases are classified by layer, not by each record.

References

[1] Zheng Zhou. BP Divine Network Manifestation Review [J]. Shanxi Electronic Technology ,2008(2): 90-92.

[2] Cancan Wu, Zhuangfu Li. Logging phase analysis and sedimentary phase identification based on BP neural network[J].Coalfield Geology and Exploration,2012,40(01):68-71.

[3] Jiao,Y.Q., Wu,L.Q.,Rong,H.,et al.,2012.Paleoecology of the Ordovician Reef-Shoal Depositional System in the Yijianfang Outcrop of the Bachu Area,West Tarim Basin.Journal of Earth Science , 23(4):408-420.

[4] Zhu,H.T.,Yang,X.H.,Zhou,X. H.,et al.,2013.Sediment Transport Pathway Characteristics of Continental Lacustrine Basins Based on 3-D Seismic Data:An Example from Dongying Formation of Western Slope of Bozhong Sag.Earth Science-Journal of China University of Geosciences,38(1):121-129 (in Chinese with English abstract).

[5] Yuan Gu,Peimin Zhu,Hui Rong,Fanping Zeng,Yang Hai. Seismic phase classification based on Bayesian network[J].Earth Sciences (Journal of China University of Geosciences),2013,38(05):1143-1152.

[6] Chen Tianyi,Wong Yiik Diew,Shi Xiupeng,Wang Xueqin. Optimized structure learning of Bayesian Network for investigating causation of vehicles on-road crashes[J]. Reliability Engineering and System Safety,2022,224.

[7] Sarker Ronobir,Kaur Amandeep,Singh D.. Noise Estimation Using Back Propagation Neural Networks[J]. Electrochemical Society Transactions,2022,107(1).

[8] ZHANG Runlian,ZHANG Zhao,PENG Xiaojin,ZENG Bing. Feature Selection Algorithm Based on Fisher Pointing and Support Vector Machine[J].Computer Engineering and Design,2014,35(12):4145-4148+4190.