Journal of Applied Geophysics 124 (2016) 17–26

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

Supervised classification of down-hole physical properties measurements using neural network to predict the lithology

Omid Mahmoodi *, Richard S. Smith, Douglas K. Tinkham

Department of Earth Sciences, Laurentian University, Sudbury, ON P3E 2C6, Canada

ARTICLE INFO

Article history:
Received 11 March 2015
Received in revised form 4 November 2015
Accepted 10 November 2015
Available online 11 November 2015

Keywords:
Victoria
Down-hole
Physical properties
Neural network
Lithology
Physical unit

ABSTRACT

The reliability of rock-type prediction using down-hole density, gamma ray response, and magnetic susceptibility measurements was evaluated at the Victoria property, Sudbury, ON. A supervised neural network, trained using lithological information from drill hole FNX1168, yields a predictive accuracy of 83% for the training data. Applying the trained network on drill hole FNX1182 resulted in 64% of the rock types being correctly classified when compared with the classification produced by geologists during logging of the core. The homogenous rock types, like quartz diorite, had a high accuracy of classification; while the heterogeneous rock types such as diabase were poorly classified. Overlap between physical properties of rock types caused by heterogeneity or inherent similarity in physical properties of rock types, which were verified by observing the cores, accounts for most of the misclassification. To reduce the misclassification, the network was trained from physical log units in FNX1168 derived from clustering of physical properties measurements. Four physical log units mainly represented four groups of rocks: i) quartz diorite; ii) metabasalt and metagabbro; iii) metasediment and quartzite; and iv) sulfide and diabase. The predictive accuracy in the training process rose to 95%. The trained network then was applied to predicting the physical log units in FNX1182. Given the relationships between physical log units and rock types from FNX1168, the results of physical-log-unit classification in FNX1182 were interpreted from a geological point of view. Although in ideal cases we would like to be able to extract the same classification that a geologist provides, the extraction of physical log units is a more realistic goal. The interpretation of the lithological units from the physical log units can be compared with the geologist's classification and discrepancies or anomalies analyzed in greater detail.

1. Introduction

Typically when a hole is drilled, a geologist will look at the core extracted and classify the lithology or rock type as a function of depth down the hole. Rock type prediction based on log data from down-hole geophysical measurements can be considered as a potential alternative to a geologist's log when the cores are not fully recovered such as ocean drilling or drilling methods which do not provide cores such as percussion drilling in mineral exploration (Benaouda et al., 1999 and Qi and Carr, 2006). Physical properties logging provides a continuous set of data down the hole, which can be effectively used to improve the understanding of the geological characteristics of the hole (Killeen, 1997; McDowell et al., 1988; Granek, 2011). The accuracy of rock-type characterization based on physical properties is proportional to the existent contrast of these data between rock types (Perron et al., 2011 and Mwenifumbo and Mwenifumbo, 2012). But, similarity of physical properties of rocks and the heterogeneity of the rock increases the overlap between physical properties of different rocks (Rabaute et al., 2003; Garcia et al., 2011). Overlap of physical properties between two rocks brings confusion to the prediction of rock types. We feel that it is more realistic to classify them as physical units. The term “physical log units” can be described as homogenous intervals of one or more rock types with consistent physical properties. The link between physical units and rock types can help geophysical studies to gain a better understanding of the geological setting (Benaouda et al., 1999 and Perron et al., 2011).

Conventional statistical techniques such as using histograms, box and whisker plots, cross plots, or the analysis of average and variance, were employed by Reed et al. (1997); Killeen (1997); McDowell et al. (1988, 2004), and Vella and Emerson (2009) to extract the pattern of variation in physical properties measurements and relate them to a geological setting. Recently, the multi-variable pattern recognition techniques have become popular, since the conventional methods were limited in terms of the number of variables, and their ability to establish a quantitative relationship between physical properties and a geological setting (Rabaute et al., 2003; Qi and Carr, 2006; Williams and Dipple, 2007; Garcia et al., 2011 and Granek, 2011). Supervised classification techniques can be used when the goal is to predict rock types based

* Corresponding author.
E-mail addresses: omahmoodi@laurentian.ca (O. Mahmoodi), rssmith@laurentian.ca (R.S. Smith), dtinkham@laurentian.ca (D.K. Tinkham).

http://dx.doi.org/10.1016/j.jappgeo.2015.11.006
0926-9851/

on the physical properties measurements. As the term 'supervised' implies, a drill hole or parts of a hole with both physical properties measurement and rock type information are used to train the classifier, and then the trained classifier can be applied to a new hole or those parts lacking the core to predict the rock types (Benaouda et al., 1999 and Qi and Carr, 2006).

Varying composition of rock types, structures, alteration and mineralization cause non-linear variation in physical properties, which justifies using a non-linear classifier to analyze these data. The neural network is a robust non-linear classifier successfully applied to down-hole logging data (Ojha and Maiti, 2013). Baldwin et al. (1990); Wong et al. (1995); Farmer and Adams (1998); Qi and Carr (2006) and Maiti et al. (2007) used the neural network to predict lithofacies based on down-hole physical properties measurements. In these works a quantitative relationship between numerical well log data and core description is simulated by the neural network, then the network is applied to the uncored wells or parts of the well lacking the core to predict lithofacies. The application of a neural network was introduced to the Ocean Drilling Program by Benaouda et al. (1999) and Ojha and Maiti (2013). They used the neural network to classify down-hole physical properties measurement to predict lithology where there is partial or zero core recovery along the hole. All of the above mentioned works have been applied in the sedimentary environment, and the prediction was considered fairly successful. Application of this method to the igneous and metamorphic environments like the Victoria property might be complicated as they are more complex than sedimentary environments due to the structural metamorphic and intrusive history.

In this research, the main objective is to compare the reliability of representing the physical properties measurements in the form of rock types and physical log units. As a first step, a neural network trained from physical properties and geological logging information collected in hole FNX1168 was used to predict the rock type in hole FNX1182. If the network was trained on the data from more than one hole, better results might be obtained, but our purpose is to test the efficacy of neural networks early in the exploration process when less data are available. The predicted rock types in FNX1182 are compared with actual rock types logged by geologists to evaluate the accuracy of the classification in the context of an igneous and metamorphic environment. With the physical properties measurements available to us, we demonstrate below that this is only moderately successful. Then rather than using the geologists' classification of rock units for training, we used physical log units in FNX1168 determined by fuzzy k-means clustering (Mahmoodi and Smith, 2015). The trained neural network was then used to predict physical units in FNX1182, with greater success. Considering the relationship between physical units and rock types is known for FNX1168, the predicted physical units of FNX1182 can be interpreted from a geological point of view. Finally, we discuss whether to represent down-hole physical properties measurements; in the form of rock type or physical log units.

2. Methods

The neural network is structurally comprised of an input layer, at least one hidden layer, and an output layer. Each layer has different numbers of neurons. Neurons of two adjacent layers are connected one-by-one by synaptic weights. The schematic structure of a simple three-layer neural network is shown in Fig. 1. The input data are represented as a data vector to the input layer. The number of neurons in the input layer equals the number of variables measured at each depth (in our case three, gamma-ray response, density and magnetic susceptibility). The number of output neurons is determined by the number of elements of the target vector; which is the number of classes in the classification problems (in our case seven rock types). The hidden neurons are computing elements of the network, which use transfer functions to generate the output. The input of a hidden neuron is a summation of bias and neurons in the previous layer multiplied by corresponding synaptic weights. Biases allow the activation function to shift to the right or left to give a desired output (Bishop, 1995; Duda, 2001; Theodoridis and Koutroumbas, 2003 and Beale et al., 2001). In multilayer networks, the sigmoid function shown in Fig. 1 is often used as the transfer function (Beale et al., 2001). There is no rigorous theoretical method available to choose the number of hidden layers and hidden neurons, and they are often determined subjectively based on trial and error. The most compact structure with acceptable performance is preferred for computational efficiency. Most classification problems can be solved by the network when one hidden layer is used. With increase in the complexity of the relationship between input and desired output on the training data, the number of the hidden neurons should increase (Lawrence et al., 1996; Wong et al., 1995 and Beale et al., 2001).

The initial weights are randomly assigned to start the training process. The main task in network training is to adjust the weights to minimize the error of the network. The error function for each iteration is a form of the difference between the actual network output and the desired or target output. Here, mean square error (mse) is described as the error function:

$$J = mse = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \tag{1}$$

where t and a represent desired and network output respectively for each neuron of the output layer for N training data. In our case N is the number of depths that physical properties are measured at. A gradient descent is an optimization method which is used to adjust the weights in the direction through which the most rapid decrease in error function is achieved. The adjusting term ($\Delta\omega_j^r$), which is added to the j^{th} weight of the r^{th} layer weight after iteration, is obtained by the derivative (gradient) of the error function with respect to the weight. To enhance optimization efficiency, this value is multiplied by a learning rate (lr). The learning rate which is greater than zero and smaller than or equal to 1 controls the speed of the convergence process and how much the weights and biases can be modified at each iteration. The new estimate of the weight ${\omega_j^r}_{(new)}$ is described as:

$$\omega^r_{j(new)} = \omega^r_{j(old)} - lr.\Delta\omega^r_j, \quad \Delta\omega^r_j = \frac{\partial J}{\partial \omega^r_j}, \tag{2}$$

where ${\omega_j^r}_{(old)}$ is the current weight, $\Delta\omega_j^r$ is the adjusting term, and lr is learning rate. The network with adjusted weights generates a new output set, and the process iterates (each iteration referred to as one epoch) until the termination criteria is fulfilled. Different criteria have been suggested to terminate the iterations, such as a threshold for the minimum performance function (cost function), the minimum decrease in the cost function in successive iterations, and the number of validation checks, which is the number of successive iterations that the cost function fails to decrease (Bishop, 1995; Benaouda et al., 1999; Theodoridis and Koutroumbas, 2003 and Beale et al., 2001).

The training data set must be sufficiently large to provide enough data to train and test the network. The data set is typically randomly divided into three parts for training, validation, and testing, constituting 70, 15, and 15% of the available data, respectively. Testing data are used to assess generalization of the trained network which determines capability of the network to be efficiently applied to new data. If the network keeps iterating and adjusting the parameters to minimize the error, the network starts to over-fit the training data. Over-fitting occurs when the network, irrespective of generalization, tries to minimize the error of training data. In this case, the impact of random noise is incorporated into the network weights; however, the error in testing and validation rises. Validation data is used to assure that the division of data is appropriate. If the errors of testing and validation data are significantly different, it indicates poor data division (Bishop, 1995; Benaouda et al.,