USE OF POLYNOMIAL NEURAL NETWORK FOR A MINERAL PROSPECTIVITY ANALYSIS IN A GIS ENVIRONMENT

V. Iyer\textsuperscript{1}, C. C. Fung\textsuperscript{1}, W. Brown\textsuperscript{2} and T. Gedeon\textsuperscript{3}

\textsuperscript{1}School of Information Technology, Murdoch University of Technology, Perth, Australia
\textsuperscript{2}Department of Geology, University of Western Australia, Perth, Australia
\textsuperscript{3}Department of Computer Science, Australian National University, Canberra, Australia
E-Mail: \{viyer|L.Fung\}@murdoch.edu.au

ABSTRACT

In the mining industry, identifying new geographic locations that are favourable for mineral exploration is very important. However, definitive prediction of such locations is not an easy task. In the recent years, artificial neural networks have received much attention in this area. This paper uses a class of neural networks known as the Polynomial Neural Network (PNN) to construct a model to correctly classify given location into deposit and barren areas. This model uses the Geographic Information Systems (GIS) data of the location. The method is tested on the GIS data for the Kalgoorlie region of Western Australia.

Keywords:
Polynomial neural network; Mineral prospectivity; Geographical Information System.

1. INTRODUCTION

The primary objective of research in mining industry is aimed towards finding new mineral deposit locations. A definite method to do this is by drilling the location of interest and to carry out extensive analysis of the composition of the retrieved samples. However this is a costly procedure as this depends on trial and error. The other step involves developing new methods that can predict with reasonable accuracy, locations where new mineral deposits can be found. During the late 1980s, with the advancement in computing software and hardware capacity, the commercial geographic information systems, that could handle large spatial data, became available. GIS is defined [2], as having functional capability to bring together spatial data from a large variety of sources into a single data base as a series of data layers that overlap correctly at all locations. Due to these features they are emerging as an important technology in mineral exploration. A typical application of GIS in the geological communication is for mineral potential mapping.

Some of the mineral prospectivity methods that use geosciences data sets include Boolean algebra [8] and the index overlay method [12, 2, 15]. While the binary and index overlay methods are simple, they have disadvantages such limited information in the map and not being suitable to model complex non-linear relationships. Statistical methods such as the multiple linear regression was one of the earliest methods used in the mineral prospectivity mapping [6, 7, 18]. But the method is based on assumptions that the relationship between the input and output variables is linear. It is obvious that in most cases, such assumption is not true.

Artificial neural networks (ANN) have been extensively used in other fields of research but are still not used exhaustively in the area of mineral exploration. Until recently, backpropagation neural network (BPNN) has made up for a majority of the neural network applications. However, there are inherent issues in using this approach. Lately, further studies are involved in investigating the performance of other neural networks. This paper is pointing to such a direction in deploying alternative ANN architecture to the problem of mineral prospectivity analysis.


In this paper a class of neural network architecture called the polynomial neural network (PNN) is used to predict deposit and barren cells. In recent years, many papers have reported using PNN architecture [1, 5, 11, 13, 16, 17] for a variety of applications with good results. But none of them are in the mineral exploration research.
2. Polynomial Neural network

PNN is a flexible neural architecture whose topology is not predetermined but developed through learning. The design is based on Group Method of Data Handling (GMDH) which was invented by Prof. A. G. Ivankhnenko in the late 1960s [10] but enhanced by others. He developed GMDH as a means for identifying nonlinear relations between input and output variables. As described in [14] the GMDH generates successive layers with complex links that are individual terms of a polynomial equation.

The individual terms generated in the layers are partial descriptions of data (PDs) being the quadratic regression polynomials with two inputs. The first layer is created by computing regressions of the input variables and choosing the best ones for survival. For example, if the first two variables a and b are taken and combined into a simple set of polynomial terms the terms would be \((1, a, b, ab)\). Next, all possible models made from these terms are checked and the best one that satisfies an evaluation criterion is retained. The second layer is created by computing regressions of the values in the previous layer along with the input variables and retaining the best candidates. More layers are built until the network stops getting better based on termination criteria. The selection criterion used in this study penalizes the model that become too complex to prevent overtraining.

3. GIS data set

The GIS database is viewed as a collection of maps of a particular data type (such as solid geology, regional-scale faults etc) for a common geographic coordinate system. Within the map layers, two data structures are used to represent the spatial objects namely the vector and raster structures. Most GIS supports both structures and allow conversion from one structure to another.

For this study, GIS data set used is the one described in [3] to examine the prospectivity of orogenic lode-gold deposits in an approximately 100 x 100 km area of the Archean Yilgarn Block, near Kalgoorlie region of Western Australia. In this study 10 GIS layers in the raster data format are used to create the feature vectors. The GIS layers correspond...
to information such as the solid geology, magnetic anomaly, gamma-ray survey and distance to faults as shown in Figure 2. The thematic layers are divided into a grid of square cells of 100m side. Each cell is represented by the cell position and set of attributes within the two dimensional matrix of cells. The map area thus results in 1,254,000 cells. Out of these only the 120 deposit cells with a total contained gold source > 1000 kg along with 148 barren or non-deposit cells are used as training and test data set for the polynomial neural networks.

4. Results

Out of the data for the 248 cells, 147 cells were used for training and 81 cells for testing. Just like the training data set, the test data set has both deposit and barren cells. All the input values are scaled to [0, 1]. Table 1 shows the number of patterns in the test and training data set.

Table 1. Number of patterns in the training and test data sets

<table>
<thead>
<tr>
<th>Training Data Set</th>
<th>Test Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposit</td>
<td>Barren</td>
</tr>
<tr>
<td>85</td>
<td>102</td>
</tr>
</tbody>
</table>

A number of neural nets were trained with 10-input single output polynomial neural networks with different model complexity. Among them the one which gives the best training set results is tested with the independent test set of 81 patterns. The resulting trained network has 14 layers.

The output values ranged from 0 to 1. These output values were classified as representing barren or deposit cells by employing different threshold or cut-off probability values. The outputs were calculated with different threshold ranging from 0.1 to 0.9 in steps of 0.1. The results are given in Table 2. For comparison purpose, a back propagation neural network (BPNN) was trained and tested with the same data set. The results for the same are given in Table 3.

From the results given above, it is evident that the BPNN method is better than the polynomial neural network method in the case of training set data. But the performance of the network is best evaluated by testing on an independent test data set. Based on that polynomial neural network test set results are better than the BPNN results. For instance for a cutoff threshold value of 0.5, the polynomial neural network results are better at 80% and 71.7% correct compared to 77.1% and 65.2% correct respectively in the case of BPNN.

5. Conclusions

The paper proposes the use of polynomial neural network (PNN) model to identify deposit and barren locations for mineral exploration. The method is tested using the GIS data of a location in Western Australia. The main advantage of PNN is its learning capability in determining the number of nodes and hidden layers. In the case of BPNN, this has to be resolved to trial and error. PNN also has an advantage in terms of a shorter training time. In particular, the results from PNN have found to be better than that from BPNN. Further investigation is currently undertaken to explore other alternate neural networks for the purpose of prospectivity analysis.

ACKNOWLEDGEMENTS

This research work is supported by an Australian Research Council SPIRT Grant.

REFERENCES


