

Article

Prospectivity Mapping of Mineral Deposits in Northern Norway Using Radial Basis Function Neural Networks

Cyril Juliani *  and Steinar L. Ellefmo 

Department of Geoscience and Petroleum, Norwegian University of Science and technology (NTNU), Sem, Sælands vei 1. N-7491 Trondheim, Norway; steinar.ellefmo@ntnu.no

* Correspondence: cyril.juliani@ntnu.no; Tel.: +33-601-766-722

Received: 18 January 2019; Accepted: 20 February 2019; Published: 24 February 2019



Abstract: In this paper, the radial basis function neural network (RBFNN) is used to generate a prospectivity map for undiscovered copper-rich (Cu) deposits in the Finnmark region, northern Norway. To generate the input data for RBFNN, geological and geophysical data, including up to 86 known mineral occurrences hosted in mafic host-rocks, were combined at different resolutions. Mineral occurrences were integrated into “deposit” and “non-deposit” training sets. Running RBFNN on different input vectors, with a k-fold cross-validation method, showed that increasing the number of iterations and radial basis functions resulted in: (1) a reduction of training mean squared error (MSE) down to 0.1, depending on the grid resolution, and (2) reaching correct classification rates of 0.9 and 0.6 for training and validation, respectively. The latter depends on: (1) the selection of “non-deposit” training data throughout the study area, (2) the scale at which data was acquired, and (3) the dissimilarity of input vectors. The “deposit” input data were correctly identified by the trained model (up to 83%) after proceeding to classification of non-training data. Up to 885 km² of the Finnmark region studied is favorable for Cu mineralization based on the resulting mineral prospectivity map. The prospectivity map can be used as a reconnaissance guide for future detailed ground surveys.

Keywords: mineral resource; prospectivity; mapping; radial basis function; artificial neural networks; Norway

1. Introduction

Data-driven mineral exploration is a cost-effective alternative to exploration by drilling in the search for undiscovered mineral deposits. It requires the analysis of collections of spatial datasets including e.g., regional geology, geochemical and structural data, and airborne magnetics from a region of interest. The analysis can be made with the help of a Geographical Information System (GIS) and geo-computational techniques whereby a set of anomaly data are compared locally with known mineral occurrences. Regionally, this comparison allows us to determine a set of coinciding geoscience data favorable for mineralization, which can be used to highlight the mineral potential of other less explored regions, with similar data attributes. As these attributes are a direct response to some unusual data variations, distinct likelihoods of the probable presence of new findings can be estimated (favorability). The aim of this research is to demonstrate how these likelihoods can be determined for undiscovered Cu-rich mineral deposits (i.e., metallic ore occurrences with Cu as the major component) within the Finnmark region, northern Norway, using a radial basis function neural network (RBFNN), a type of artificial neural network (ANN) that can be employed for mineral prospectivity mapping (MPM) (e.g., [1–4]). MPM is a process whereby a set of geoscientific data

(e.g., magnetic and geochemical anomalies, structural data and regional geology) are combined to produce a map which ranks areas according to their potential to host a deposit(s) of particular type(s) (e.g., [3–14]). The ranking derives from the combining of layers of geoscientific data with the help of GIS, which is usually based on some implicit weighting scheme that attempts to discover or learn to assign appropriate weights to the different data layers (see e.g., [7,8,12]). The learned weights determine the importance of the layers for particular data attributes of mineral deposits or, in other terms, a degree of favorability of an area for mineralization. The favorability, sometimes interpreted as a probability (e.g., [12]), is thus, a quantitative measure with low-to-high range of values that describes the likelihood that some areas contain mineral deposit(s).

RBFFNN calculates favorability by interrogating and synthesizing spatial data through a network of interconnected computational units or mathematical functions [15,16]. The network maps each input feature vector of a training set to its output target vector i.e., a vector of predicted conditions, being either barren or mineralized for the deposit type(s) considered. This way, the network can be trained to indicate the degree of similarity of unknown input vector (unseen during training) to a composite of known deposit vectors (one of each deposit type) used during training. As a result, RBFFNN deduces knowledge by learning from samples of training data, and then, uses this knowledge for generalization beyond the training data. Both, the concept and application of RBFFNN will be demonstrated in this study for base-metal prospectivity in Finnmark. A mineral prospectivity map will be produced to suggest where future exploration surveys could be conducted.

2. Radial Basis Function Neural Networks (RBFFNN)

ANNs combine and weight information through computational units (the neurons) in a network consisting of input, hidden and output layers (Figure 1). Input signals ($x_0, x_1, x_2, \dots, x_n$), that must be classified, are connected to and weighted by hidden neurons through “synaptic weights” (or “weights”), and the neurons’ responses are called their “activation” values (see [17] for details). These values are calculated through non-linear activation functions by taking the weighted sum of their input and adding a bias:

$$y = \sum(\text{weight} * \text{input}) + \text{bias} \quad (1)$$

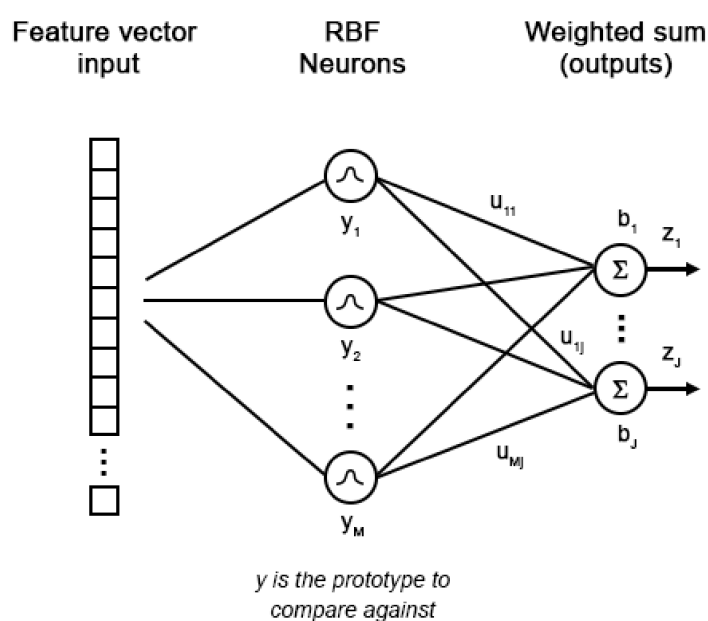


Figure 1. Architecture of a standard radial basis function neural network.

The set of weights describes the strength of connections between processing units among different layers of the network. A weight from unit A to unit B, that has greater magnitude, means that A has greater influence over the behavior of B, i.e., in increasing or decreasing the level of activation of B. ANNs are trained to optimize these weights so that input signals correspond to some degree to output predictions. The back-propagation method (see [18]) is commonly employed for tuning the synaptic weights in ANNs to minimize the network output error on training examples. The error can be expressed as, e.g., the mean squared error (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2, \quad (2)$$

where n is the number of predictions, y_i the observed (target) value and \hat{y}_i the predicted value returned by the network. Here, the computer tries to increase or decrease each synaptic weight to see how it reduces MSE, with the purpose of reaching the lowest MSE given an optimal combination of synaptic weights. As the network outputs \hat{y} depend on each individual weight in the input and/or hidden layer(s) of the ANN, the weights dictate the relevance of the input variables. The weights can be either be positive or negative. The increase or decrease of parameters value (weights and bias) is decided given the error gradient E , which is calculated for every parameter [17]:

$$\nabla E = \left(\frac{\partial E}{\partial w_{11}}, \dots, \frac{\partial E}{\partial w_{21}}, \dots, \frac{\partial E}{\partial w_{jk}}, \frac{\partial E}{\partial b} \right), \quad (3)$$

with w_{jk} , a given synaptic layer-to-layer weight considering that the two layers in question have j and k neurons, respectively, and b , the related bias. For each parameter, the derivative of E can be either positive or negative with respect to the error function (which increases or decreases, respectively); only the negative derivative is considered to minimize the error E . A weight can be updated as follow (the same is applied for b):

$$w'_{jk} \leftarrow w_{jk} - \alpha \frac{\partial E}{\partial w_{jk}}, \quad (4)$$

with w_{jk} , the former synaptic weight from node j to node k , $\frac{\partial E}{\partial w_{jk}}$, the weight increment, and α , the learning rate, i.e., a parameter controlling how much we are adjusting the weights of our network in order to minimize the network's error signal. The ability of the trained network to generalize to the prediction of new outputs is finally tested by applying the network to a set of test examples which had not been seen during training [17].

A three-layer ANN model, which consists of one hidden layer, processes data through a number of neurons by which mathematical functions map an input vector to respective output target vector. In a RBFNN, neurons of the hidden layer use the radial basis function (RBF) to process input data [15,16]. RBF measures the average distance between all points of a N -dimensional feature vector x , and a center position v corresponding to a prototype vector taken from the training set. A width (or spread) parameter σ normalizes this distance measure, i.e., it determines the receptive field where all input data x equidistant from v yield the same value y [15,16]. Given the output from the m th node of the hidden layer, RBF can be written as follow [16]:

$$y_m = \exp\left(-x - v_m^2 / (2\sigma_m^2)\right) \quad (5)$$

and,

$$\sigma_m = \frac{1}{n} \sum_{i=1}^n x_i - v_m \quad (6)$$

Here, as the distance between the input and prototype grows, the response y_m falls off exponentially towards 0. The RBF neuron's response y_m generates a bell curve (Gaussian shape)

for which the height is controlled by the coefficient $\frac{1}{2\sigma^2}$. The double bar notation indicates that we are taking the Euclidean distance between both vectors x and v_m , and squaring the result. The Euclidean distance between the two vectors allows to evaluate the similarity between an input vector and a prototype. Each RBF neuron will produce its largest response when the input vector corresponds to the prototype vector. Otherwise, neurons whose prototypes are far from the input vector in the Euclidean space will actually contribute very little to the prediction results. The weighted sum of outputs from RBF neurons corresponds to an output prediction. The output prediction from the j th node of the output layer is [15,16]:

$$z_j = \left(\frac{1}{M} \right) \sum_{m=1}^M u_{mj} y_m + b_j, \quad (7)$$

with m , the number of hidden nodes, u_{mj} , the synaptic weight and b_j , the bias. Training RBFNN involves determining (1) the number of RBFs (hidden neurons), (2) the center and spread parameters of RBFs, and (3) the synaptic weights for classification of all training vectors. During training, the output nodes will learn the weights u_{mj} that minimize the network's prediction error E . In the network, each RBF will have some influence over the classification decision given by z_j , according to the weights u_{mj} .

3. Regional Geology and Mineralization

The study area, located within the Finnmark region, northern Norway, covers an area of about 44 thousand km² and extending for some about 400 km from west to east (Figure 2). Associated bedrock geology (see [19] for an overview of regional geology) consists of (1) autochthonous Archean gneiss complexes (Jergul and Ráiseatnu) in addition to early Paleoproterozoic tectonic belts, the Kautokeino and Karasjok Greenstone Belts (KkGB and KjGB, respectively), overlain by (2) a succession of Neoproterozoic to Cambrian nappe complexes mostly made of metasedimentary units [20,21], some of which were intruded by numerous igneous mafic rocks (e.g., at the Seiland Igneous Province) [22–25] in the Caledonides (Figure 2). The greenstone belts expose tectonically reworked sedimentary-volcanic successions made mainly of mafic volcanic rocks, and clastic sedimentary rocks and plutonic complexes formed between 2.5 and 1.8 Ga [26–29]. KjGB is an east-dipping tectonic wedge unconformably bounded by the Archean Jergul Gneiss Complex [30]. It hosts some important deposits, including placer gold and komatiites in volcano-sedimentary rocks, in addition to Ni-Cu-PGE deposits in ultramafic and mafic intrusions; several gold-hosted lithologies are also associated to copper deposits [31]. KkGB, on the other hand, consists of supracrustal metavolcanic rocks varying in composition from tholeiitic to komatiitic, and clastic metasedimentary units deposited during an early Paleoproterozoic rifting [32]; the area is known to contain widespread gold-copper and copper mineralization [31].

Table 1. Metallic ore deposits in the studied Finnmark region (according to NGU's classification). Cu: copper; Zn: zinc; Pb: lead; Fe: iron; As: arsenic; Sb: antimony; Bi: bismuth; Sn: tin; Mn: manganese; Ti: titanium; Co: cobalt; Cr: chrome; Ni: nickel; V: vanadium; Mo: molybdenum; W: tungsten; U: uranium; Th: thorium; Au: gold; Ag: silver; PGE: platinum group elements.

Metallic Ore Category	Total in Area	Main Ore Component
Base metals (Cu, Zn, Pb, Fe-sulfide, As, Sb, Bi, Sn)	231	Cu (218), Pb (1), Zn (6), Sulfides (6)
Ferrous metals (Fe, Mn, Ti)	94	Fe (84), Mn (9), Ti (1)
Iron Alloy Metals (Cr, Ni, Co, V, Mo, W)	17	Co (3), Cr (1), Mo (1), Ni (12)
Energy metals (U, Th)	2	U (2)
Noble Metals (Au, Ag, PGE)	13	Au (13)

Several tectonic windows in the Caledonian part of Finnmark represent apparent continuations of the Archean–Paleoproterozoic basement beneath the Caledonian nappes (e.g., [33]) The Alta–Kvaenangen Tectonic Window (AKTW) (Figure 2), for example, contains diverse metamorphosed sedimentary rocks associated with abundant mafic extrusive and intrusive rocks dated 2–1.9 Ga [34–36].

Sedimentary and volcanic rocks in AKTW constitute the Raipas Supergroup, which includes the Kenvik formation [34], a formation mainly composed of gabbro and mafic tuff and tuffite, but also massive and pillowed tholeiitic basalt [36]. Another example is the Repparfjord Tectonic Window (RTW), which is of Early Palaeoproterozoic age and consists predominantly of metavolcanic and metasedimentary rocks intruded by mafic and ultramafic intrusive rocks [37]. RTW can be subdivided in two groups, (1) the Holmvatn Group, made up by a sequence of various metalavas and tuffites of calc-alkaline affinity, and (2) the Nussir Group, consisting of tholeiitic metavolcanites deposited within a submarine rift [27,33,38]. Both, the AKTW and RTW are largely similar with respect to lithology, stratigraphy and age and, in addition, they host several volcanic- and sediment-hosted Cu deposits with epigenetic and syngenetic origin (e.g., [31,39]).

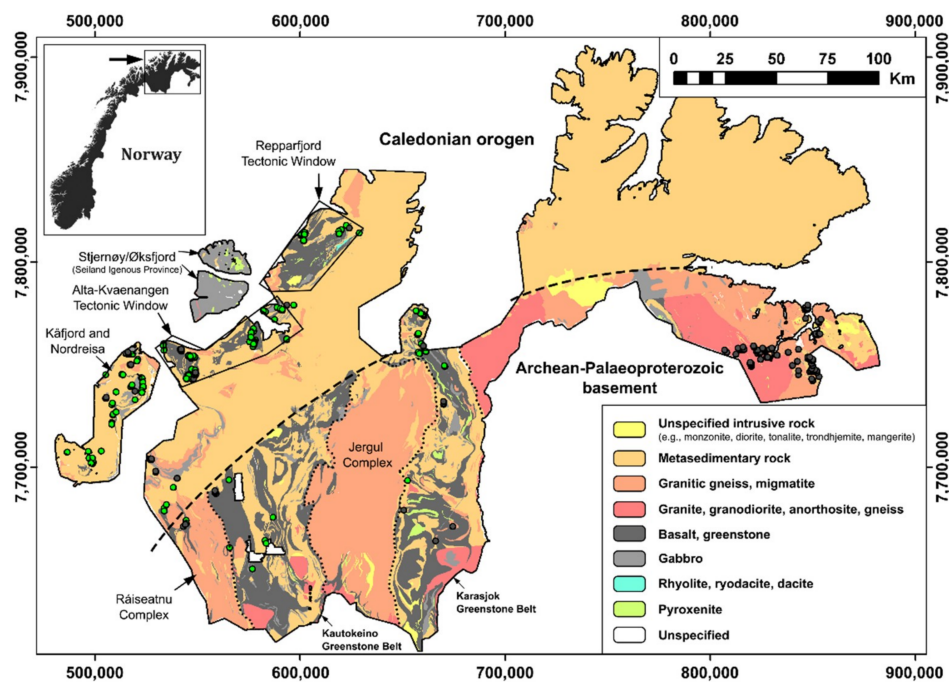


Figure 2. Geology of the studied area in Finnmark. The location of known Cu-rich deposits and other types of deposits are indicated in green and dark grey circles, respectively (see Table 1 for details). The dashed black lines delimit the northern Caledonian orogen from the southern Archean schist and gneiss complexes and Palaeoproterozoic sedimentary, volcanic and plutonic rocks (based on [19]). The dotted lines delineate boundaries between the greenstone belts and the Archean gneiss complexes. UTM zone 34, European Datum 1950.

Not only copper mineralization, but also several iron, chrome-nickel, uranium and gold deposits make up the metallogeny of the Finnmark region (see [40]). For predictive mapping of favorable Cu mineralization, the present study involves an analysis of GIS data of Finnmark, which was acquired and distributed by the Geological Survey of Norway (NGU) (Datasets available online: <http://www.ngu.no/en/topic/datasets>). These data include (1) a geological map of the subject area at a scale of 1:250,000 (see [16] for details), (2) a compilation of airborne magnetic and radiometric data (U, K and Th) acquired between 1979 and 2015 (see [41–43] and references therein) (Figure 3) with diverse line spacing and flight altitudes (mostly 200 m spacing and 60 m altitude; see Figure A1), (3) the regional gravity field acquired based on measurements at gravity stations established by NGU with a minimum spacing of 800 m (Figure 4) (see [44]) and (4) the coordinates of known mineral deposits (Table 1; Figure 2), notably several sulfidic Cu-rich occurrences, which are mainly hosted by mafic rocks. These occurrences were used to train the classifier, although no distinction was made among the different types of ore deposits (such as, e.g., porphyry or volcanogenic massive sulfide deposits)

possibly associated with mafic host-rocks. Here, given the size of the training set used (see Section 4), we assume to develop a composite model that combines the statistical attributes of different deposit types. A normal modeling procedure would establish separate prediction models for the deposit types considered, but this may require diverse and a fairly large number of representative training data.

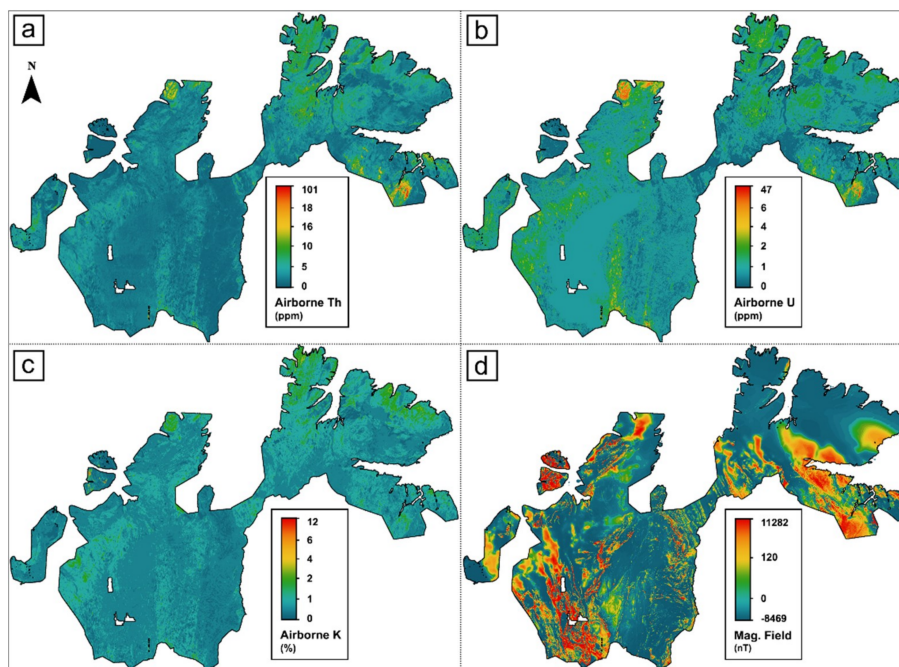


Figure 3. Maps amplitude of (a) thorium (Th), (b) uranium (U), (c) potassium (K) and (d) total magnetic field.

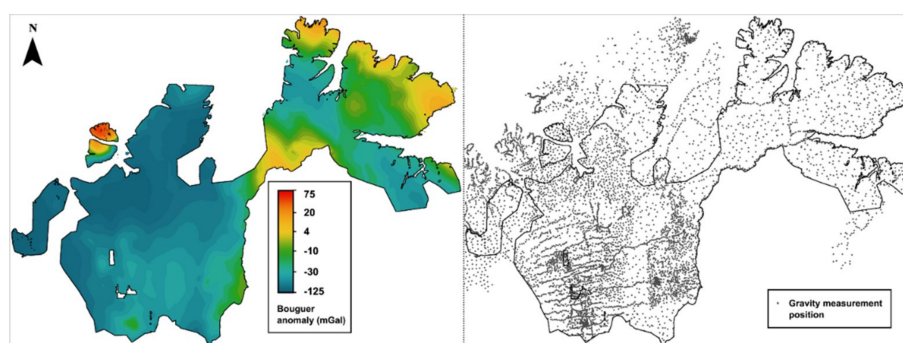


Figure 4. Bouguer anomaly map of the study area (left) based on measurements from gravity stations (right). The magnitude-per-unit area from point was calculated using the Kernel Density method. Mean distance between measurement positions is 1.5 km.

4. Data Processing

Data pre-processing and integration in GIS were carried out using ArcGIS 10.3 (Esri, Redlands, CA, USA). Each data layer has been stored in regular grid networks (square grid) of 250 and 500 m cell sizes. This required (1) to grid-cell average radiometric, gravity and magnetic records, and (2) to take the major local lithology for each cell. Note that the choice of grid resolution was motivated given (1) the line spacing of aeromagnetic and radiometric surveys (50–250 m) and (2) the problem evoked about averaging data at the grid resolution used later on (see Section 6.1 and Figure A3), which avoid obtaining too high (or too low) spatial variability in mineral favorability. The grids were overlaid to create a unique conditions grid where unique overlay conditions are considered as

n-dimensional (n = number of data layers) input vectors for RBFNN. Training RBFNN requires two sets of points: one that defines the presence of the condition to be predicted (i.e., the presence of a Cu-rich mineral occurrence(s)) and a second that defines the absence of this condition (i.e., locations where Cu mineralization is known to not occur); given the first and latter conditions, a particular input vector, representing a unique overlay condition (or data attributes) of the studied terrain, can be categorized as “deposit” or “non-deposit”, respectively, although the notion of “non-deposit” has no geological meaning.

Depending on grid resolution, 86–115 “deposit” input vectors are associated to mafic host-rocks in the Finnmark region analyzed. This is little compared with the 177,077–690,953 remaining vectors categorized “non-deposit”. For this reason, the size of the training set was limited by the number of known deposits to avoid preference learning by the model over the more numerous “non-deposit” vectors representations. In this regard, selecting truly barren grid cells is a difficult exercise because one might not certify whether a particular location without known deposit(s) contains or not the deposit type sought. A suitable solution is to consider locations with mineral deposits that are (1) not of interest for prospectivity (i.e., not Cu-rich; see Table 1), (2) nor consistent with favorable lithostratigraphic units (i.e., mafic rocks in this study) as barren grid cells. Correspondingly, a number of these grid cells, approximately equal to the number of deposit vectors, is selected homogeneously for each rock type (among those presented in Figure 2). Note that several training vectors (deposits or non-deposits) may have the same unique condition (i.e., unique combinations of attributes; see Figure A2) especially if e.g., most of the non-deposit grid cells are located in the same neighborhood; this can bias the training performance of the model. To handle this issue, the cosine similarity measure, $\cos(\theta)$, is applied among the deposit and non-deposit input vectors so that only one condition is kept for each training data. The cosine similarity is written as follows:

$$\cos(\theta) = \frac{v_1 \cdot v_2}{\|v_1\| \|v_2\|} = \frac{\sum_{i=1}^n v_{1i} v_{2i}}{\sqrt{\sum_{i=1}^n v_{1i}^2} \sqrt{\sum_{i=1}^n v_{2i}^2}} \quad (8)$$

where v_{1i} and v_{2i} are components of vectors v_1 and v_2 , respectively. After analysis, it was decided to consider training data with an average dissimilarity value higher than 0.3. Resulting number of deposit input vectors was refined to 67–87.

A common procedure to estimate a model’s generalization error is to subdivide the training set into training, validation and testing data sets; the latter is used to evaluate the classifier (see Section 5) while the validation set evaluates the network prediction performance during training to avoid the effect of preferential sampling (see “early stopping” in Section 5.1). Given the modest size of the training set used in this work (134–174 vectors), only the training and validation sets are used; they represent 70% and 30% of the original training set, respectively, assuming that vectors labeled “deposit” and “non-deposit” are homogeneously distributed between the two data sets. To obtain the lowest generalization error, a k-fold cross-validation method (see e.g., [10]) was applied. This method subdivides the training set into k subsets of approximately equal size. Then, the network is trained k times, each time leaving out one of the subsets from training, which in turn is used to monitor (validate) the network parameters and the prediction error.

5. Network Implementation

The network was implemented using Python and Tensorflow, a library that can be used to create models released by Google. The implementation required three main steps: (1) developing the RBFNN architecture with a number of RBFs, (2) training the network with an adequate learning rate, and (3) evaluating the predictive performance of the model with appropriate metrics. Prior to being sent as inputs to the network, non-lithologic data are normalized by standard deviation to make all units roughly comparable to each other. Output variable of the analysis is expressed in binary form with respect to the “deposit” and “non-deposit” conditions.

5.1. Architecture and Training Phase

Selecting a number of RBF is critical as it affects the quality of predictions (e.g., [4,13]). The predictive accuracy and cost function error of the network were estimated iteratively using various numbers of RBFs. Related prototypes were randomly selected from the training set by assuming an equal number of vectors labeled “deposit” and “non-deposit” for training. The training was done using the back-propagation method for the error optimization task, considering a learning rate of 0.5. The learning rate value was decided based on MSE, which was calculated after successive training tests. Training was stopped after a number of training cycles when the minimum error is reached during validation; a cycle of training is reached once all the training data have been used by the network.

For the classification task, the softmax activation function [45] was used to calculate the probability σ of an output z_j (see Section 2) to belong to a given class, either “deposit” or “non-deposit”:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \text{ for } j = 1, \dots, K \quad (9)$$

The training procedure has been repeated multiple times to capture the highest average performance over samples taken from the training set. To avoid overfitting, the regularization term L2-norm, written $\frac{\lambda}{2} \|\vec{w}\|^2$, was added to the error function to maintain all the weights homogeneously small; λ defines how much the coefficients are penalized, i.e., if λ is large, these coefficients are penalized significantly [46]. Note that overfitting is the use of network models that do not generalize well on unseen data (validation or testing data). With an overfitted model, the calculated error on the training set can be very small, but when new data is presented to the network, the error is large. Standard techniques, other than regularization, exist for preventing overfitting such as “early stopping” [17], where training is stopped before overfitting begins to occur by monitoring the error rate estimate of the trained model on validation data. This approach was used in this study to determine when to stop training.

The entire dataset, including unseen data, was finally presented to the trained network to calculate the probability whether grid cell contains the deposit type sought or not. Note that according to the output variables selected for the study, namely “deposit” and “non-deposit”, grid cells with known Cu-rich deposits have a value of 1 while the non-deposit sites have value of 0. Thus, from the model’s perspective, we calculate the probability of having 1 or 0 for every grid cell. For simplicity, these probability values will be rescaled between -1 and 1 (i.e., an input vector is less and more likely to be categorized “deposit”, respectively), which are described as favorability values. It will be decided further on a minimum value of favorability (threshold) to categorize a grid cell as “deposit”, considering the estimated favorability of training and validation data.

5.2. Metrics

Evaluating the classification performances of an ANN requires informative metrics to test capabilities of classification algorithms in detecting true and false positives (e.g., the proportion of true and false deposit input vectors, respectively) and negatives (e.g., the proportion of true and false non-deposit input vectors, respectively) [47]. The receiver operating characteristic curve (ROC) [9] is an example of metric where the true positive rate (TPR) is plotted against the false positive rate (FPR) (Figure 5). Both are expressed as follows:

$$TPR = \frac{TP}{TP + FN} \quad (10)$$

$$FPR = \frac{FP}{FP + TN} \quad (11)$$

with TN, TP, FN, FP as true negatives, true positives, false negatives and false positives, respectively. The area under the ROC curve (AUC) is a common measure of neural network performance. AUC can have values from 0 to 1. The closer AUC is to 1, i.e., with 1 for TPR and 0 for FPR, the more accurate is

the classification (Figure 5). Prediction results with an AUC of 0, however, are completely incorrect. The lower limit of ROC curve area is 0.5; it is called the chance diagonal.

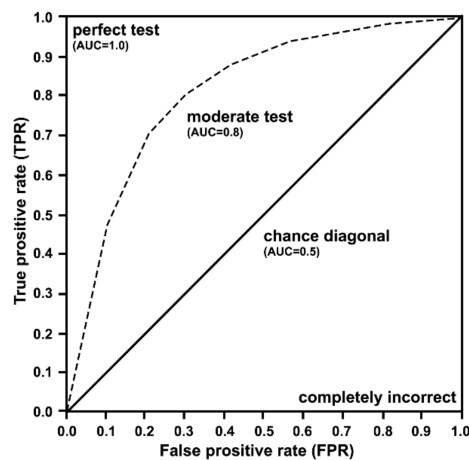


Figure 5. Receiver operating characteristic (ROC) curves with respective area under the curve (AUC) evaluating hypothetical data.

6. Results and Discussion

6.1. Training and Validation

As previously said, both the number of training cycles (iterations) and RBFs influence the training prediction results (e.g., [4,13]). These parameters must be adjusted, not only to select a model, but to avoid overfitting. As shown in Figure 6a, e.g., for a grid of 500 m resolution, MSE was optimized on the training and validation feature vectors by increasing the number of iterations; MSE reached 0.13 and 0.24, respectively, and remained at these levels after 50 iterations. AUC, on the other hand, stabilizes at 0.90 and 0.65 after 500 iterations for training and validation, respectively (Figure 6b). The range of output values for a model trained with 500 iterations is -0.73 – 0.73 . The simplest form of network that generated the smallest generalization error was with 90 radial basis functions (i.e., 0.15 and 0.20 for training and validation, respectively; Figure 7a). Although adding more hidden neurons increases the learning performance, it does not improve the prediction results (as shown by the validation curve on Figure 7b). A summary of training and validation performance for different grid resolutions is shown in Table 2.

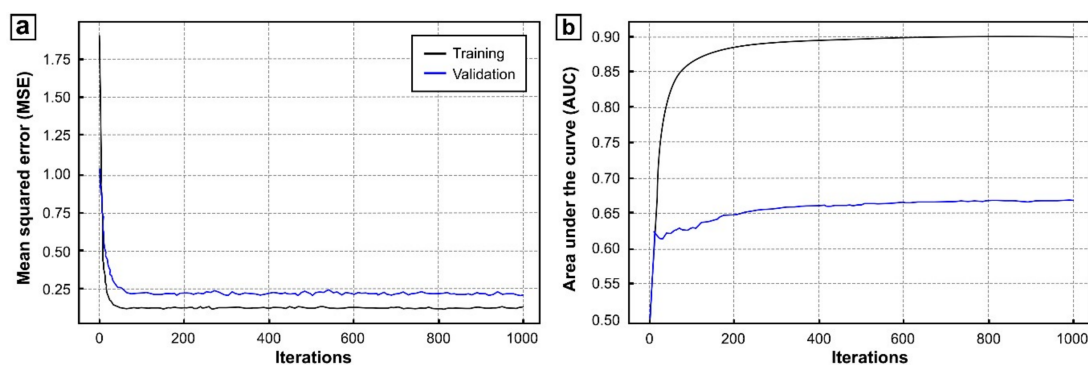


Figure 6. Example of training results generated for 500-m grid cells comparing (a) mean squared error and (b) area under the curve against number of iterations for a radial basis function neural network with 120 RBFs. The number of input vectors is 94 (70% of dataset) and 40 (30%) for training and validation, respectively.

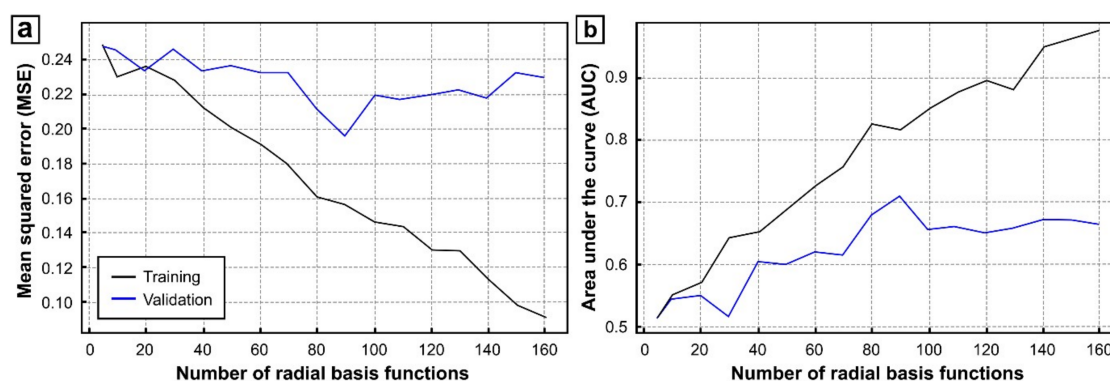


Figure 7. Example of training results generated for 500-m grid cells, comparing mean squared error (a) and area under the curve (b) against number of radial basis functions calculated for training runs with 500 iterations. The number of input vectors is 94 (70% of dataset) and 40 (30%) for training and validation, respectively.

Table 2. Summary of training and validation performances.

Resolution	Phase	Iterations	Number of Hidden Neurons	Mean Squared Error (MSE)	Area under the Curve (AUC) *
500 m	training	500	90	0.15	0.90
	validation			0.20	0.65
250 m	training	500	150	0.10	0.94
	validation			0.22	0.63

AUC values * calculated for different grid resolutions are quite similar to one another and generally close enough to 1. However, the model performance on validation data is lower (0.6), which diagnoses either potential overfitting or a model too complicated to generalize. The second option is preferred as several techniques were applied to reduce overfitting in this research (cross-validation, regularization, early stopping and cosine similarity; Sections 4 and 5.1) and because the difference of MSE between the training and validation sets (Table 2) is not large. Rather, peculiar attributes of input data were not depicted by the model and/or there is no sufficient dissimilarity between deposit and non-deposit input vectors. The latter can be inherent to e.g., (1) the use of small-sized grid cells compared to data with coarse spatial resolution, notably the measured regional gravity field derives from widely spaced gravity stations (Figure 4), (2) sampled deposit and non-deposit locations (Figure 2), and related multivariate spatial data signatures, which do not match as much the “true” population of target data (deposit and non-deposit vectors), and/or (3) with regard to the second point, the dataset used in this study (geoscience and/or target data) is quite limited for predictive modelling. In addition, data are averaged at the grid resolution used, which may decorrelate the real “ground-truth” (mineral) data with former related geophysical signals (see examples in Figure A3). Some modelling procedures, other than those applied in this study, may be applied to change the model’s predictive capabilities e.g., by selecting a coherent training set, i.e., deposit and non-deposit vectors having strongly similar characteristics (as shown by [48]), or using one deposit vector at a time in the learning procedure to identify similar deposit locations and assembling the prediction scores of the deposit locations into a two-dimensional prediction matrix (as shown by [49]); here, it can also be suggested to combine with other techniques limiting overfitting such as ensemble learners, (e.g., Random Forest).

6.2. Classification and Limitations

Figure 8 presents the predictive map of Cu mineralization using the classification results obtained from training on 500 iterations. Areas are considered with high potential for Cu mineralization where the favorability of such mineralization is equal to or higher than the favorability obtained for “deposit” training and validation data, which was evaluated to be 0.5 (threshold). On the predictive map, areas having the highest favorability of Cu mineralization are constrained where favorable lithologies (mafic rocks) are specified geographically (see Figure 2). In addition, there is no representational similarity between the prospectivity map calculated and non-lithological input layers (Figures 3 and 4). The absence of representational similarity suggests that none of the non-lithological data layers presents a major influence on the calculated favorability and, thus, the model behaves well within the range of study i.e., a generalization seems possible. These areas are located around most known Cu-rich occurrences (up to 83%), which were here adequately detected by the model (although some discrepancies can be noticed). Other areas, such as in zones A, B and C (Figure 8), indicate high favorability for Cu mineral resources, which suggest additional potential targets for further exploration.

Out of the 250 and 500 m resolution grids, 1.5–2% of the grid extent is favorable for Cu mineralization considering the above-mentioned threshold. Here, decreasing the cell size increases the overall extent of favorable area. An intuitive explanation is as more known deposits with unique conditions are used from the 250-m resolution grid for training the model (see Figure A3), the model learns more unique conditions (or data attributes) in return, which increases the prediction possibilities. It is worthwhile to note that the spatial extent of known deposits is not documented, and as such, the choice of grid resolution and the expected number of deposit(s) within a favorable grid cell, is debatable.

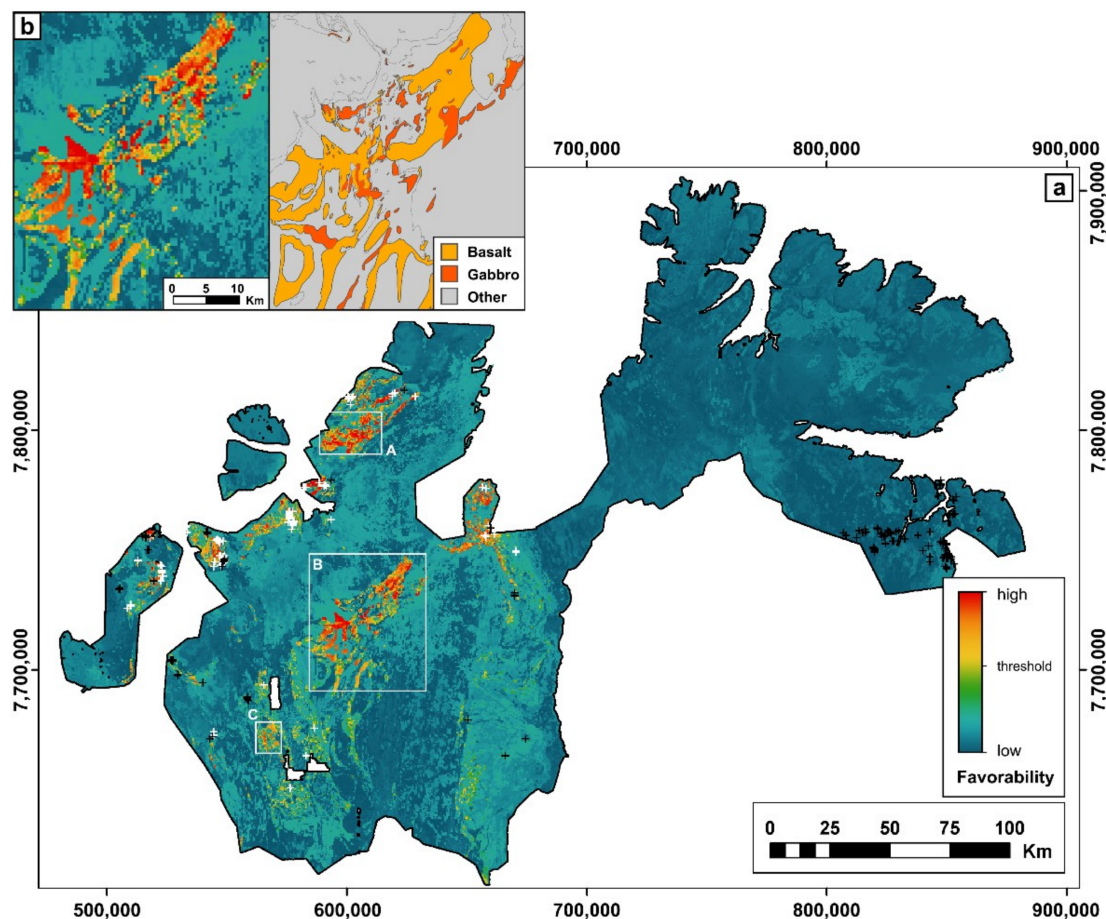


Figure 8. (a) Favorability measure of Cu-rich mineralization estimated from the radial basis function neural network for the Finnmark region (500-m cell resolution). Favorability values derive from the model’s prediction value range (−0.73–0.73). Shades of yellow, orange and red (equal or above the threshold) represent areas favorable for Cu mineralization. Locations of “deposit” and “non-deposit” sites, which were used as training and validation data, are shown by white and black crosses, respectively. Zones A (within the Repparfjord Tectonic Window), B and C (both within the Kautokeino Greenstone Belt) are examples of prospective zones unveiled by the model, which do not contain any local training or validation data. (b) Predictions in zone B compared with geology.

The area occupying up to 2% of the studied region (885 km²) corresponds to a high-potential zone for future mineral exploration. Note, however, that similarities can be identified between deposit and non-deposit training data (Figure A2), which brings redundant information to the training phase. This may decrease the model performance, and, for this reason, a number of target sites not regarded as favorable on Figure 8 could plausibly be of interest.

7. Conclusions

Initial results of this study demonstrated that RBFNN is a feasible approach for integrating geological variables when conducting mineral prospectivity mapping. The model selection was carried out using a k-fold cross-validation method and subsequent analysis of classification errors during validation. The end-product of predictions was a map depicting up to 885 km² of high favorability areas. Although spatial distribution of high favorability zones is consistent with locations, of known mineralized sites, the difference between training and validation AUC (0.9 and 0.6, respectively) suggests that training data vectors do not hold all information to distinguish most known Cu-rich deposits in the studied Finnmark region. Selecting non-deposit data prior to training, in particular, may contribute to bias the RBFNN analysis. For these reasons, additional data (e.g., structural data and geochemical soil anomalies) should be considered to improve prospectivity mapping and/or new data analysis techniques should help selecting adequate training data. Another result of this study is that the resolution of data analysis generates gain or loss of information, which influences the model performance. More research on data integration procedures and how to evaluate integrated data may be critical to solve such issues.

Author Contributions: C.J. developed the algorithms, applied the methodology, performed the analyses and wrote the article. S.L.E. reviewed the original draft.

Funding: This research was funded by the Norwegian University of Science and Technology (NTNU).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Aeromagnetic and radiometric surveys carried out in the study area.

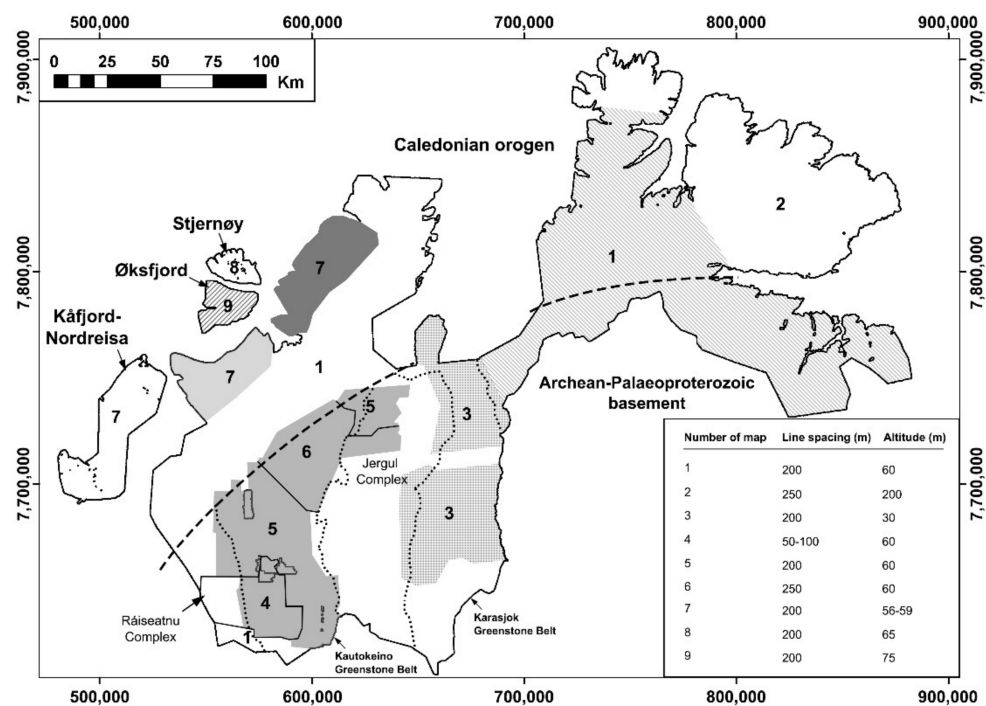


Figure A1. Aeromagnetic and radiometric survey data compilation (helicopter and fixed-wing surveys). See [41–43] and references therein.

Histograms of training data attributes: comparison between deposit and non-deposit vectors.

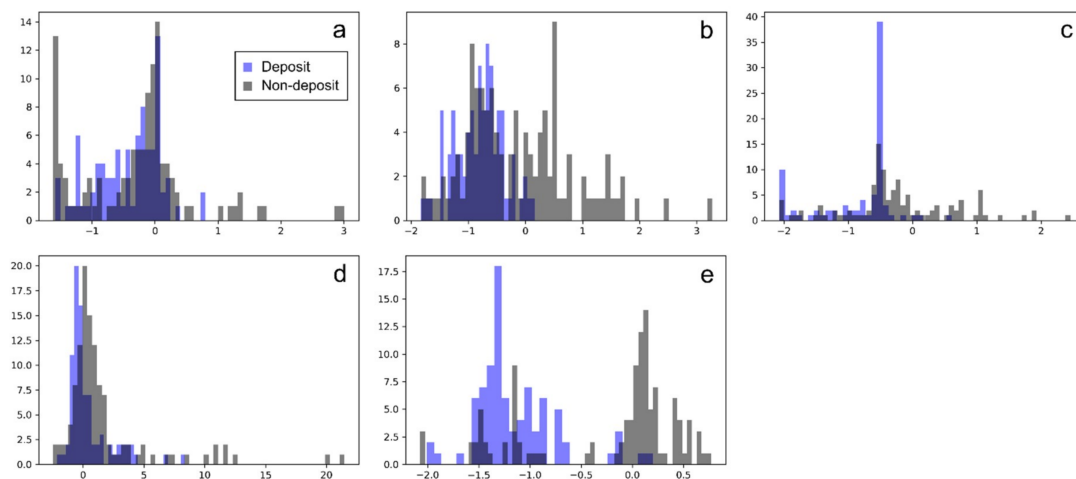


Figure A2. Histograms of (a) U, (b) Th, (c) K, (d) gravity and (e) magnetic anomaly data taken from the training set (deposit and non-deposit vectors). Data are normalized by standard deviation (x-axis). A noticeable dissimilarity is indicated by K and the magnetic records while there exist overlaps of data distributions for U and Th data.

Histograms of training data attributes for different grid resolutions.

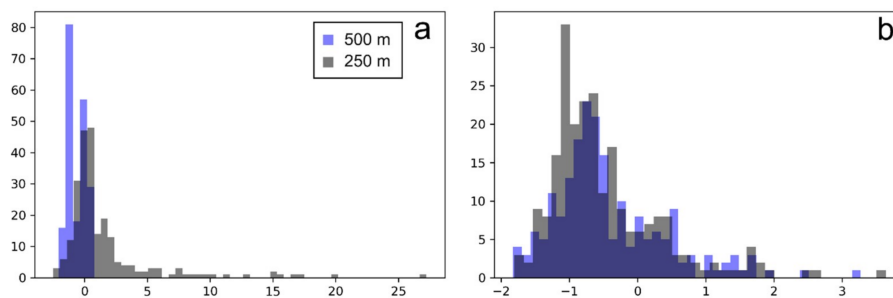


Figure A3. Histograms of (a) magnetic and (b) Th training data taken from grids at different resolutions. Data are normalized by standard deviation (x-axis). A loss of information is noticeable when changing the cell size from 250 to 500 m.

References

1. Porwal, A.K.; Carranza, E.J.M.; Hale, M. Artificial Neural Networks for Mineral-Potential Mapping: A Case Study from Aravalli Province, Western India. *Nat. Resour. Res.* **2003**, *12*, 155–171. [[CrossRef](#)]
2. Porwal, A.K. Mineral Potential Mapping with Mathematical Geological Models. Ph.D. Thesis, University of Utrecht, Utrecht, The Netherlands, 2006. [[CrossRef](#)]
3. Behnia, P. Application of Radial Basis Functional Link Networks to Exploration for Proterozoic Mineral Deposits in Central Iran. *Nat. Resour. Res.* **2007**, *16*, 147–155. [[CrossRef](#)]
4. Nykänen, V. Radial Basis Functional Link Nets Used as a Prospectivity Mapping Tool for Orogenic Gold Deposits Within the Central Lapland Greenstone Belt, Northern Fennoscandian Shield. *Nat. Resour. Res.* **2008**, *17*. [[CrossRef](#)]
5. Singer, D.A.; Kouda, R. Application of a Feedforward Neural Network in the Search for Kuroko Deposits in the Hokuroku District, Japan. *Math. Geol.* **1996**, *28*. [[CrossRef](#)]
6. Singer, D.A.; Kouda, R. Use of a Neural Network to Integrate Geoscience Information in the Classification of Mineral Deposits and Occurrences. Integrated Exploration Information Management. In Proceedings of the Exploration 97: Fourth Decennial International Conference on Mineral Exploration, Toronto, ON, Canada, 14–18 September 1997; pp. 127–134.
7. Harris, D.-V.; Pan, G. Mineral Favorability Mapping: A Comparison of Artificial Neural Networks, Logistic Regression, and Discriminant Analysis. *Nat. Resour. Res.* **1999**, *8*, 93–109. [[CrossRef](#)]

8. Brown, W.M.; Gedeon, T.D.; Groves, D.I.; Barnes, R.G. Artificial neural networks: A new method for mineral prospectivity mapping. *Aust. J. Earth Sci.* **2000**, *47*, 757–770. [[CrossRef](#)]
9. Bradley, A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **1997**, *30*, 1145–1159. [[CrossRef](#)]
10. Rigol-Sanchez, J.P.; Chica-Olmo, M.; Abarca-Hernandez, F. Artificial neural networks as a tool for mineral potential mapping with GIS. *Int. J. Remote Sens.* **2003**, *24*, 1151–1156. [[CrossRef](#)]
11. Skabar, A.A. Mineral Potential Mapping Using Bayesian Learning for Multilayer Perceptrons. *Math. Geol.* **2007**, *39*, 439–451. [[CrossRef](#)]
12. Skabar, A.A. Mapping Mineralization Probabilities using Multilayer Perceptrons. *Nat. Resour. Res.* **2005**, *14*, 109–123. [[CrossRef](#)]
13. Leite, E.P.L.; de Souza Filho, C.R. Artificial neural networks applied to mineral potential mapping for copper-gold mineralizations in the Carajás Mineral Province, Brazil. *Geoph. Prosp.* **2009**, *57*, 1049–1065. [[CrossRef](#)]
14. Abedi, M.; Norouzi, G.-H. Integration of various geophysical data with geological and geochemical data to determine additional drilling for copper exploration. *J. Appl. Geophys.* **2012**, *83*, 35–45. [[CrossRef](#)]
15. Broomhead, D.S.; Lowe, D. Multivariable functional interpolation and adaptive networks. *Complex Syst.* **1988**, *2*, 321–355.
16. Moody, J.E.; Darken, C.J. Fast learning in networks of locally-tuned processing units. *Neural Comput.* **1989**, *1*, 281–294. [[CrossRef](#)]
17. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin, Germany, 2006; 758p.
18. Rumelhart, D.E.; McClelland, J.L. *Parallel Distributed Processing*; Volume 1: Foundations; MIT Press: Cambridge, MA, USA, 1986.
19. Siedlecka, A.; Roberts, D. *Finnmark Fylke. Berggrunnsgeologi M 1:500 000*; Norges Geologiske Undersøkelse (NGU): Trondheim, Norway, 1996.
20. Gayer, R.A.; Hayes, S.J.; Rice, A.H.N. The structural development of the Kalak Nappe Complex of eastern and central Porsangerhalvøya, Finnmark, Norway. *Nor. Geol. Unders. Bull.* **1985**, *400*, 67–87.
21. Gayer, R.A.; Rice, A.H.N.; Roberts, D.; Townsend, C.; Welbon, A. Restoration of the Caledonian Baltoscandian margin from balanced cross-sections: The problem of excess continental crust. *Trans. R. Soc. Edinb. Earth Sci.* **1987**, *78*, 197–217. [[CrossRef](#)]
22. Sturt, B.A.; Pringle, I.R.; Ramsay, D.M. The Finnmarkian phase of the Caledonian orogeny. *J. Geol. Soc.* **1978**, *135*, 597–610. [[CrossRef](#)]
23. Laird, M.G. Stratigraphy and sedimentology of the Laksefjord Group. Finnmark. *Nor. Geol. Unders. Bull.* **1972**, *278*, 13–40.
24. Roberts, D. The terrane concept and the Scandinavian Caledonides: A synthesis. *Nor. Geol. Unders. Bull.* **1988**, *413*, 93–99.
25. Siedlecka, A.; Roberts, D.; Nystuen, J.P.; Olovyanishnikov, V.G. Northeastern and northwestern margins of Baltica in Neoproterozoic time: Evidence from the Timanian and Caledonian Orogens. In *The Neoproterozoic Timanide Orogen of Eastern Baltica*; Gee, D.G., Pease, V., Eds.; Geological Society of London Memoirs: London, UK, 2004. [[CrossRef](#)]
26. Pharaoh, T.C.; Warren, A.; Walsh, N.J. Early Proterozoic volcanic suites of the northernmost of the Baltic Shield. In *Geochemistry and Mineralization of Proterozoic Volcanic Suites*; Pharaoh, T.C., Beckinsale, R.D., Rickard, D., Eds.; Geological Society of London Special Publications: London, UK, 1987. [[CrossRef](#)]
27. Pharaoh, T.C.; Brewer, T.S. Spatial and temporal diversity of early Proterozoic volcanic sequences - comparisons between the Baltic and Laurentian shields. *Precambrian Res.* **1990**, *47*, 169–189. [[CrossRef](#)]
28. Hanski, E.; Huhma, H. Central Lapland greenstone belt. In *Precambrian Geology of Finland. Key to the Evolution of the Fennoscandian Shield*; Lehtinen, M., Nurmi, P.A., Rämö, O.T., Eds.; Elsevier Science B.V.: Amsterdam, The Netherlands, 2005; pp. 139–193.
29. Melezhik, V.A.; Hanski, E.J. Palaeotectonic and palaeogeographic evolution of Fennoscandia in the Early Palaeoproterozoic (Chapter 3.3). In *Reading the Archive of Earth's Oxygenation; Volume 1: The Palaeoproterozoic of Fennoscandia as Context for the Fennoscandian Arctic Russia-Drilling Early Earth Project*; Melezhik, V.A., Prave, A.R., Fallick, A.E., Kump, L.R., Strauss, H., Lepland, A., Hanski, E.J., Eds.; Springer-Verlag: Berlin, Germany, 2013; pp. 111–178. [[CrossRef](#)]

30. Braathen, A.; Davidsen, B. Structure and stratigraphy of the Palaeoproterozoic Karasjok Greenstone Belt, North Norway—Regional implications. *Nor. Geol. Tidsskr.* **2000**, *80*, 33–50. [[CrossRef](#)]
31. Sandstad, J.S.; Bjerkgård, T.; Boyd, R.; Ihlen, P.; Korneliussen, A.; Nilsson, L.P.; Often, M.; Eilu, P.; Hallberg, A. Metallogenic areas in Norway. In *Mineral Deposits and Metallogeny of Fennoscandia*; Eilu, P., Ed.; Geological Survey of Finland Special Paper: Helsinki, Finland, 2012. [[CrossRef](#)]
32. Olesen, O.; Sandstad, J.S. Interpretation of the Proterozoic Kautokeino Greenstone Belt, Finnmark, Norway from combined geophysical and geological data. *Nor. Geol. Unders. Bull.* **1993**, *425*, 43–64.
33. Pharaoh, T.C.; Pearce, J.A. Geochemical Evidence for the Geotectonic Setting of Early Proterozoic Metavolcanic Sequences in Lapland. *Precambrian Res.* **1984**, *25*, 283–308. [[CrossRef](#)]
34. Zwaan, K.B.; Gautier, A.M. Alta and Gargia. Description of the geological maps (AMS-M711) 1834 1 and 1934 IV 1:50 000. *Nor. Geol. Unders. Bull.* **1980**, *357*, 1–47.
35. Pharaoh, T.C.; Ramsay, D.M.; Jansen, Ø. Stratigraphy and structure of the Repparfjord-Komagfjord Window, Finnmark, Northern Norway. *Nor. Geol. Unders. Bull.* **1983**, *377*, 1–45.
36. Bergh, S.G.; Torske, T. Palaeovolcanology and tectonic setting of a Proterozoic metatholeiitic sequence near the Baltic Shield margin, northern Norway. *Precambrian Res.* **1988**, *39*, 227–246. [[CrossRef](#)]
37. Torgersen, E.; Viola, G.; Sandstad, J.S. Revised structure and stratigraphy of the northwestern Repparfjord Tectonic Window, northern Norway. *Nor. J. Geol.* **2015**, *95*. [[CrossRef](#)]
38. Pharaoh, T.C. The stratigraphy and sedimentology of autochthonous metasediments in the Repparfjord-Komagfjord Tectonic Window, west Finnmark. In *The Caledonide Orogen—Scandinavia and Related Areas*; Gee, D.G., Sturt, B.A., Eds.; John Wiley & Sons Ltd.: Chichester, UK, 1985; pp. 347–372. [[CrossRef](#)]
39. Bjørlykke, A.; Olerud, S.; Sandstad, J.S. Metallogeny of Finnmark, North Norway. *Nor. Geol. Unders. Bull.* **1985**, *403*, 183–196.
40. Koistinen, T.; Stephens, M.B.; Bogatchev, V.; Nordgulen, Ø.; Wennerström, M.; Korhonen, J. *Geological map of the Fennoscandian shield, scale 1:2,000,000, Geological Surveys of Finland, Norway and Sweden and the North-West Department of Natural Resources of Russia*; Karelian Research Center: Petrozavodsk, Russia, 2001.
41. Dumais, M.-A. *Reprocessing and Compilation of Radiometrics Data from Finnmark and Northern Troms*; Report No. 2014.015; Geological Survey of Norway (NGU): Trondheim, Norway, 2013; 48p.
42. Rodionov, A.; Ofstad, F.; Stampolidis, A.; Tassis, G. *Helicopter-Borne Magnetic, Electromagnetic and Radiometric Geophysical Survey at Stjernøy, Finnmark County*; Report No. 2014.033; Geological Survey of Norway (NGU): Trondheim, Norway, 2014.
43. Nasuti, A.; Roberts, D.; Dumais, M.-A.; Ofstad, F.; Hyvönen, E.; Stampolidis, A.; Radionov, A. New high-resolution aeromagnetic and radiometric surveys in Finnmark and North Troms: Linking anomaly patterns to bedrock geology and structure. *Nor. J. Geol.* **2016**, *96*. [[CrossRef](#)]
44. Olesen, O.; Roberts, D.; Henkel, H.; Lile, O.B.; Torsvik, T.H. Aeromagnetic and gravimetric interpretation of regional structural features in the Caledonides of West Finnmark and North Troms, northern Norway. *Nor. Geol. Unders. Bull.* **1990**, *419*, 1–24.
45. Bridle, J.S. Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters. *Adv. Neural Inf. Process. Syst.* **1990**, *2*, 211–217.
46. Chen, Z.; Haykin, S. On different facets of regularization theory. *Neural Comput.* **2002**, *14*, 2791–2846. [[CrossRef](#)] [[PubMed](#)]
47. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]
48. Carranza, E.J.M.; Hale, M.; Faassen, C. Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping. *Ore Geol. Rev.* **2008**, *33*, 536–558. [[CrossRef](#)]
49. Stensgaard, B.M.; Chung, C.J.; Rasmussen, T.M.; Stendal, H. Assessment of mineral potential using cross-validation techniques and statistical analysis: A case study from the Paleoproterozoic of West Greenland. *Econ. Geol.* **2006**, *101*, 1297–1413. [[CrossRef](#)]

