Journal of Mining and Environment (JME), Vol. 15, No. 2, 2024, 761-790

Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

A Comparative Analysis of Artificial Neural Network (ANN) and Gene Expression Programming (GEP) Data-driven Models for Prospecting Porphyry Cu Mineralization; Case Study of Shahr-e-Babak Area, Kerman Province, SE Iran

Bashir Shokouh Saljoughi[1*], and Ardeshir Hezarkhani[2]

Department of Mining and Metallurgy Engineering, Amirkabir University of technology (Tehran Polytechnic), Tehran, Iran

Article Info

Received 16 November 2023

Received in Revised form 3 December 2023

Accepted 14 December 2023

Published online 14 December 2023

DOI: 10.22044/jme.2023.13852.2573

Keywords

Mineral prospectivity mapping

Artificial neural network

Gene expression programming

Cu mineralization

Abstract

The porphyry Cu-mineralization potential area studied in this article is located in the southern section of the Central Iranian volcano–sedimentary complex, contains large number of mineral deposits, and occurrences that are currently facing a shortage of resources. Therefore, prospecting potential areas in the deeper and peripheral spaces has become a high priority in this region. Different direct and indirect methods try to predict promising areas for future explorations that most of them are very time-consuming and costly. The main goal of mineral prospecting is applying a transparent and robust approach for identifying high potential areas to be explored further in the future. This study presents the procedure taken to create two different Cu-mineralization prospectivity maps. This study aims to investigate the results of applying the ANN technique, and to compare them with the outputs of applying GEP method. The geo-datasets employed for creating evidential maps of porphyry Cu mineralization include solid geology map, alteration map, faults, dykes, airborne total magnetic intensity, airborne gamma-ray spectrometry data (U, Th, K and total count), and known Cu occurrences. Based on this study, the ANN technique (10 neurons in the hidden layer and LM learning algorithm) is a better predictor of Cu mineralization compared to the GEP method. The results obtained from the P-A plot showed that the ANN model indicates that 80% (vs. 70% for GEP) of the identified copper occurrences are projected to be present in only 20% (vs. 30% for GEP) of the surveyed area. The ANN technique due to capabilities such as classification, pattern matching, optimization, and prediction is useful in identifying anomalies associated with the Cu mineralization.

1. Introduction

The first stage in a systematic exploration is to prospect new mineralization in a region of interest, which is known as Mineral Potential Mapping (MPM) [1, 2]. MPM is the process of creating a map that illustrates the favorability of mineralization occurrence in a specified area. Identifying high potential areas within a promising region is one of the fundamental objectives in mineral exploration projects [3-5]. The primary objective of MPM is to optimize the efficiency of mineral exploration programs by minimizing costs and time of prospecting, while simultaneously maximizing the prospecting benefit of a mineral exploration program [6].

✉ *Corresponding author: bashir.shokouh@gmail.com (B. Shokouh Saljoughi)*

Different exploratory geo-datasets including geological, geo-physical, and geo-chemical thematic layers are, respectively, collected, processed, and integrated for mineral potential mapping. Consequently, mineral potential mapping can be considered as a multiple criteria decision-making (MCDM) procedure, which creates a predictive model to delineate prospective zones [7-10].

The main object of mineral potential mapping is to apply a transparent and robust method to determine favorable zones of a specific mineralization. The most important steps in mineral potential mapping include selecting proper targeting criteria and applying advanced, robust, and reproducible approach for deriving evidential features [4, 5]. Therefore, the methodological aspects and choosing an appropriate algorithm or methodology in order to obtain the spatial relationships between evidential features and known deposit locations are very vital. As a result, an accurate mineral potential map is generally achieved depending on the capacity of the algorithm to learn complex relationships between the input evidential features and the occurrence of mineralization, as well as interpretability and transparency of the algorithm used. Furthermore, easy implementation of the algorithm and availability of the software tools in practical applications are also important [11, 12].

The most definitive way to know if an area contains mineral deposits, is drilling but it is time-consuming and involves high costs along with a trial and error process [13]. Consequently, geological, geo-chemical, geo-physical, and remote sensing data, which are related to the mineralization are fed to a prediction algorithm to help mining companies on decision-making for locating best drill-hole points. Also advances in computer technology play an increasingly important role in the MPM [12, 14].

In the past decades, several MPM approaches in order to deal with multiple datasets or layers of diverse character from various sources, quantitative integration, and evaluation of these layers have been developed [15]. These approaches can be categorized into three classes including data driven, knowledge-driven, and hybrid methods. Data driven or empirical techniques like logistic regression [11,15-16], weights of evidence (WofE) [17-20], extended weights of evidence [21], evidential belief functions [22- 24], fuzzy weights of evidence [25, 26], support vector machines [1, 2], fuzzy logic [14], feed-forward neural networks [27], multi-layer perceptrons [28], Bayesian networks [25], radial basis functional link net [29], probabilistic neural networks [30, 31], and restricted Boltzmann machine [32] apply the relationship between discovered mineral deposits and their surrounding map patterns to create a mineral potential map. The second category, i.e. knowledge-driven methods, like fuzzy logic [7, 22, 33-34], Boolean logic [35], and wildcat mapping [36], outranking method [37] is based on the expert judgment of the analyst. The hybrid methods, such as fuzzy weights-of-evidence modeling [26, 38], data-driven fuzzy modeling [22], and neuro-fuzzy modeling [39] has been proposed in order to optimize the application of both conceptual knowledge of mineral systems and empirical spatial associations between mineral deposits and evidential features in geo-computational modelling.

In the recent years, new types of the MPM techniques have been developed. These techniques compared to the conventional ones, assign weights to the continuous spatial evidence without considering the location of known mineral occurrences and expert judgments. Fuzzy logic MPM with continuous evidential data, data-driven index overlay, data-driven Boolean logic, expected value and geometric average are from these techniques. They are expected to overcome the problem of exploration bias resulting from the expert judgments in knowledge-driven MPM and from using characteristics of known mineral occurrences in data-driven MPM [40, 41]. Also, it should also be noted that in the recent decades, fractal and multifractal methods have been used to improve mineral potential mapping [42-45].

Furthermore, Yousefi *et al.* (2019) [46] classified ore-forming processes into three sub-systems, and explained how an exploration information system (EIS) can be constructed in conjunction with GIS and in the context of "mineral sub-systems" to help design information system architecture for better understanding of complex ore-forming processes and to help translate mineral systems understanding into practical applications. Exploration "data" is gradually converted into (1) "information" regarding ore-forming geological processes, (2) "knowledge" of the processes, and (3) "insight" critical to establishing mineral exploration tactics using this approach. As a result, by facilitating the translation of ore-forming processes into spatially-mappable criteria, an EIS can be immensely valuable in exploration targeting by gradually converting "data" to "information," "knowledge,"

and "insight." Importantly, an EIS integrates data, information, and expertise from multiple sources [47, 48].

This study aims to apply data-driven Genetic Expression Programming (GEP) and Artificial Neural Network (ANN) models for mineral potential mapping and to compare the outputs with each other. Shahr-e-Babak district in Kerman Province, Iran, which is located in the Cenozoic Sahand-Bazman orogenic belt and has a complex geological setting, was selected as the case study area, and the GEP and ANN integration approaches are applied to map Cu mineralization.

2. Geological Setting

The studied area in the west of Kerman (Figure 1a) is located in the southern portion of the Central Iranian Volcanic Sedimentary Complex in Iran, from a geographical perspective (Figure 1c). The NE portion of the studied area is mountainous, with the Paqaleh range in the E, the Kuh-e-Mosahim volcano in the center, and the Narkuh range in the NW. The Shahr-e-babak dasht plain, known as Dasht-e Yekkeh Baneh, is located in the southern and southwestern regions of the area. Situated at an altitude ranging from 1800 to 2000 m, the drainage pattern in this region is primarily centrifugal inside the mountains and predominantly parallel within the dasht plain. Notably, Rud-e Kang in the western part of the area serves as the principal watercourse. The majority of the sheet's land is drained in a southwestern direction [49]. The selection of the study location was based on a review of regional copper potential in the Shahr-e-Babak district, indicating a high potential for mineral deposits. The aforementioned area has a multitude of mineral occurrences, accompanied by a wealth of geological, geochemical, and geo-physical survey data that is readily accessible.

The geological map of the studied area is illustrated in Figure 1b, with a scale of 1:100,000. This map was derived from the geological survey conducted by a Yugoslavian research team during the 1970s. The most ancient geological formations within the region are comprised of Cretaceous sediments, specifically found in the northernmost section of the sheet. These sediments are commonly referred to as Cenomanian-Turonian flysch, often exhibiting bioglyphs. Additionally, the southern region of the sheet is characterized by a mélange formation with distinct coloration [49]. The Eocene flysch sediments are found inside the exposed central region of an open-anticline axis that runs in a northwest-southeast direction in the northern portion of the research area. The dominant geological unit in the area is represented by the Eocene volcanic-sedimentary Units, comprising around 80 percent pyroclastics, 10 percent lava flows, and the remaining 10 percent sediments [49].

The extensive volcano-sedimentary sequence, with a minimum thickness of 15 kilometers, has been classified into three complexes, namely “Bahr Aseman”, “Razak”, and “Hezar”, as delineated by Dimitrijevic et al. (1973) [50].

- The Lower Eocene “Bahr Aseman” complex commences with basal arenites that are several meters in thickness. The subsequent units in the sequence include pyroclastics, rhyodacite, trachyandesite, andesite, and rhyolite.
- The “Razak” complex is situated over a sedimentary sequence that dates back to the Middle Eocene period. The complex is partitioned into three distinct components. The lowest volcanic unit primarily consists of basic rock types, whereas the intermediate unit is predominantly composed of acidic rock types. The higher unit, once again, is primarily composed of basic rock types.
- The “Hezar” complex is indicative of the most recent stage of volcanic activity during the Eocene period. The geological formation consists of a sequence of fast alternating lava flows including andesite-basalt, andesite, and rhyolite.

The sequence of intrusions can be attributed to the Miocene epoch, perhaps occurring within the temporal range from the uppermost Oligocene to the Upper Miocene. Igneous rocks can be classified into two distinct categories [50].

- The geological formations in Jebal Barez consist of granodiorite rocks, including granite, granodiorite, quartz diorite, monzonite, and syenite. The research region is characterized by the presence of the Chenar massif and several minor intrusions.
- The Kuh-e-Panj type is characterized by the presence of sub-volcanic rocks that typically exhibit a porphyritic texture and a wide range of compositions including quartz diorite, diorite, dacite, and others. Certain volcanic necks within the studied area have been identified as "dacite and andesite," falling under this particular classification.

The primary volcano activity of Kuh-e-Mosahim commenced with a pyroclastic

explosion, followed by the deposition of mostly pyroclastic debris, accompanied by a limited number of andesite lava flows. Within the caldera, there are occurrences of diminutive diorite bodies and dykes that are exposed in areas where the surrounding rocks have undergone significant alteration.

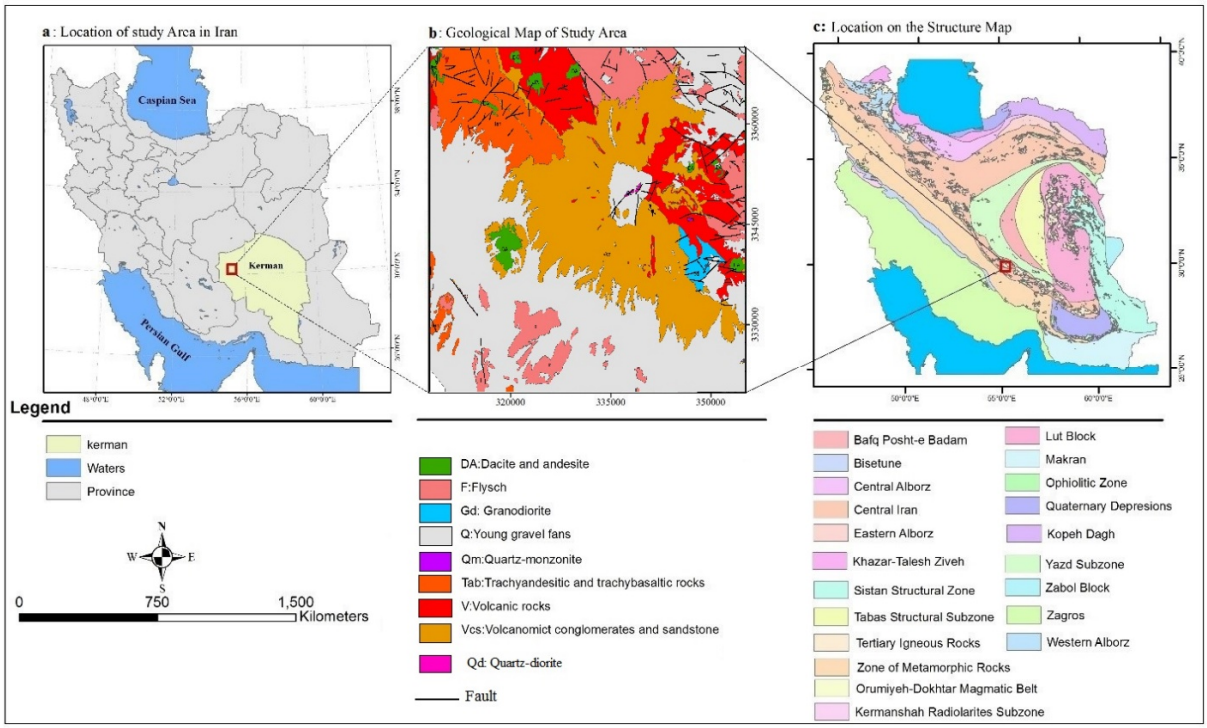


Figure 1. (a) Location of study area in Iran, (b) Geological Map of interested area [49] and (c) position of study area on the Urumieh–Dokhtar volcanic belt [51].

In the region, a considerable number of dykes with a northwest-southeast orientation were seen. The research area exhibits a prevalence of N-S and E-W strikes, with the former being more prominent in the northern region and the latter in the eastern region. According to several sources, the age of dykes can be traced back to the Eocene period and extends up until the most recent volcanic activity during the Pliocene epoch. The composition of dykes exhibits significant variability; nonetheless, it is worth noting that dacite and andesite dykes are particularly prominent in this context [49].

The structural interpretation of the north-central portion of the research area suggests that the Eocene and Cretaceous flysch occupy a central position within a regional-scale "open" anticline. The region's southwest flank is characterized by the significant Dehno fault, which serves as a prominent boundary between Eocene flysch deposits and volcanic formations. A zone of weakness that runs parallel to the southwest, spanning approximately 10 kilometers, is associated with aberrant readings of scintillometric observations. This observation has been further supported by aircraft data [50]. The Miduk district is located in the vicinity of the Kuh-e Mosahim volcano, which serves as the central caldera of this region. Additionally, the Chenar granitoid massif is also situated within this geographical zone [49].

2.1. Mineralization of studied area

Porphyry copper deposits are the most important copper sources in the world which supply nearly three-fourths of the world's copper. Porphyry copper deposits (PCDs) in Iran mainly occur in the Cenozoic Sahand-Bazman orogenic belt. The belt was formed during the Alpine orogeny when Arabian plate subducts under central Iran [49]. Sahand-Bazman orogenic belt hosts major porphyry copper deposits like Sar-Cheshmeh, Sungun and Miduk (Figure 1. (c)). These ones and some of other economic deposits are associated with mid to late Miocene diorite/granodiorite to quartz-monzonite stocks. Miduk and Sar-Cheshmeh, which are located in south-southeast central Iran in NW-trending

middle mountain range of Kerman province, are the largest porphyry copper deposits there, generated in quartz diorite, and granodiorite-quartz monzonite plutons, respectively. The Miduk PCD is located on La Chah mountain about 4 km northeast of the village of Miduk and 85 km northwest of Sar-Cheshmeh PCD. The regional structures of Miduk area consist of a series of NW-SE trending Paleogene volcano-sedimentary rocks. The porphyry system of Miduk forms a semi-radial quartz diorite stock, and is restricted to severely biotitized volcanic wall rock [49].

The northern region of the research area experiences the presence of hydrothermal copper and lesser amounts of lead occurrences. The Lachah occurrence exhibits characteristics consistent with the porphyry copper type, wherein the presence of copper carbonates, chrysocolla, and limonite is observed in the superficial alteration zone. Additionally, primary pyrite and disseminated chalcopyrite are found in the deposit. Vein-type copper mineralization is observed at Chah Massi, characterized by the presence of pyrite, chalcopyrite, and galena. Similarly, in the Chehel Dokhtaran region, copper mineralization is found within quartz veins containing magnetite and malachite. The southern slopes of Kuh-e-Masahim exhibit a hydrothermally altered zone characterized by an east-west trend. In this region, a mineralized fault zone oriented in the NNW-SSE direction is present, like the fault zone situated to the south of Kuh-e-sara. In this region, there exists a mineralized fault zone oriented in the NNW-SSE direction, which has resemblance to the fault zone situated south of Kuh-e-sara. Copper (Cu) and lead (Pb) mineralization have been identified within these zones. The events taking place in the area are related to the Neogene volcanic activity [49].

3. Spatial Geo-dataset Descriptions

The database utilized in this case study pertains to the Shahr-e-Babak 1:100,000, sheet and has been compiled by the Geological Survey of Iran [49]. The selection of this particular region serves as the basis for a preliminary investigation aimed at evaluating the efficacy of ANN and GEP in the context of MPM. Due to the abundance of mineral occurrences in this region, a substantial quantity of samples can be procured for the purpose of training in order to employ the ANN methodology. Subsequently, the GEP methodology can be applied for comparative analysis. Additionally, all thematic layers cover the entire studied area uniformly and completely.

The solid geology map, alteration map, faults, dykes, airborne total magnetic intensity and airborne gamma-ray spectrometry (U, Th, K, and total count), make up the input data for ANN and GEP (Table 1 and Figure 2). Mineral occurrences and small- and medium-sized mines are included in the output layer. A 200 m square on the ground is represented by each cell in the gridded data.

Table 1. Spatial geo-dataset of study area.

Category	Factors	Data type	Scale	Remarks
Occurrence	Cu	Point	-	36 occurrences and mines
Geo-chemical data	As, B, Ba, Co, Cr, Cu, Mo, Ni, Pb, Sb, Sn, W, Zn	Point	1:100000	IDW (Inverse Distance Weight) interpolation
Geological data	Geology	Polygon	1:100000	Shahr-e-Babak map sheet
Faults	Distance from fault	Polyline	1:100000	Fault density map
Dyke	Distance from dyke	Polyline	1:100000	Dyke density map
Geo-physical data	Magnetic anomaly	Point	1:100000	IDW (Inverse Distance Weight) interpolation
Airborne gamma ray spectrometry	U, Th, K and total count anomaly	Point	1:100000	IDW (Inverse Distance Weight) interpolation
Satellite data	Alterations	Raster	-	Band ratios

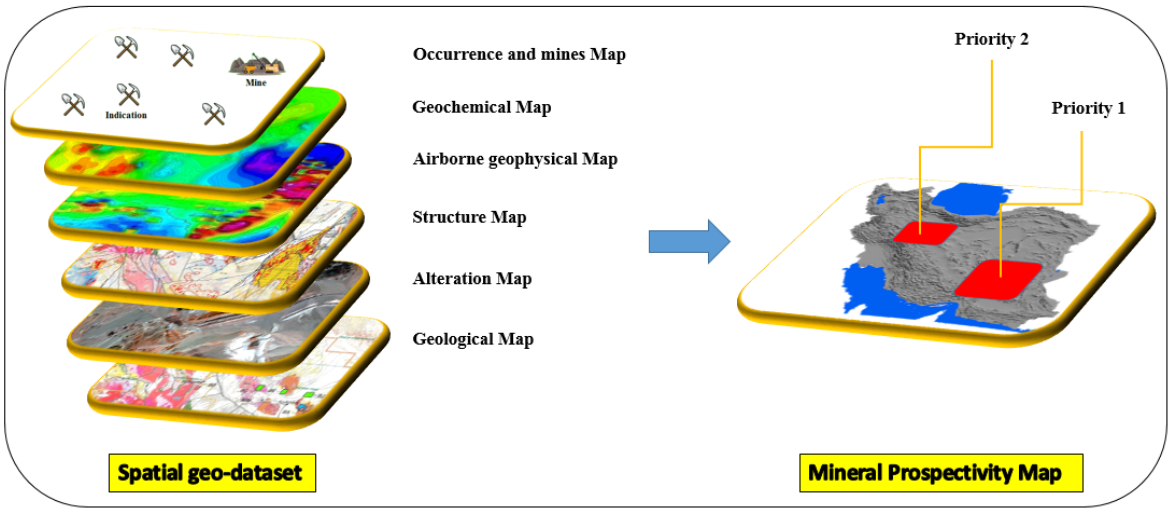


Figure 2. Schematic image of spatial geo-dataset for mineral prospectivity map.

4. Methodology

4.1. Creating ANN model

The utilization of artificial neural networks (ANNs) has garnered substantial attention from exploration geologists in the recent decades as a reliable and replicable method for generating mineral potential maps. ANNs are valued for their transparency, robustness, and reproducibility in this context. ANNs possess a non-linear mathematical framework that enables them to execute curve-fitting operations in multidimensional space. Therefore, it possesses the capability to effectively depict a data-generating mechanism of considerable complexity, establishing a connection between the inputs and outputs of said mechanism [52-54].

In the field of mineral potential mapping, various neural network models including multilayer perceptrons [28], radial basis function network [29], probabilistic neural networks [30], and self-organizing map [55] have been utilized in the recent times. In every ANN system, three key components have significant importance: (a) the configuration of the nodes, (b) the network's topology, and (c) the learning method employed to determine the weights of the ANN [56, 57]. In contrast, within an ANN, every processing unit functions as a conceptual neuron, whereby it accepts input, performs activation computations, and afterwards sends the resulting activation to other processing units. Each link between these processing units is coupled with a weight value, which is defined to indicate the strength of the relationship. The ideal determination of the connection weight of each processing unit is achieved by presenting known examples and applying a learning rule. After the determination of connection weight by neural network learning, the inter-connection between the input and output that is embedded in the data is captured [52].

ANNs are a computational approach that operates based on the same underlying principles as neurons in the human brain. During the process of data processing using an ANN, an artificial neuron provides output data based on the Equation (1) [58]:

$$y_i = f\left(\sum_{i=1}^{N} x_i w_i\right) \quad (1)$$

Equation 1 is defined as follows: w_i denotes the weights, x_i represents the input data, f is the activation function, and y_i represents the output for the i^{th} element. The operational mechanism of an artificial neuron unit is depicted in Figure 3.

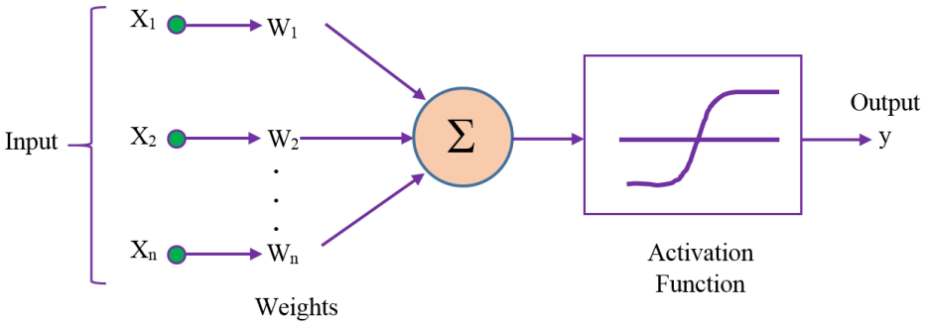


Figure 3. Schematic representation of an artificial neuron.

ANN model was constructed using three input variables and one output variable. The topology chosen for this study was the Multilayer Perception (MLP) model, which consisted of a single hidden layer. According to the universal approximation theorem, it has been proven that a neural network consisting of a single hidden layer is capable of accurately approximating any continuous and non-linear function [54]. The activation function employed in this study was the sigmoid activation. The ANN model is depicted in Figure 4.

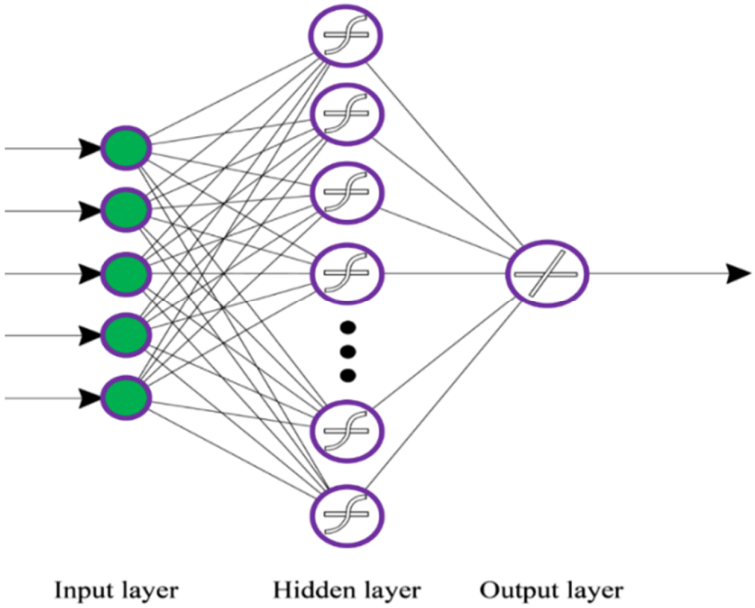


Figure 4. Schematic illustration of the ANN topology.

4.2 Creating GEP model

Genetic algorithms (GAs) are the straightforward case of evolutionary algorithms. In a genetic algorithm (GA), an iterative process is employed to generate a sequence of solutions for the purpose of optimizing a given issue, with the aim of obtaining increasingly accurate solutions. The process of evolution involves successive repeats known as generations. The algorithm encompasses both the solution domain and an objective function. GEP, which was developed by Ferreira, can be considered as a natural progression of genetic programming (GP) as originally proposed by Koza [59]. Unlike GA, the individuals of GEP are non-linear entities of different sizes and shapes. The primary benefit of Genetic Programming (GP) is in its ability to generate a solution by systematically exploring a vast search space comprised of computer programs that exhibit high fitness [60]. The primary components of the GEP can be delineated as follows: the set of functions, the set of terminals, the fitness function, control parameters, and a termination condition [61]. The Genetic Programming (GP) algorithm generates a population of solutions that are structured by an expression tree (ET).

The Genetic Expression Program (GEP) defines the encoding of ramified expression trees (ETs), which represent the expression of genetic information, in chromosomes. This encoding involves the use of one or more linear symbolic strings of a pre-determined length, known as genes. Each gene consists of two distinct domains, namely the head and tail, as seen in Figure 5a. The head of a production rule includes symbols that are present in both functions and terminals, whereas the tail exclusively comprises terminal symbols. The set of functions encompasses a wide range of mathematical operations, including addition, subtraction, multiplication, division, exponentiation, logarithmic functions, and logic operators, all of which are valuable for problem-solving purposes. The set of terminals encompasses the variables and constants employed inside a given task [62]. For a practical problem, the value of the head length (h) is selected with a priori, whereas the value of the tail length (t) is evaluated by Equation (2):

$$t = h(n-1)+1 \quad (2)$$

where n is the maximum arity, which refers to the highest number of arguments that may be accepted by the functions. Furthermore, it is important to note that each gene possesses a coding region that is specifically characterized as an open reading frame (ORF). In the context of evolutionary computation, an open reading frame (ORF) is then translated into an expression tree (ET), which represents a potential solution to the given problem (see Figure 5b). The initiation site of an open reading frame (ORF) is invariably located at the initial position of a gene, although the termination site does not necessarily align with the final location of the gene. In this context, it is worth noting that while the length of genes in GEP remains constant, each gene possesses the capacity to encode for ETs of varying sizes and forms. In instances involving multigenic chromosomes, each gene encodes a sub-entity, referred to as a sub-ET. These sub-ETs then interact with one another by a linking function, which can encompass many mathematical functions, including Boolean logic operations with numerous arguments. This mechanism enables the comprehensive representation of the entire individual, as depicted in Figure 5c. The evolutionary computing processes employed by the GEP algorithm are illustrated in Figure 6 [62].

The initial steps of the GEP require human control over several aspects: (i) the selection of terminals, (ii) the determination of primitive functions for each branch of the program being evolved, (iii) the establishment of a fitness measure, (iv) the specification of certain parameters to govern the execution, and (v) the definition of a termination rule and method to determine the outcome of the execution [63].

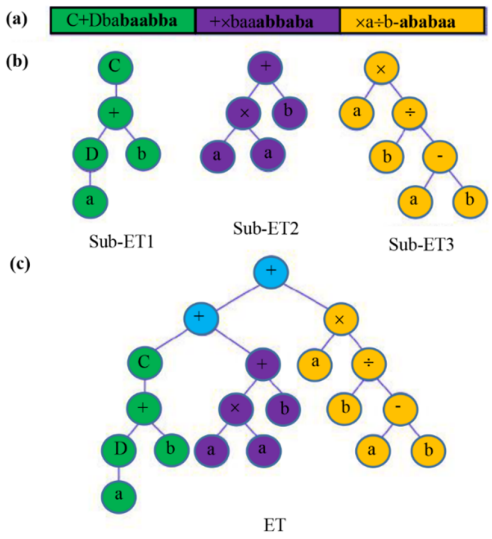


Figure 5. Expression of multigenic chromosomes encoding sub-ETs associated with a two-argument function. (a) A three-genic chromosome with the tails displayed in bold. (b) The sub-ETs codified by each gene. (c) The result of posttranslational linking with addition. The linking functions are presented in yellow.

The flowchart representing the process of genetic programming can be succinctly represented as outlined in reference [64].

(1) Create an initial population of randomly generated compositions consisting of the functions and terminals relevant to the situation at hand, specifically computer programs.

(2) Perform the execution of each program inside the population and allocate a fitness value to each program based on its effectiveness in solving the problem.

(3) Create a new population of computer programs:

a. Copy the best existing programs.

b. Create new computer programs by mutation.

c. Create new computer programs by crossover (sexual reproduction).

(4) The best program that appeared in any generation, the best solution so far, is

designated as the genetic result of genetic programming.

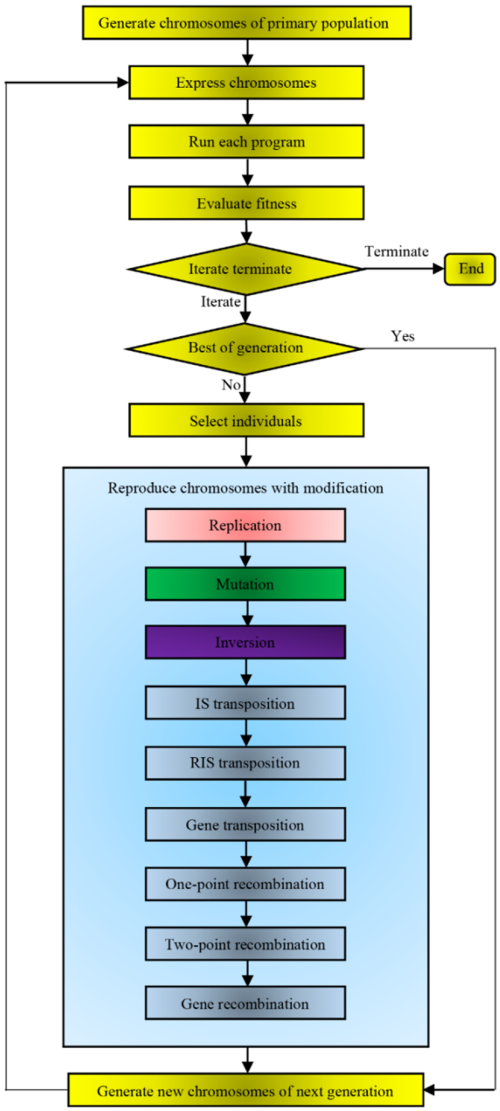


Figure 6. Flowchart of GEP algorithm (adapted from [64]).

This study employs the GEP approach to develop empirical models for the estimation of MPM based on geo-spatial data. The generalization capability of the GEP algorithm is influenced by the selection of parameters including the chromosome, head size, and function set. Hence, the selection of these factors was informed by empirical investigations and existing values proposed in the scholarly literature [65, 66].

5. Discussion and Results

The following issues should be answered to better understand geological controls on deposit development within mineral systems relevant to mineral exploration targeting [67, 68]: (1) What is the fundamental nature of the system?; (2) What are the stress, thermal and dynamic processes of the system?; (3) What are the fluid sources of the system?; (4) What are the channels/pathways for moving the fluids and drivers of fluid mobility?; and (5) What are the ore depositional/precipitation mechanisms? These questions stem from efforts to comprehend the crucial geological factors that must be present in order for mineral deposits to emerge [69, 46]. The critical geological elements involved in the mineral system concept include: (i) tectonic processes that activate and define the favorable time intervals for mineralizing events; (ii) geochemical and tectonic processes that produce metal sources, hydrothermal fluids and magmas; (iii) geological processes that make conduits, which can act as pathways for fluid/magma flow from lithospheric, to district scales; (iv) processes of fluid flow; (v) focusing mechanisms that throttle the fluids flow into depositional trap sites; (vi) physico-chemical processes that lock in metal at the trap; and (vii) preservation, exhumation, and upgrading processes allowing later detection and economic extraction of mineralization. When these processes are taken into account, mineral systems can be divided into pre-, syn-, and post-mineralization components, which are referred to as sub-systems. This framework assists us in better understanding the underlying geological processes in order to translate them into geographical proxies and prediction maps. The role of scale dependency of geological processes that affect mineral system components is particularly important and should be emphasized. In mineral exploration targeting, there is typically a sequence of scale-dependent targeting stages for vectoring toward mineral deposit sites, depending on the datasets available. Ignoring this and erroneously integrating mineral exploration targeting criteria pertinent to different scales will unavoidably lead to poor exploration targets. Therefore, in this article, it has been tried to use the EIS approach as much as possible [47].

5.1. ANN model results

5.1.1. Data preparation

Precursors (inputs) to the formation of mineralization, namely, source, transport,

physiochemical trap, and deposition [70], were used to define a suite of mappable targeting criteria. The input layers can exhibit variability based on the data models, structures, and formats employed. The input layers depending on the data models, structures and formats can be varied. The evidential maps due to the controlling factors of porphyry copper mineralization in the Urumieh-Dokhtar magmatic belt encompass several thematic layers. These layers consist of solid lithology, tectonic features such as faults and dykes, airborne measurements of total magnetic intensity, airborne gamma-ray spectrometry data including measurements of uranium (U), thorium (Th), potassium (K), and total count, remotely sensed data indicating hydrothermal alteration haloes, geochemical data, and the spatial distribution of known occurrences.

Source: In light of the spatial-temporal association [70], the correlation between magma type and metal content [71], and isotopic signatures of fluids [72], there is a consensus on the source of metals and sulfur for porphyry copper deposits being the magmas forming shallow porphyritic plutons [73, 74]. In the area confined by this study, shallow granitoid intrusions of calc-alkaline affinity, Oligocene-Miocene in age, have emplaced within an Eocene volcanic sequence [75-78]. These are deemed the source of metals and sulfur for porphyry copper mineralization. The outcropping Oligocene-Miocene plutons can be mapped by their polygons portrayed in the geological map of the area [75]. Prior to conducting additional processing, it is necessary to transform all thematic layers inside the geo-database into raster format. The lithological layer utilized in this study is derived from the 1:100,000 geological map published by the Geological Survey of Iran (GSI). The geological base layer comprises a total of 23 distinct rock units, a significant proportion of which exhibit either negligible or minimal depositional occurrences. Units belonging to the same stratigraphic group are integrated. The outcome is a simplified geology layer including nine distinct rock units. Figure 7a illustrates the ultimate geological map.

Transport: Structural data containing the map of faults and dykes can be provided by various sources. Fracture/fault systems have been a focus of research for decades owing to their importance in contexts such as hydrocarbon accumulation, contaminant transport, engineering geology and seismogenic faults. Rock is rarely a continuous medium and contain many natural breaks such as faults, joints, bedding planes, and schistosity planes. Fractures in a rock mass usually control its overall behavior so that deformability, stability of underground excavations, and fluid flow depend significantly on the fracture network. Fault/fracture systems are also significant to mineralization [79-83], because faulting activities confine magmatic activities into certain spatial scales and temporal stages, which provide both hydrothermal fluids and heat for mineralization within fracture zones. The concept that fracture/fault patterns show a degree of self-similarity over a wide range of scales has long been familiar to geologists [84, 85] and follow a power law in all scales. In general, some methods such as fractal [86], Fry [87- 90] and fracture density [91] analyses as well as lineament factor value method [92, 93] can be used for structural control analysis on mineralization. In these methods, spatial distribution of a set of points (e.g. occurrences of mineral deposits and fault intersection points) can be characterized by point pattern analysis [94]. In this paper, they are obtained from geologic maps and remotely sensed data, then classified and coded according to their strike. The presence of east-trending lineaments in the Kerman region has a strong correlation with the occurrence of porphyry copper deposits [50]. Linear geological features such as faults and dykes are identified and retrieved from the geological map. These features are then adjusted and refined based on the analysis of Landsat 8 satellite imagery. Subsequently, the distribution of faults and dykes is cartographically shown. The layer consisting of faults and dykes is transformed into a grid format, wherein the values assigned to each cell represent the proximity to the closest fault (see Figure 7b, c).

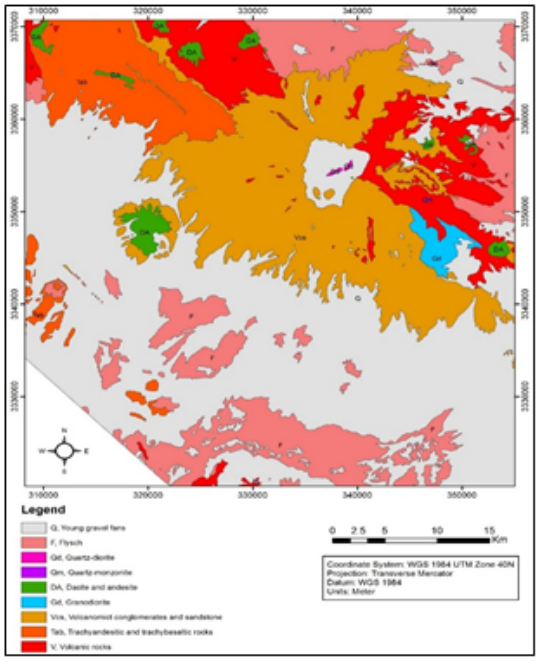
The utilization of airborne magnetic data is an additional dataset employed in the present investigation. The dataset was obtained by the Atomic Energy Organization of Iran (AEOI) between the years 1977 and 1978. The acquisition of airborne magnetic data was conducted with flight lines spaced at intervals of 500 m and at an altitude of 120 m. The analysis of aerial geophysical data pertaining to the nearby mines and exploratory geology has played a crucial role in elucidating the overall geological characteristics and structural elements in the Shahr-e-Babak region of Kerman Province, southeastern Iran. Initially, the magnetic data was subjected to pole reduction using an RTP filter. This approach eliminates the reliance of magnetic data on the inclination. The utilization of this

approach facilitates the process of interpretation as it converts the complex asymmetric responses of sub-vertical prisms or contacts into more straightforward symmetric and anti-symmetric forms. The symmetric "highs" are directly centered on the body, while the maximum gradient of the anti-symmetric dipolar anomalies coincides exactly with the body edges. The removal of the dipolar nature of the magnetic field in the study area is achieved through the application of Fast Fourier filtering programs on the digitized aeromagnetic data [66]. Next, the data obtained from the previous phase is subjected to the analytic signal method, and the outcome is depicted in Figure 7d.

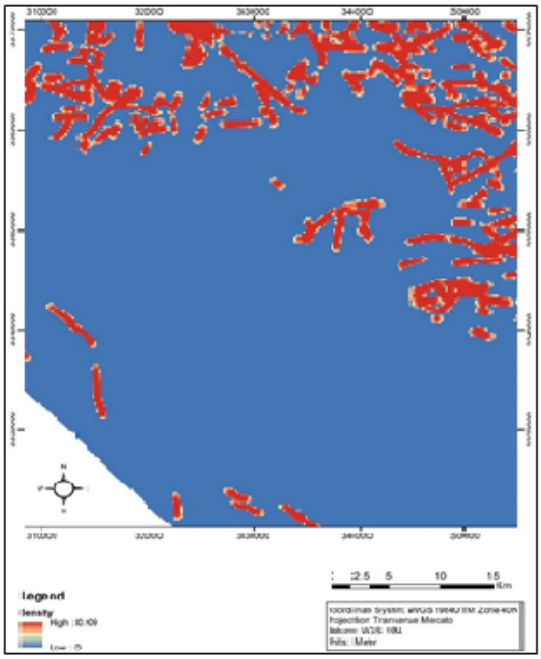
Physico-chemical trap: Several physico-chemical processes including fluctuations in the pressure and temperature, redox interactions, and interactions with wall rocks, exert strict controls over the precipitation of metallic compounds in hydrothermal mineral systems [95]. As such, the interactions between fluids and country rocks lower the temperature of fluids, facilitating the precipitation of metals [95]. These also alter the composition of country rocks, developing zones of hydrothermal alteration. Hydrothermally altered rocks are expressions of the constituent processes linked with physicochemical interactions, yielding specific mineral assemblages. Clay minerals, kaolinite, montmorillonite, dickite, and halloysite, are typical minerals presented in the argillic mineral assemblage. The zone of phyllic alteration is dominated by muscovite, sericite, pyrite, illite, and quartz. Leached caps overlying hypogene zones are also prevalent in porphyry systems, in which hematite, limonite, goethite, and jarosite are typical minerals. These mineral assemblages are of diagnostic spectral absorption features allowing for the remotely sensed recognition of hydrothermal zones using ASTER multispectral data [96]. The utilization of remotely sensed data, specifically the Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER), is employed as an additional dataset. The aforementioned layer has undergone processing subsequent to the application of necessary corrections. The band ratio approach is used on the ASTER data [97] in order to delineate the alteration haloes surrounding the intrusive bodies. One evidential map is derived from the processed images as the alteration haloes (Figure 7e).

Airborne geo-physical surveys have also been conducted to produce gamma-spectrometry measurements of uranium, thorium, potassium, and total count. The process of radiometric data pre-processing involves the identification and elimination of spikes. The aforementioned task is achieved by the utilization of a fourth-difference analysis technique, wherein the radiometric field value at a specific location is substituted with the discrepancies observed between the values at the four closest neighboring places along a designated flight trajectory. In instances when the fourth difference exceeds the standard deviation of the dataset, an examination is conducted on the radiometric value to identify potential problems related to location or measurement. Following the removal of spikes, the data is subjected to interpolation through the utilization of a minimum curvature technique as outlined in reference [98]. The selection of the grid cell size is based on a fraction of the flight line spacing, specifically one-fourth (equivalent to 200 meters). By utilizing a specific cell size, the spectral composition of the original field is effectively conserved, hence preventing the occurrence of aliasing. The findings are depicted in Figure 7f, g, h, i.

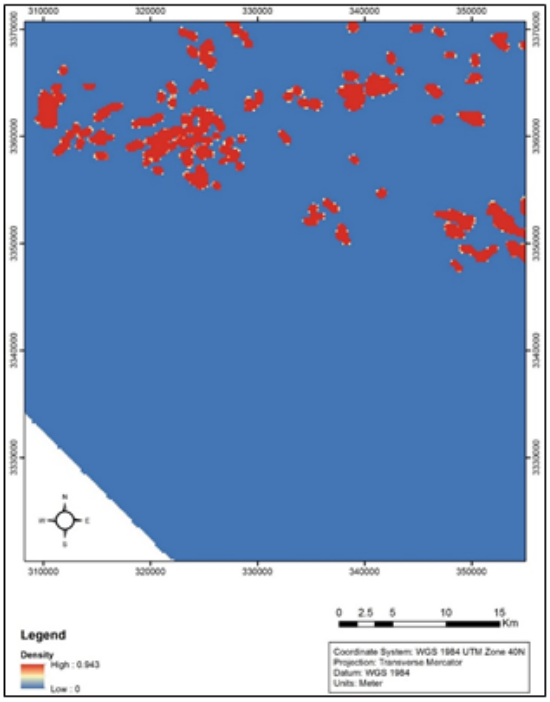
Deposition: The emplacement of mineralization is accompanied by the enrichment and depletion of specific element associations in a porphyry system (Sillitoe, 2010). Hence, mapping mineralization-related multi-element associations can help recognize the footprints of this style of mineralization [99, 100]. Multi-element maps of geochemical enrichments pertaining to the deposition of mineralization have long been applied to the context of exploration targeting [101, 102]. Stream sediment geochemical data is another dataset utilized for this investigation. This dataset is extracted from the GSI-published 1:100,000 sheet of the Shahr-e-Babak. This study employs Factor Analysis (FA) to determine the elemental associations associated with Cu mineralization. Initially, the data for selected geochemical elements Zn, Pb, Cr, Ni, Cu, As, Sb, Co, Sn, Ba, B, W, and Mo are transformed using the ilr transformation [103], which eliminates the effects of data closure. Next, a factor analysis (FA) with a varimax rotation is conducted on the ilr-transformed data. The final map of stream sediment geochemical data is prepared as an evidential map which is given in Figure 7j. Finally, the mineral deposit layer depicted in Figure 7k encompasses both mineral occurrences and small- to medium-sized mines.



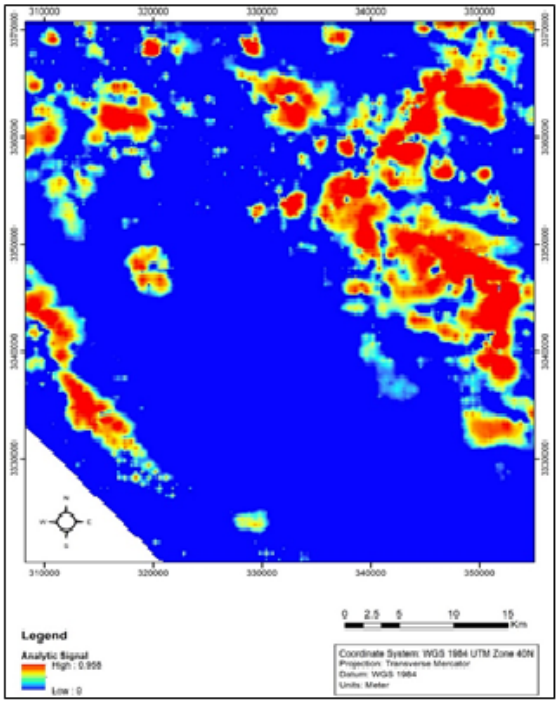
(a)



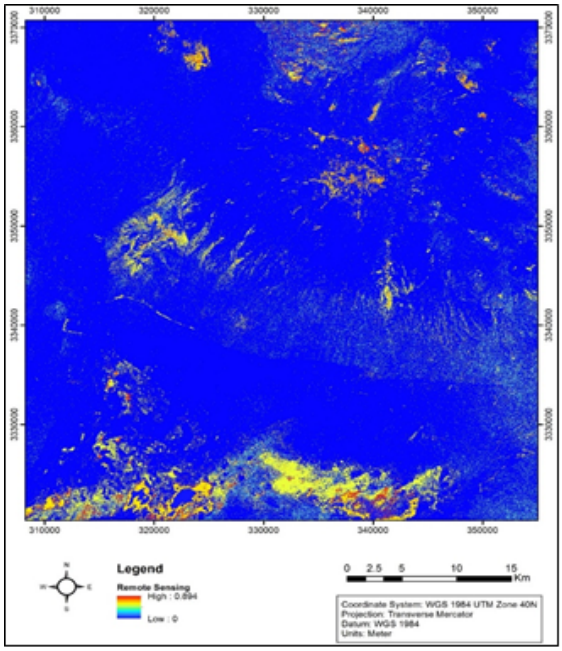
(b)



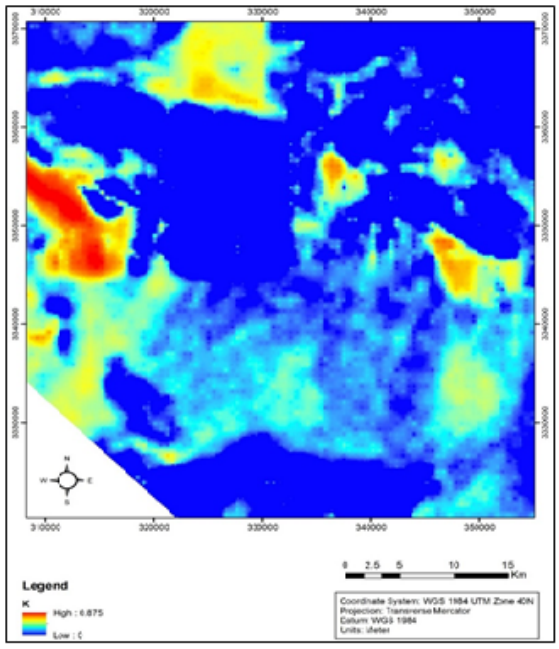
(c)



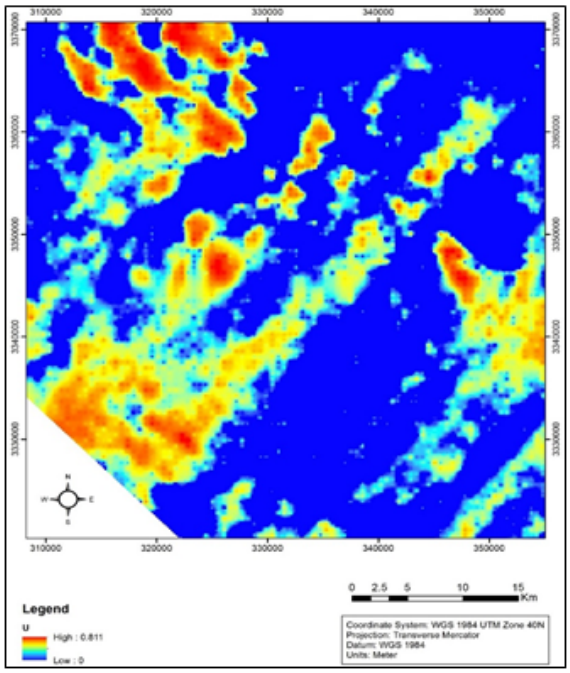
(d)



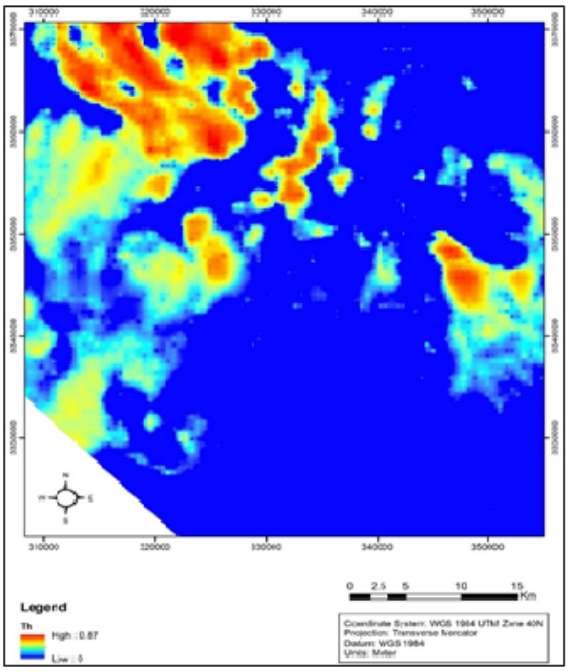
(e)



(f)



(g)



(h)

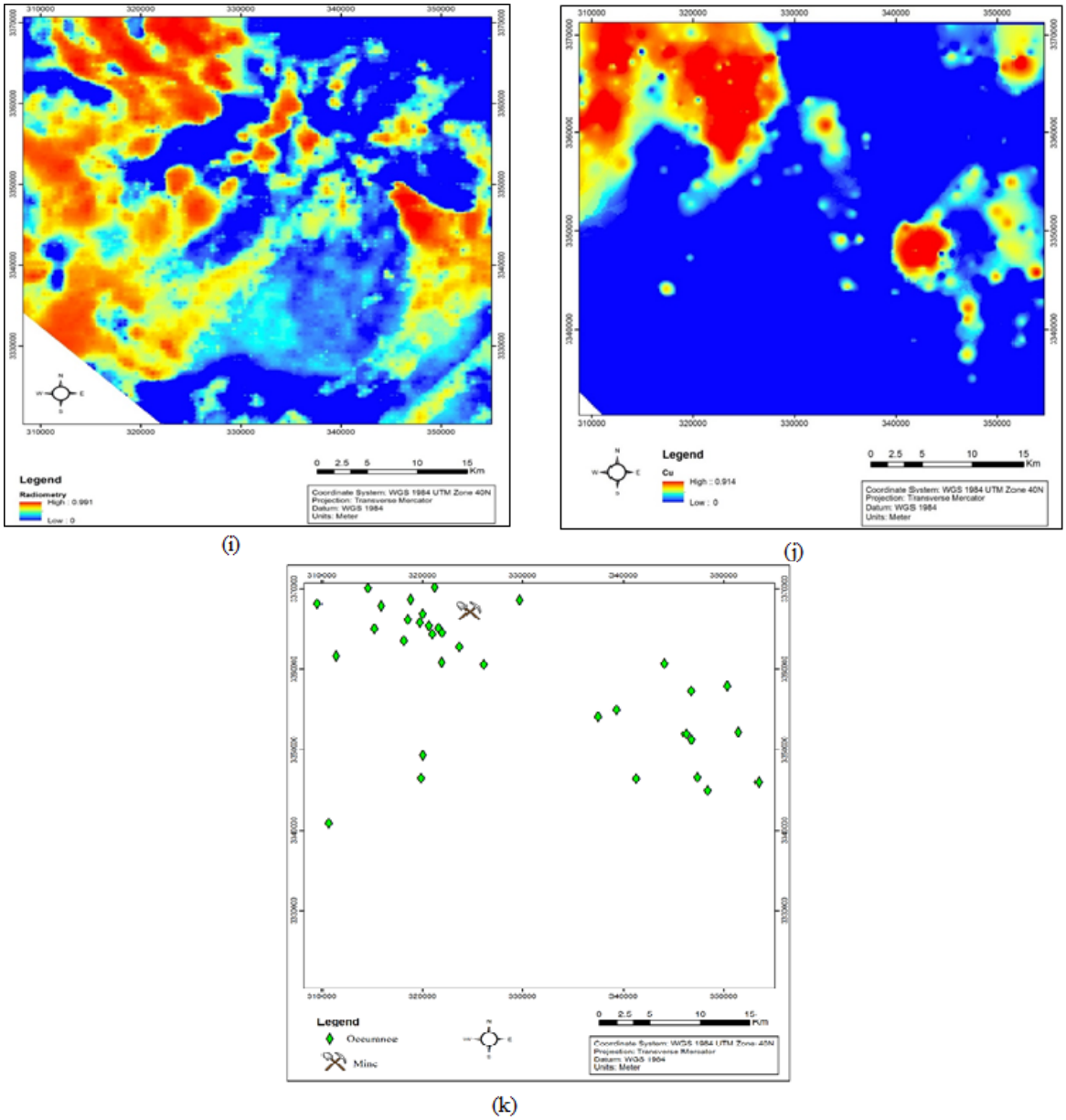


Figure 7. Evidential layers after processing the primary data. a) lithological map; b) fault density; c) dyke density; d) analytic signal map of the magnetic data; e) hydrothermal alterations extracted from the ASTER satellite data; f) radiometric data resulted from gamma spectrometry of f) potassium, g) uranium, h) thorium, i) total count, j) Geo-chemical interpolated map of the factor related to the Cu mineralization resulted from processing stream sediment samples, k) mineral-deposit locations.

All the evidential maps have been digitally superimposed with class scores. The statistical description of the data utilized in this paper, for both the input and output variables, is presented in the Table 2. The combination of all evidential maps results in the production of 62,086 feature vectors, each consisting of six dimensions. The operation is conducted within a Geographic Information System (GIS) framework, resulting in the automatic generation of a corresponding database. This database is responsible for storing the various components of the feature vectors. Target vectors are utilized in artificial neural networks (ANNs) to establish the desired output vectors that input feature vectors are mapped.

The process of training involves utilizing input feature vectors alongside their corresponding known target vectors, which together form training vectors. Validation vectors, which are often referred to as target vectors, are solely employed for the purpose of validating the training process of artificial neural networks

(ANNs). In the context of mineral potential mapping, a singular one-dimensional binary target vector is utilized, wherein the values are encoded as either 1 or 0. These values correspond to the presence or absence, respectively, of a target mineral deposit. The identification of the deposited samples is straightforward as these samples represent the feature vectors associated with the pre-identified locations of mineral deposits. However, the process of selecting non-deposited samples has certain challenges. In order to achieve this objective, strategies that are based on knowledge and driven by data can be employed. In the data-driven approach, non-deposited samples can be selected randomly from pre-modeled areas with low probability of hosting mined deposits. Additionally, in the knowledge-driven approach, the feature vectors of the points with extremely low probabilities are used as non-deposit training/validation vectors based on prior knowledge of the process of producing the target mineral deposit type. The training/validation vectors consist of feature vectors that represent the presence or absence of a target mineral deposit. These vectors are referred to as deposit or non-deposit training/validation vectors, respectively. In this paper, we used the well-explored parts of the area to select training and validation data. As a result, the Miduk site was chosen because it comprises 36 well-known occurrences and deposits (Figure 7k). Figure 8 depicts a schematic representation of the items discussed above.

Table 2. The comprehensive statistical description of the applied data in this study for both input and output.

	Layer	Evidential map	Number	Average	Minimum	Maximum	Standard deviation
Inputs	Lithological	Solid geology	62086	-	0.1	0.9	-
	Geo-chemical	Selected factor score map	62086	-0.098	-0.92	12.55	0.55
	Alteration	Presence and proximity to alteration haloes	62086	-	0	1	-
	Geo-physical	Total magnetic intensity	62086	16.06	-10190.81	15149.60	3128.60
		Potassium	62086	64	8.04	185.46	17.36
		Uranium	62086	208.23	23.39	445.25	55.32
		Thorium	62086	592.87	47.02	1884.92	193.08
		Total count	62086	856.38	149.02	2218.68	226.27
	Structural	Fault density	62086	0.191	0	5.55	0.53
		Dyke density	62086	0.059	0	4.32	0.259
output	Deposit/occurrence location	Mineral deposit location	62086	-	0	1	-

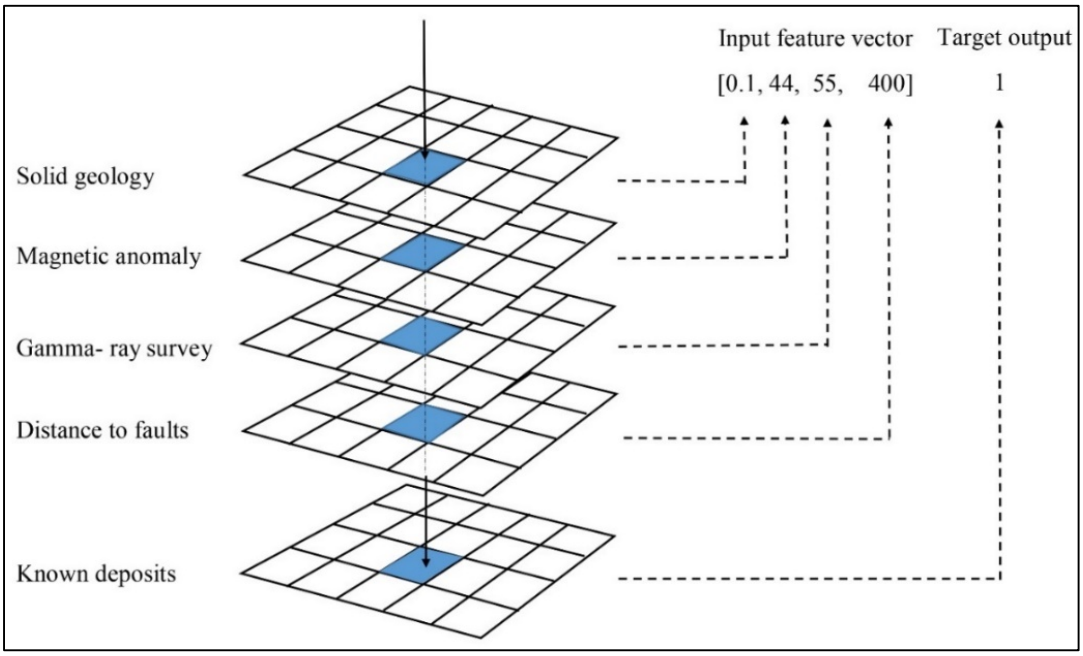


Figure 8. Association between GIS thematic layers and feature vectors applied as input to the ANN model. For each cell on the Shahr-e-babak grid, the components of the input feature vector for that location are set to the values stored in the thematic layers. Each pattern in the dataset used to train the network consists of an input feature vector paired together with the desired output (i.e. the value indicating the presence or absence of a deposit).

5.1.2. Data preprocessing

In the initial stages of ANN modeling, it is imperative to normalize evidential layers containing continuous values. The implementation of noise reduction techniques, input, and output scaling can contribute to improved prediction accuracy. The process of scaling the inputs and output variables prior to using an ANN is of utmost significance. The primary benefit lies in mitigating the dominance of attributes with larger numeric ranges over those with lower numeric ranges. Another advantage is that numerical complications are avoided throughout the calculation. Large attribute values may pose numerical problems because kernel values are frequently determined by the inner products of feature vectors, such as the linear kernel and the polynomial kernel [104, 105]. Over the course of recent decades, a range of normalization techniques have been examined in order to enhance the training of networks [106-108]. In this study, the normalization of each variable is conducted by implementing three distinct procedures in order to determine the most optimal and accurate approach. The applied methods encompass the utilization of various data normalization techniques. These techniques consist of the original data, the normalized data within the range of [-1, 1] achieved by employing the maximum and minimum values of the dataset, the normalized data obtained by utilizing the mean and standard deviation of the dataset, and lastly, the normalized data within the range of [0, 1] achieved through the application of the Equation (3):

$$X_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3)$$

Let x represent the input data that requires normalization, while x_{max} and x_{min} denote the maximum and minimum values of the original data, respectively. X_{norm} refers to the transformed data that has been normalized. The findings suggest that the utilization of Equation (3) for the purpose of normalization yields superior outcomes. Consequently, in the present study, the normalization of input and output data is accomplished by the utilization of this particular method. The assignment of classes to lithology groups is based solely on the discrete values of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 for the lithology layer. These values correspond to the 1st to 9th lithology groups, respectively.

5.1.3. Divisionism and performance indices

Divisionism is a crucial aspect in the field of ANN modelling, as it plays a significant role in ensuring the logical coherence and accuracy of the final model. In general, the data can be divided into three categories: 1) the training data refers to a distinct and unambiguous group of data points that are utilized in the training phase. 2) Test data is utilized after the training process, but their target is not clearly defined. 3) Validation data, which is utilized during the training process and serves to mitigate the issue of over fitting, is not inherently definitive. In this research, the entirety of the accessible datasets is randomly partitioned into three separate subsets, namely the training, validation, and testing data. The proportions of each of these subsets are 70%, 15%, and 15% of the whole number of the data (Table 3).

Table 3. Range of different parameters for train, test, and validation data sets.

Effective parameters			Train data (43460 samples)			Test data (9313 samples)			Validation data (9313 samples)		
			Min.[a]	Ave.[b]	Max.[c]	Min.	Ave.	Max.	Min.	Ave.	Max.
Input parameters	Lithological	Solid geology	0.1	-	0.9	0.1	-	0.9	0.1	-	0.9
	Geo-chemical	Selected factor score map	-0.86	-0.07	12.55	-0.88	-0.2	1.94	-0.924	-0.093	1.69
	Alteration	Presence and proximity to alteration haloes	0	-	1	0	-	1	0	-	1
	Geo-physical	Total magnetic intensity	-7296.73	158.03	11284.76	-10162.89	423.97	10975.31	-10190.81	-1054.27	15149.66
		Potassium	8.04	65.29	18.46	19.74	58.70	97.41	9.36	63.31	118.59
		Uranium	23.39	212.48	444.58	46.41	198.41	328.48	31.16	198.23	445.25
		Thorium	47.02	620.14	1884.92	53.96	525.33	1425.8	148.62	533.19	1198.73
		Total count	149.02	890.63	2218.68	288.43	765.99	1374.68	162.48	786.92	1713.62
	Structural	Fault density	0	0.141	5.55	0	0.384	4.92	0	0.231	3.58
		Dyke densiy	0	0.05	4.32	0	0.035	2.70	0	0.089	3.32
Output parameter	Deposit/occurrence location	Mineral deposit location	0	-	1	0	-	1	0	-	1

[a] Minimum, [b] Maximum and [c]Average

Mineral potential mapping states that there is just a single, binary target vector that is represented as 1 for the existence and 0 for the lack of a target mineral deposit. The training and validation vectors are built using these feature vectors. The terms "deposit" and "non-deposit" training/validation vectors, respectively, are used to describe these vectors. The input-output spatial relationship is discovered by the network by repeatedly analyzing the training set, which comprises the first subset and roughly 70% of the training/validation dataset. The validation and test sets (approximately 15% of the validation dataset and roughly 15% of the test dataset) make up the second and third subsets.

The evaluation of prediction models such as Artificial Neural Networks (ANNs) can be conducted by assessing their fitting accuracy. This assessment is performed after calibrating each model structure using both the training/validation dataset and the testing dataset. Additionally, the utilization of the Root Mean Squared Error (RMSE) and the coefficient of determination (R^2) can serve as performance metrics for evaluating the accuracy of the anticipated values generated by the models in relation to the desired output. These two indices are used in this study, that R^2 measures the degree of correlation between the observed and predicted values. The model's strength is assessed by establishing a correlation between the input and output variables. R^2 values range from 0 to 1, with 1 indicating a perfect fit between the data and the line drawn through them, and 0 representing no statistical correlation between the data and a line. The expression for R^2 is provided in Equation (4) as referenced by [109, 110]:

$$R^2 = 1 - \frac{\sum_{k=1}^{N}(t_k - y_k)^2}{\sum_{k=1}^{N}(t_k - \bar{t}_k)^2} \quad (4)$$

where t_k and y_k represent the target and network output, respectively, for the k^{th} input is the average of the targets, and N is the total number of events being examined.

The root mean square error (RMSE) is a statistical measure used to assess the variability of errors, independently of the size of the sample. It is mathematically represented by Equation (5) [109]:

$$RMSE = \sqrt{\frac{SEE}{N}} \quad (5)$$

The sum of the squared errors (SEEs) is calculated by summing the squared differences between the observed values and the predicted values. The number of data points utilized in the calculation is denoted by N. The expression for the calculation of SEE is provided in Equation (6) [111]:

$$SEE = \sum_{i=1}^{N}(y_i - \bar{y}_i)^2 \quad (6)$$

Given that the variables have been previously defined. Root Mean Square Error (RMSE) is a statistical metric used to quantify the level of disagreement between the observed data and the corresponding anticipated values. A perfect fit between observed and forecasted values would have an RMSE of 0. The most effective models are those that maximize R^2 while minimizing RMSE measurements. The analysis is iterated multiple times throughout the study.

5.1.4. ANN modelling

ANN models can be conceptualized as graphical depictions of parametric functions that accept a given set of input values and subsequently associate them with a corresponding set of output values. For the purpose of training, this study has used a feed forward-scaled conjugate gradient network and a feed forward-Levenberg-Marquardt network. In this study, the transfer function chosen for the ANN in the hidden layer neurons is the sigmoid function. The utilization of MATLAB is employed for the purposes of training and testing neural networks. Results of observed and predicted data obtained from the ANN is shown in the Tables 4. The best obtained network is the network with 8 neurons in the hidden layer.

Table 4. Results of observed and predicted data obtained from the ANN. The best obtained network is in bold.

Number of network	The number of neurons in the hidden layer	Training Algorithms	Train			Test			Validation		
			R^2	SSE	RMSE	R^2	SSE	RMSE	R^2	SSE	RMSE
1	2	SCG	0.66	1.71	0.61	0.53	1.79	0.66	0.55	1.79	0.68
2	3	SCG	0.68	1.68	0.55	0.59	1.74	0.62	0.57	1.71	0.59
3	4	LM	0.72	1.62	0.51	0.64	1.71	0.59	0.64	1.69	0.59
4	5	LM	0.79	1.59	0.49	0.72	1.66	0.55	0.73	1.65	0.56
5	6	LM	0.81	1.54	0.44	0.73	1.63	0.52	0.74	1.66	0.54
6	7	LM	0.87	1.41	0.43	0.78	1.59	0.47	0.78	1.59	0.47
7	8	LM	0.89	1.37	0.39	0.81	1.57	0.44	0.81	1.57	0.44
8	**10**	**LM**	**0.96**	**1.24**	**0.37**	**0.95**	**1.54**	**0.39**	**0.93**	**1.56**	0.38
9	12	SCG	0.90	1.46	0.063	0.84	1.56	0.43	0.84	1.57	0.45
10	14	SCG	0.87	1.51	0.072	0.81	1.59	0.45	0.81	1.62	0.42

The MLP architecture is well-recognized as one of the most ubiquitous and commonly used architectures in the field of Artificial Neural Networks (ANNs). This architectural design offers numerous advantages from various perspectives. Hence, the present study assesses the performance of the aforementioned network (as depicted in Figure 9) in order to establish a comparative analysis with the outcomes obtained by the application of GEP. The optimal configuration for a feed forward neural network entails the utilization of a hidden layer consisting of seven neurons, determined by a process of trial and error. Within this hidden layer, the neurons are equipped with sigmoid transfer functions. Figure 10 depicts the favorable map that was obtained through the implementation of the ANN technique.

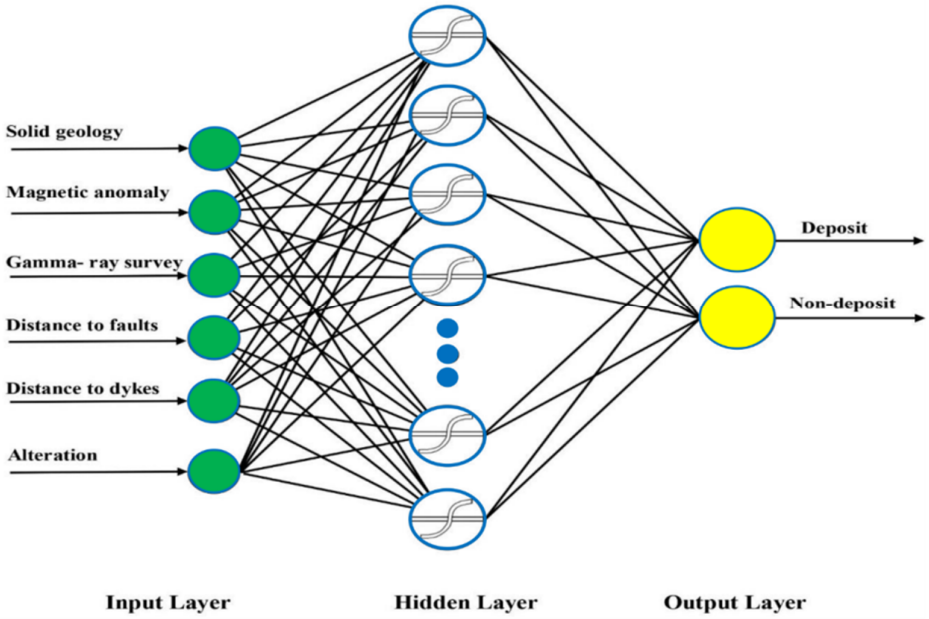


Figure 9. ANN architecture with sigmoid activation function for predicting Cu mineralization.

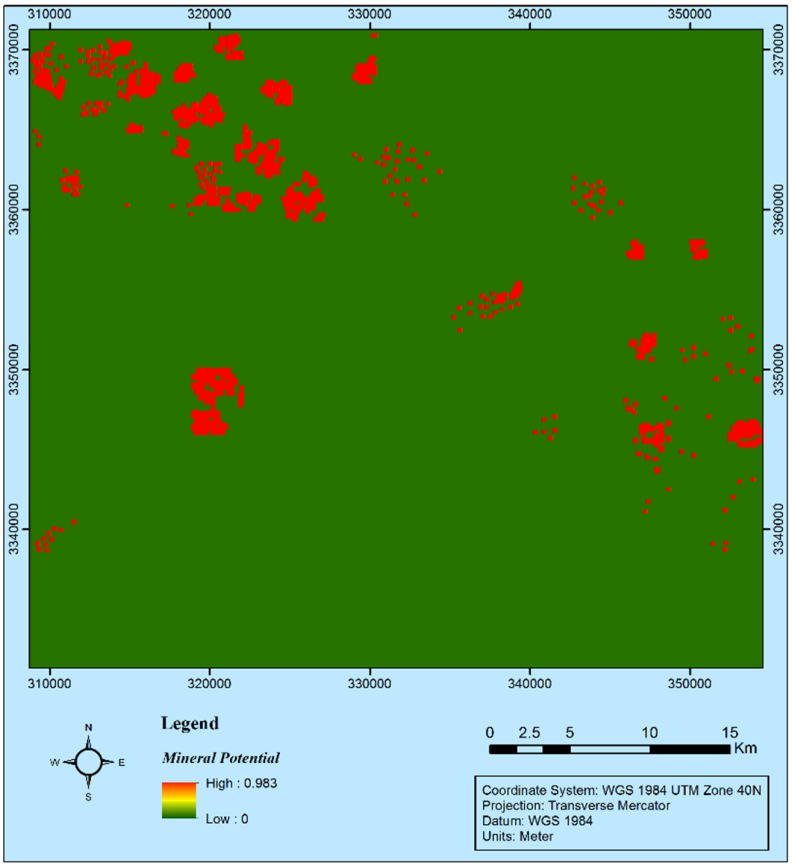


Figure 10. Favorable map resulted from ANN modeling.

5.2. GEP model results

The study involved an investigation of several combinations of GEP models developed by manipulating different GEP parameters such as chromosomal number, head size, gen number, linking function, and function set. The objective was to identify the most optimal GEP model that could effectively explain the MPM values, as presented in Table 5. The function sets utilized in GEP models are presented in Table 6.

Table 5. GEP models with applied parameters.

GEP model	Number of chrom.	Head size	Number of genes	Linking func.	Function set	R^2	Fitness
1	35	7	3	Addition	F4	0.8891	970.49
2	35	7	3	Multiplication	F4	0.8633	969.23
3	40	10	4	Addition	F1	0.8587	981.30
4	40	8	4	Addition	F1	0.8410	977.74
5	40	8	4	Multiplication	F1	0.8631	982.35
6	45	8	4	Addition	F1	0.8515	979.82
7	45	8	4	Multiplication	F1	0.8467	978.93
8	**50**	**10**	**4**	**Addition**	**F1**	**0.8985**	**989.01**
9	50	10	4	Multiplication	F1	0.8698	984.02
10	50	10	4	Addition	F2	0.8725	971.90
11	50	10	4	Addition	F3	0.8960	976.60
12	50	10	4	Addition	F4	0.8628	982.40
13	50	12	4	Addition	F1	0.8423	977.71
14	50	12	4	Multiplication	F1	0.8478	978.96
15	50	15	4	Addition	F1	0.8649	980.97

The study presented a compilation of GEP models characterized by varying characteristics, accompanied with statistical performance metrics including the correlation coefficient (R^2) and fitness. As expected, more chromosome and gene numbers increased accuracy of the GEP models. The linking functions of addition and multiplication demonstrated comparable outcomes. Based on the R^2 values, it can be concluded that the optimal GEP model is model 8. This model was also evaluated using various function sets in models 10-12. Statistical performance and parametric details of the selected model 8 is listed in Table 6. Figure 11 depicts the favorable map that was obtained through the implementation of the GEP technique. As shown in Figure 11, the GEP model is successful on generation of MPM. The calculated R^2 value is 0.9858, indicating statistical performance of the GEP model is acceptable.

Table 6. Parametric and statistical details of the selected GEP model.

Parameter	Value
chromosomes	50
Head size	10
Genes	4
Linking function	+
Function set	F1
Mutation rate	0.044
Inversion rate	0.1
Constant per gene	2
Independent variables	3
Training samples	17
Testing samples	4
Fitness function	RMSE
R^2 training	0.8985
Training fitness	989.01
R^2(testing)	0.8821
Testing fitness	963.09

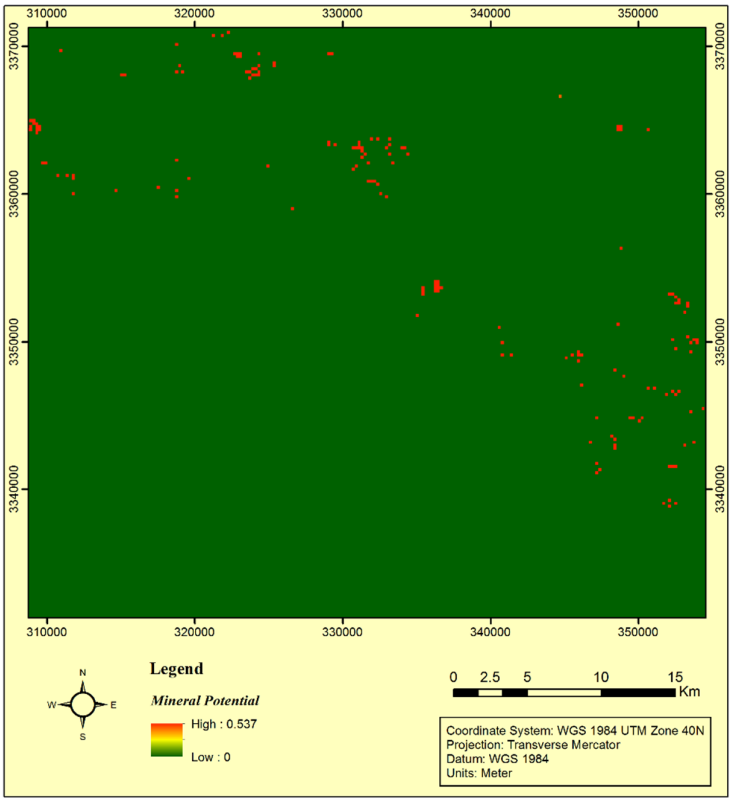


Figure 11. Favorable map resulted from GEP modeling.

5.3. Performance measures of ANN and GEP models

Two distinct models were created to predict the MPM using ANN and GEP approaches. The ANN and GEP models were subjected to analysis and subsequently compared with each other. The statistical comparison of the models employed for predicting the performance of MPM was conducted using the R^2 metric. The models underwent evaluation using several external model validation criteria to verify their accuracy. The ANN and GEP models were validated based on the evaluation of R^2 and RMSE metrics. The

ANN and GEP models have larger R^2 values, which is a predictable outcome owing to their greater complexity, as indicated in Table 7. It is widely recognized that the coefficient of determination (R^2) can be enhanced by incorporating additional terms into the model, hence increasing the complexity of the model. ANN models exhibited a reduced time need for construction. The process of identifying a GEP model with desirable features necessitates a comparatively greater amount of time.

Table 7. Comparison of ANN and GEP models.

Model	R^2 (Test)	R^2 (Train)	Complexity	Tim REquation
ANN	0.95	0.96	High	Low
GEP	0.88	0.89	High	High

6. Discussion

Mineral exploration is nearly impossible nowadays with only one information layer. In order to achieve this objective, it is imperative to take into account multiple data layers including geology, geo-physics, and geo-chemistry. The process of mineral potential mapping in Earth science incorporates both empirical and conceptual components. The empirical component involves the utilization of an exploration database, while the conceptual component relies on an expert knowledge-base [1, 2].

This study employed Artificial Neural Networks (ANNs) and GEP methodologies to integrate exploration data at a regional scale for the purpose of mapping copper potential in the Shahr-e-Babak area of Kerman. This location is known to possess significant potential for porphyry copper deposits. ANNs are empirical, data-driven methods which attempt to emulate features of biological neural networks (i.e. the human brain and nervous system) in order to address a range of difficult information processing, analysis and modeling problems [112]. Artificial neural networks possess strong capabilities in pattern recognition and classification, demonstrating the capacity to make generalizations from input data that may lack precision [113]. There are several benefits associated with the utilization of ANNs in the context of MPM: (i) in most cases, the modeling approach does not necessitate the use of a physics-based algorithm, resulting in a faster and more adaptable model construction process; (ii) ANNs are capable of effectively and efficiently handling non-linear relationships, and (iii) the model structure may easily incorporate the expertise and user experiences [114]. During the recent years, different types of ANN have been used more than any other methods in mineral potential mapping [29, 31]; however, there are several problems with ANN's training and designing. The majority of artificial neural networks (ANNs) often require a time-intensive process for designing their architecture. Moreover, these methods often exhibit poor learning rates and are prone to converging towards local minima or over fitting. In order to mitigate the issue of over-fitting, it is imperative to employ supplementary methodologies such as cross-validation and regularization. These techniques serve the purpose of providing insights into situations where continued training fails to yield improved generalization.

GEP is an artificial intelligence technique that utilizes an evolutionary algorithm to simulate a biological genetic process, drawing inspiration from the neutral theory of molecular evolution [115]. It serves as a potential substitute or supplement to existing genetic-based methodologies in computer programming, such as genetic algorithms (Gas), [116] and genetic programming (GP), [117]. The literature has shown that GEP is an effective method for automatically identifying and understanding the connections between target and response variables within a given dataset. This enables the construction of mathematical models that can accurately explain these relationships [118, 119]. Symbolic regression is a widely recognized approach in which GEP is utilized to construct a mathematical model that effectively elucidates the association between predictive and response variables, while minimizing the occurrence of errors.

Figures 12 and 13 display mineral prospectivity maps generated using the ANN and GEP techniques, respectively. These maps are accompanied by the inclusion of known copper occurrence and deposit sites. The ANN method produces geologically meaningful results. The outcomes obtained through the utilization of the ANN technique exhibit a high level of concurrence with the established mineral deposits, specifically Meiduk and Sara. The ANN technique has the capability to predict regions

exhibiting high, moderate, and low levels of favorability. The results of this method is compared to the GEP results. Both techniques exhibit a strong concurrence in identifying areas of high to moderate favorability, primarily located in the vicinity of the Meiduk and Sara deposits, as well as in the western regions of the Abdar deposit. The areas with moderate to low favorability are mostly scattered in the west, center and east to northeast of the study area. These areas are far from the well-known deposits and are good candidates for further exploration.

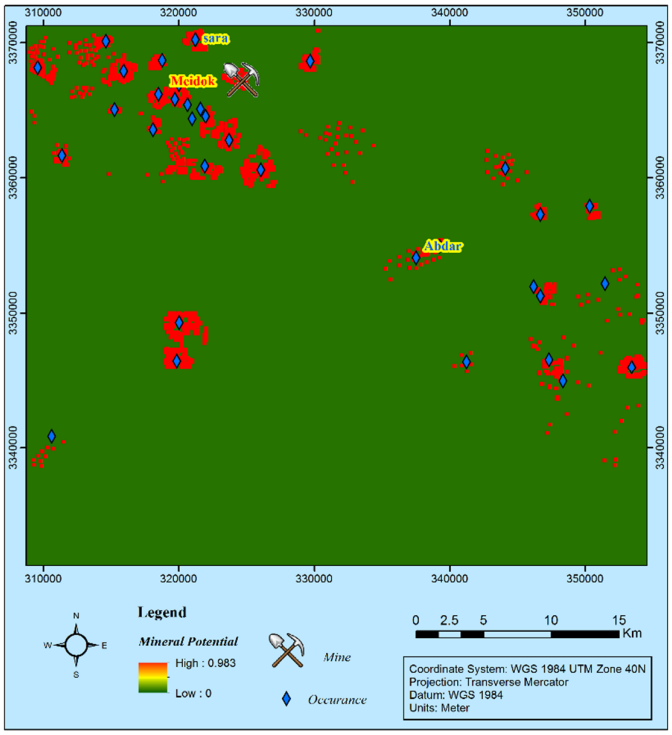


Figure 12. The mineral potential map created by ANN along with the occurrences and mines.

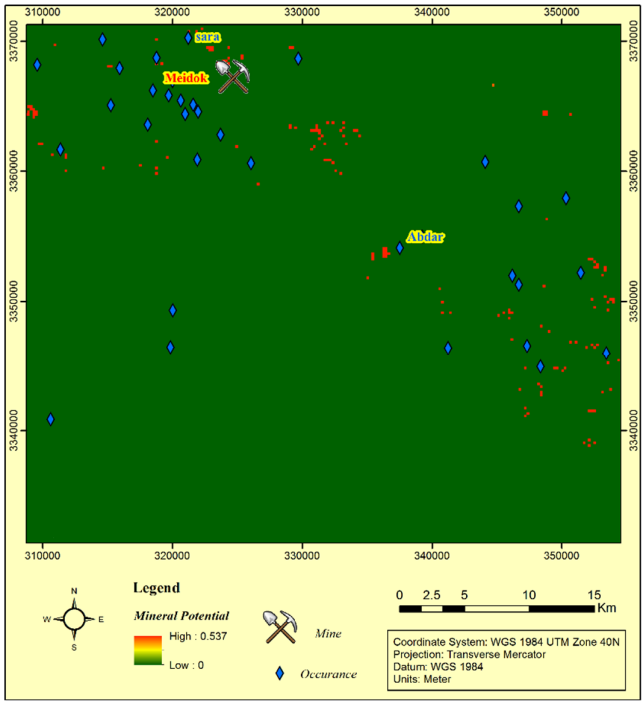
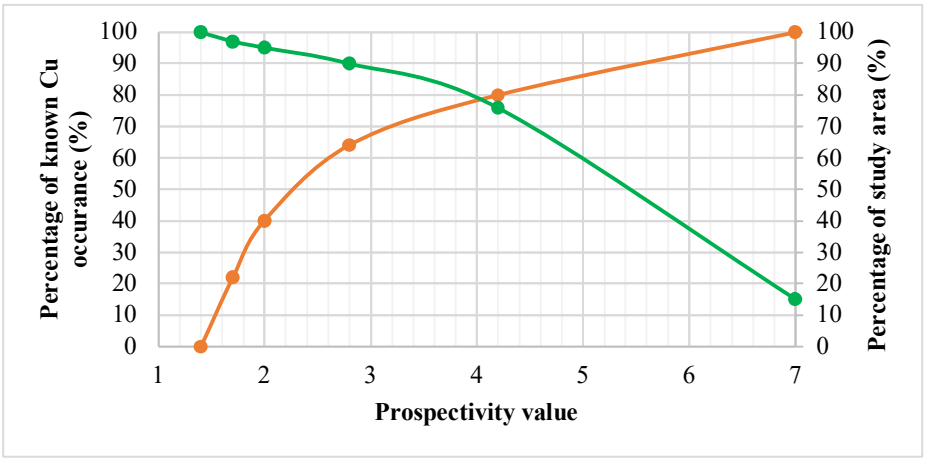
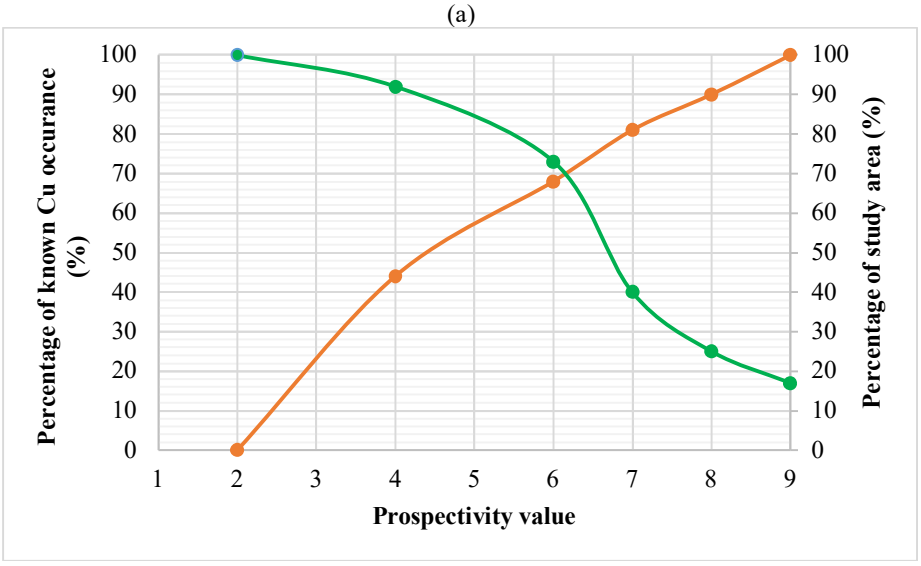


Figure 13. The mineral potential map generated with GEP along with the occurrences and mines.

This study employed the prediction-area (P-A) plot as a quantitative tool to validate the acquired results from the two methodologies. The P-A plot illustrates the relationship between the cumulative percentage of known occurrences predicted by integrating evidentiary classes and their associated cumulative occupied areas, relative to the entire examined area. This figure is constructed by plotting these values against the prospectivity values. Thus, the prediction ability of the evidence layer and its ability to delimit the studied area for further exploration are evaluated. Within a P-A plot, two distinct curves may be observed. The first curve represents the percentage, also known as the prediction rate, of known mineral occurrences that align with the various classes depicted on the prospectivity map. The second curve illustrates the proportion of occupied regions that match to the classes displayed on the prospectivity map. When an intersection point of the two curves is at a higher place in the P-A plot, it portrays a small area containing large number of mineral deposits. Furthermore, it chooses objectively a better model to give priority for mineral exploration [120, 121]. The significance of evaluating the predictability of prospectivity models is demonstrated by comparing the prediction rates depicted in Figure 14 of the study. The point of intersection on the P-A plot (Figure14a) of the ANN prospectivity model indicates that 80% of the identified copper occurrences are projected to be present in only 20% of the surveyed area. Similarly, the point of intersection on the P-A plot (Figure 14b) of the genetic expression programming (GEP) prospectivity model reveals that 70% of the known copper occurrences are estimated to be located within 30% of the studied area.



(a)



(b)

Figure 14. P-A plot for prospectivity model generated by integration of a) ANN, and b) GEP model.

The parameters of the intersection points in the P-A plots of the integrated maps are shown in Table 8. The analysis of the provided data reveals that the prospectivity models developed through the integration of ANN exhibit more efficiency compared to the prospectivity models built using GEP.

Table 8. Extracted parameters from intersection point of P-A plots.

Integration model	Prediction rate (%)	Occupied area (%)
ANN	80	20
GEP	70	30

7. Conclusions

Due to the importance of the MPM to minimize the cost and time as well as maximizing the benefit of a mineral exploration program and also decreasing the uncertainty caused by the unknown/complex geological variables, a mining engineer and exploration geologist should be able to apply a robust and transparent method to determine the favorable zones of mineralization for further exploration. For this purpose, two techniques of data-driven GEP and ANN were employed for predicting favorable areas associated with the Cu mineralization in the Shahr-e-Babak area, Kerman Province, SE Iran. We used solid lithology, tectonic (faults and dykes), airborne total magnetic intensity, airborne gamma-ray spectrometry data (U, Th, K and total count), remotely sensed data (hydrothermal alteration haloes), geochemical data, plus 34 deposits/occurrences representing known Cu mineralizations as the evidential maps. The results indicate that, in this particular application, the ANN method produces a plausible mineral prospectivity map from the geological aspect and superior to the GEP map. According to this study, the ANN technique with ten neurons in the hidden layer and Levenberg–Marquardt learning algorithm is a better predictor of Cu mineralization compared to the GEP method. The results obtained from the P-A plot showed that the ANN model indicates that 80% (vs. 70 % for GEP) of the identified copper occurrences are projected to be present in only 20% (vs. 30% for GEP) of the surveyed area. The ANN method is regarded as objective, because it is data-driven, and it is not necessary to select the weights based on the researcher's knowledge. In this paper, the artificial neural network data driven approach due to the capabilities such as classification, pattern matching, optimization and prediction shows better results than GEP technique so that, the areas with moderate to low favorability which are candidates for further exploration, are detected very well with this method while the GEP failed to identify them.

References

[1]. Zheng, C., Yuan, F., Luo, X., Li, X., Liu, P., Wen, M., & Albanese, S. (2023). Mineral prospectivity mapping based on support vector machine and random forest algorithm-A case study from Ashele copper-zinc deposit, Xinjiang, NW China. *Ore Geology Reviews*, 159, 105567.

[2]. Li, B., Yu, Z., & Ke, X. (2023). One-dimensional convolutional neural network for mapping mineral prospectivity: A case study in changba ore concentration area, Gansu province. *Ore Geology Reviews*, 160, 105573.

[3]. Saljoughi, B. S., & Hezarkhani, A. (2018). A comparative analysis of artificial neural network (ANN), wavelet neural network (WNN), and support vector machine (SVM) data-driven models to mineral potential mapping for copper mineralizations in the Shahr-e-Babak region, Kerman, Iran. *Applied Geomatics*, 10, 229-256.

[4]. Zuo, R. (2020). Geodata science-based mineral prospectivity mapping: A review. *Natural Resources Research*, 29, 3415-3424.

[5]. Xu, Y., Li, Z., Xie, Z., Cai, H., Niu, P., & Liu, H. (2021). Mineral prospectivity mapping by deep learning method in Yawan-Daqiao area, Gansu. *Ore Geology Reviews*, 138, 104316.

[6]. Chen, Y., & Wu, W. (2016). A prospecting cost-benefit strategy for mineral potential mapping based on ROC curve analysis. *Ore Geology Reviews,* 74, 26-38.

[7]. Ghezelbash, R., Maghsoudi, A., Bigdeli, A., & Carranza, E., J., M. (2021). Regional-scale mineral prospectivity mapping: Support vector machines and an improved data-driven multi-criteria decision-making technique. *Natural Resources Research*, 30, 1977-2005.

[8]. Ghezelbash, R., Maghsoudi, A., & Carranza, E., J., M. (2019). An improved data-driven multiple criteria decision-making procedure for spatial modeling of mineral prospectivity: adaption of prediction–area plot and logistic functions. *Natural Resources Research*, 28, 1299-1316.

[9]. Cheng, H., Zheng, Y., Wu, S., Lin, Y., Gao, F., Lin, D., & Chen, L. (2023). GIS-based mineral prospectivity mapping using machine learning methods: a case study from Zhuonuo ore district, Tibet. *Ore Geology Reviews*, 161, 105627.

[10]. Yousefi, M., & Carranza, E., J., M. (2015a). Geometric average of spatial evidence data layers: a GIS-based multi-criteria decision-making approach to mineral prospectivity mapping. *Computers & Geosciences*, 83, 72-79.

[11]. Daviran, M., Maghsoudi, A., Ghezelbash, R., & Pradhan, B. (2021). A new strategy for spatial predictive mapping of mineral prospectivity: Automated hyperparameter tuning of random forest approach. *Computers & Geosciences*, 148, 104688.

[12]. Lachaud, A., Adam, M., & Mišković, I. (2023). Comparative study of random forest and support vector machine algorithms in mineral prospectivity mapping with limited training data. *Minerals*, 13(8), 1073.

[13]. Acosta, I., C., C., Khodadadzadeh, M., Tusa, L., Ghamisi, P., & Gloaguen, R. (2019). A machine learning framework for drill-core mineral mapping using hyperspectral and high-resolution mineralogical data fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote* Sensing, 12(12), 4829-4842.

[14] Luo, X. & Dimitrakopoulos, R. (2003). Data-driven fuzzy analysis in quantitative mineral resource assessment. *Computers & Geosciences*, 29, 3-13.

[15]. Wei, H., Xiao, K., Shao, Y., Kong, H., Zhang, S., Wang, K., & Wen, C. (2020). Modeling-based mineral system approach to prospectivity mapping of stratabound hydrothermal deposits: A case study of MVT Pb-Zn deposits in the Huayuan area, northwestern Hunan province, China. *Ore Geology Reviews*, 120, 103368.

[16]. Xiong, Y., & Zuo, R. (2018). GIS-based rare events logistic regression for mineral prospectivity mapping. *Computers & Geosciences*, 111, 18-25.

[17]. Fu, C., Chen, K., Yang, Q., Chen, J., Wang, J., Liu, J., & Rajesh, H., M. (2021). Mapping gold mineral prospectivity based on weights of evidence method in southeast Asmara, Eritrea. *Journal of African Earth Sciences*, 176, 104143.

[18]. Zeghouane, H., Allek, K., & Kesraoui, M. (2016). GIS-based weights of evidence modeling applied to mineral prospectivity mapping of Sn-W and rare metals in Laouni area, Central Hoggar, Algeria. *Arabian Journal of Geosciences*, 9, 1-13.

[19]. Behera, S., & Panigrahi, M., K. (2022). Gold prospectivity mapping and exploration targeting in Hutti-Maski schist belt, India: Synergistic application of Weights-of-Evidence (WOE), Fuzzy Logic (FL) and hybrid (WOE-FL) models. *Journal of Geochemical Exploration*, 235, 106963.

[20]. Zhang, Z., Zuo, R., & Xiong, Y. (2016). A comparative study of fuzzy weights of evidence and random forests for mapping mineral prospectivity for skarn-type Fe deposits in the southwestern Fujian metallogenic belt, China. *Science China Earth Sciences*, 59, 556-572.

[21]. Porwal, A., Carranza, E. J., M., & Hale, M. (2001). Extended weights-of-evidence modelling for predictive mapping of base metal deposit potential in Aravalli province, western India. *Exploration and Mining Geology*, 10, 273-87.

[22]. Mohammadpour, M., Bahroudi, A., & Abedi, M. (2021). Three dimensional mineral prospectivity modeling by evidential belief functions, a case study from Kahang porphyry Cu deposit. *Journal of African Earth Sciences*, 174, 104098.

[23]. Carranza, E., J., M. (2015). Data-driven evidential belief modeling of mineral potential using few prospects and evidence with missing values. *Natural Resources Research*, 24, 291-304.

[24]. Geranian, H., Tabatabaei, S., H., Asadi, H., H., & Carranza, E., J., M. (2016). Application of discriminant analysis and support vector machine in mapping gold potential areas for further drilling in the Sari-Gunay gold deposit, NW Iran. *Natural Resources Research*, 25, 145-159.

[25]. Xie, S., Huang, N., Deng, J., Wu, S., Zhan, M., Carranza, E. J. M., ... and Meng, F. (2022). Quantitative prediction of prospectivity for Pb–Zn deposits in Guangxi (China) by back-propagation neural network and fuzzy weights-of-evidence modelling. Geochem.: *International Journal of Environmental Analytical* Chemistry, 22(2), 2021-085.

[26]. Porwal, A., Carranza, E., J., M., & Hale, M. (2006b). A hybrid fuzzy weights-of-evidence model for mineral potential mapping. *Natural Resources Research*, 15, 1-14.

[27]. Brown, W., M., Gedeon, T., Groves, D., & Barnes, R. (2000). Artificial neural networks: a new method for mineral prospectivity mapping. *Australian Journal of Earth Sciences*, 47, 757-70.

[28]. Skabar, A. (2007). Mineral potential mapping using Bayesian learning for multilayer perceptrons. *Mathematical Geosciences,* 39, 439-51.

[29]. Juliani, C., & Ellefmo, S., L. (2019). Prospectivity mapping of mineral deposits in northern Norway using radial basis function neural networks. *Minerals*, 9 (2), 131.

[30]. Leite, E., P., & de Souza Filho, C., R. (2009b). Probabilistic neural networks applied to mineral potential mapping for platinum group elements in the Serra Leste region, Carajás Mineral Province, Brazil. *Computers & Geosciences*, 35, 675-87.

[31]. Wang, J., Zuo, R., & Xiong, Y. (2020). Mapping mineral prospectivity via semi-supervised random forest. *Natural Resources Research*, 29, 189-202.

[32]. Chen, Y. (2015). Mineral potential mapping with a restricted Boltzmann machine. *Ore Geology Reviews*, 71, 749-60.

[33]. Zhang, N., Zhou, K., & Du, X. (2017). Application of fuzzy logic and fuzzy AHP to mineral prospectivity mapping of porphyry and hydrothermal vein copper deposits in the Dananhu-Tousuquan island arc, Xinjiang, NW China. *Journal of African Earth Sciences*, 128, 84-96.

[34]. Boadi, B., Raju, P. S., & Wemegah, D., D. (2022). Analysing multi-index overlay and fuzzy logic models for lode-gold prospectivity mapping in the Ahafo gold district–Southwestern Ghana. *Ore Geology Reviews*, 148, 105059.

[35]. Yousefi, M., & Carranza, E., J., M. (2016). Data-driven index overlay and Boolean logic mineral prospectivity modeling in greenfields exploration. *Natural Resources Research*, 25, 3-18.

[36]. Carranza, E., J., M. (2010). Improved wildcat modelling of mineral prospectivity. *Resource Geology*, 60, 129-49.

[37]. Ghaeminejad, H., Abedi, M., Afzal, P., Zaynali, F., & Yousefi, M. (2020). A fractal-based outranking approach for integrating geo-chemical, geological, and geo-physical data. *Bollettino Di Geofisica Teorica Ed Applicata*, 61(4), 555-588.

[38]. Cheng, Q. & Agterberg, F. (1999). Fuzzy weights of evidence method and its application in mineral potential mapping. *Natural Resources Research*, 8, 27-35.

[39]. Chudasama, B., Torppa, J., Nykänen, V., Kinnunen, J., Lerssi, J., & Salmirinne, H. (2022). Target-scale prospectivity modeling for gold mineralization within the rajapalot Au-Co project area in northern Fennoscandian Shield, Finland. Part 1: application of knowledge-driven-and machine learning-based-hybrid-expert systems for exploration targeting and addressing model-based uncertainties. *Ore Geology Reviews*, 147, 104937.

[40]. Yousefi, M., Carranza, E., J., M. (2015b). Fuzzification of continuous-value spatial evidence for mineral prospectivity mapping. *Computers & Geosciences,* 74, 97-109.

[41]. Yousefi, M., & Carranza, E., J., M. (2016). Data-driven index overlay and Boolean logic mineral prospectivity modeling in greenfields exploration. *Natural Resources Research*, 25, 3-18.

[42]. Mahdiyanfar, H., and Seyedrahimi-Niaraq, M. (2023). Integration of Fractal and Multivariate Principal Component Models for Separating Pb-Zn Mineral Contaminated Areas. *Journal of Mining and Environment*, 14(3), 1019-1035.

[43]. Mahdiyanfar, H., & Seyedrahimi-Niaraq, M., (2022). Improvement of geochemical prospectivity mapping using power spectrum–area fractal modelling of the multi-element mineralization factor (SAF-MF). *Geochemistry: Exploration, Environment, Analysis*, 22(4), 2022-015.

[44]. Seyedrahimi-Niaraq, M., & Mahdiyanfar, H. (2021). Introducing a new approach of geochemical anomaly intensity index (GAII) for increasing the probability of exploration of shear zone gold mineralization. *Geochemistry*, 81(4), 125830.

[45]. Lei, Y., Gan, Q., Du, Y., Liang, Y., Wang, G., Zheng, Y., & Xu, G. (2010). Mineral potential mapping based on GIS technology and fractal method. *The 2nd Conference on Environmental Science and Information Application Technology*, 2, 335-338.

[46]. Yousefi, M., Kreuzer, O.P., Nyk¨anen, V., and Hronsky, J.M.A. (2019). Exploration information systems—a proposal for the future use of GIS in mineral exploration targeting. *International Geology Review*, 111, 103005.

[47]. Yousefi, M., Carranza, E. J. M., Kreuzer, O. P., Nykänen, V., Hronsky, J. M., and Mihalasky, M. J. (2021). Data analysis methods for prospectivity modelling as applied to mineral exploration targeting: State-of-the-art and outlook. *Journal of Geochemical Exploration*, 229, 106839.

[48]. Yousefi, M., & Hronsky, J. M. (2023). Translation of the function of hydrothermal mineralization-related focused fluid flux into a mappable exploration criterion for mineral exploration targeting. *Applied Geochemistry*, 149, 105561.

[49]. Geological Survey of Iran, (1973). Exploration for ore deposits in Kerman region, Report no. Yu/53.

[50]. Dimitrijevic, M.D. (1973). Geology of the Kerman region, Report no. Yu/52. Geological Survey of Iran, Iran.

[51]. Hezarkhani, A. (2008). Hydrothermal evolution of the Miduk porphyry copper system, Kerman, Iran: a fluid inclusion investigation. *International Geology Review*, 50(7), 665-684.

[52]. Haykin, S. (1994). Neural networks: a comprehensive foundation: Prentice Hall PTR.

[53]. Pattanayak, S., Loha, C., Hauchhum, L., & Sailo, L. (2021). Application of MLP-ANN models for estimating the higher heating value of bamboo biomass. *Biomass Conversion and Biorefinery*, 11, 2499-2508.

[54]. Gandomi, A. H., & Roke, D. A. (2015). Assessment of artificial neural network and genetic programming as predictive tools. *Advanced Engineering*, 88, 63-72.

[55]. Miljković, D. (2017). Brief review of self-organizing maps. *In 2017 40th international convention*

on information and communication technology, electronics and microelectronics (MIPRO), 1061-1066.

[56]. Rojas, R. (2013). Neural networks: a systematic introduction: *Springer Science & Business Media.*

[57]. Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology (IJEIT)*, 2(1), 189-194.

[58]. Fausett, L. V. (2006). Fundamentals of neural networks: architectures, algorithms and applications. *Pearson Education India*.

[59]. Koza, J.R. (1992). Genetic Programming: on the Programming of Computers by Means of Natural Selection, *MIT Press*.

[60]. Afradi, A., & Ebrahimabadi, A. (2020). Comparison of artificial neural networks (ANN), support vector machine (SVM) and gene expression programming (GEP) approaches for predicting TBM penetration rate. *SN Applied Sciences*, 2, 1-16.

[61]. Bagatur, T., & Onen, F. (2014). A predictive model on air entrainment by plunging water jets using GEP and ANN. *KSCE Journal of Civil Engineering*, 18, 304-314.

[62]. Xiao, F., Chen, W., Wang, J., and Erten, O. (2021). A hybrid logistic regression: gene expression programming model and its application to mineral prospectivity mapping. *Natural Resources Research*, 31, 1-24.

[63]. Sivanandam, S.N., & Deepa, S.N. (2008). Introduction to Genetic Algorithms, *Springer*, 1-455.

[64]. Koza, J.R. (1992). Genetic Programming: on the Programming of Computers by Means of Natural Selection, *MIT Press*.

[65]. Karahan, İ. H., & Özdemir, R. (2010). A new modeling of electrical resistivity properties of Zn-Fe alloys using genetic programming. *Journal of Optoelectronics and Advanced Materials*, 4(6), 812-815.

[66]. Ozdemir, R., & Karahan, I. H. (2015). Grain size calculation of Cu-Zn alloys using genetic programming; an alternative for Scherer's formula. *Journal of Optoelectronics and Advanced Materials*, 17, 14-26.

[67]. Hagemann, S.G., Lisitsin, V.A., & Huston, D.L. (2016a). Mineral system analysis: Quo vadis. *Ore Geology Reviews*, 76, 504–522.

[68]. Carranza, E.J.M. (2017). Natural resources research publications on geochemical anomaly and mineral potential mapping, and introduction to the special issue of papers in these fields. *Natural Resources Research*, 26 (4), 379–410.

[69]. McCuaig, T.C., Beresford, S., Hronsky, J. (2010). Translating the mineral systems approach into an effective exploration targeting system. *Ore Geology Reviews,* 38, 128–138.

[70]. Knox-Robinson, C.M., & Wyborn, L.A.I. (1997). Towards a holistic exploration strategy: using geographic information systems as a tool to enhance exploration. Australian Journal of Earth Sciences, 44, 453–463

[71]. Vigneresse, J.L., Truche, L., & Richard, A. (2019). How do metals escape from magmas to form porphyry-type ore deposits? *Ore Geology Reviews*, 105, 310–336.

[72]. Chiaradia, M. (2020). Gold endowments of porphyry deposits controlled by precipitation efficiency. *Nature Communications*, 11, 1–10.

[73]. Tosdal, R.M., and Richards, J.P. (2001). Magmatic and structural controls on the development of porphyry Cu±Mo±Au deposits. *Reviews In Economic Geology*, 14, 157–181.

[74]. Arndt, N., Kesler, S., & Ganino, C. (2015). Metals and Society: An Introduction to Economic Geology. *Springer*.

[75]. Dimitrijevic, M.D., Dimitrijevic, M.N., Djordjevic, M., & Vulovic, D. (1971). Geological Map of Pariz, Scale 1: 100,000. *Geological Survey of Iran, Tehran*.

[76]. Atapour, H., & Aftabi, A. (2007). The geochemistry of gossans associated with Sarcheshmeh porphyry copper deposit, Rafsanjan, Kerman, Iran: implications for exploration and the environment. *Journal of Geochemical Exploration*, 93, 47–65.

[77]. Boomeri, M., Nakashima, K., & Lentz, D.R. (2010). The Sarcheshmeh porphyry copper deposit, Kerman, Iran: a mineralogical analysis of the igneous rocks and alteration zones including halogen element systematics related to Cu mineralization processes. *Ore Geology Reviews*, 38, 367–381.

[78]. Aftabi, A., & Atapour, H. (2011). Alteration geochemistry of volcanic rocks around Sarcheshmeh porphyry copper deposit, Rafsanjan, Kerman, Iran: implications for regional exploration. *Resource Geology*, 61, 76–90.

[79]. Zhao, J., Chen, S., Zuo, R., & Carranza, E.J.M. (2011). Mapping complexity of spatial distribution of faults using fractal and multifractal models: vectoring towards exploration targets. *Computers & geosciences*, 37, 1958-1966.

[80]. Kim, Y. S., & Sanderson, D.J. (2005). The relationship between displacement and length of faults: a review. *Earth Science Reviews*, 68, 317-334.

[81]. Faulkner, D., Jackson, C., Lunn, R., Schlische, R., Shipton, Z., Wibberley, C., & Withjack, M. (2010). A review of recent developments concerning the structure, mechanics and fluid flow properties of fault zones. *Journal of Structural Geology*, 32, 1557-1575.

[82]. Micklethwaite, S., Sheldon, H.A., & Baker, T. (2010). Active fault and shear processes and their implications for mineral deposit formation and discovery. *Journal of Structural Geology*, 32, 151-165.

[83]. Torabi, A., & Berg, S.S. (2011). Scaling of fault attributes: A review. *Marine and Petroleum Geology*, 28, 1444-1460.

[84]. Tchalenko, J. (1970). Similarities between shear zones of different magnitudes. *Geological Society of America Bulletin*, 81, 1625-1640.

[85]. Agterberg, F., Cheng, Q., Brown, A., & Good, D. (1996). Multifractal modeling of fractures in the Lac du Bonnet batholith, Manitoba. *Computers & geosciences*, 22, 497-507.

[86]. Mandelbrot, B.B. (1983). The fractal geometry of nature. *Macmillan*.

[87]. Fry, N. (1979). Random point distributions and strain measurement in rocks. *Tectonophysics*, 60, 89-105.

[88]. Vearncombe, J., & Vearncombe, S. (1999). The spatial distribution of mineralization; applications of Fry analysis. *Economic Geology*, 94, 475-486.

[89]. Vearncombe, S., & Vearncombe, J.R. (2002). Tectonic controls on kimberlite location, southern Africa. *Journal of Structural Geology*, 24, 1619-1625.

[90]. Stubley, M. (2004). Spatial distribution of kimberlite in the Slave craton, Canada: a geometrical approach. *Lithos*, 77, 683-693.

[91]. Casas, A.M., Cortes, A.L., Maestro, A., Soriano, M.A., Riaguas, A., & Bernal, J. (2000). LINDENS: a program for lineament length and density analysis. *Computers & geosciences*, 26, 1011-1022.

[92]. Hardcastle, K.C. (1995). Photolineament factor: a new computer-aided method for remotely sensing the degree to which bedrock is fractured. *Photogrammetric Engineering and Remote Sensing*, 61, 739-747.

[93]. Ranjbar, H., & Roonwal, G.S. (2002). Digital image processing for lithological and alteration mapping using SPOT multispectral data: a case study of Pariz area, Kerman Province, Iran, International Symposium on Remote Sensing. I*nternational Society for Optics and Photonics*, 207-215.

[94]. Carranza, E.J.M. (2009b). Objective selection of suitable unit cell size in data-driven modeling of mineral prospectivity. *Computers & geosciences*, 35, 2032-2046.

[95]. Pirajno, F. (2012). Hydrothermal Mineral Deposits: Principles and Fundamental Concepts for the Exploration Geologist. *Springer Science & Business Media*.

[96]. Pour, A.B., & Hashim, M. (2012). The application of ASTER remote sensing data to porphyry copper and epithermal gold deposits. *Ore Geology Reviews*, 44, 1–9.

[97]. Tangestani, M., & Moore, F. (2003). Mapping porphyry copper potential with a fuzzy model, northern Shahr-e-Babak, Iran. *Australian Journal of Earth Sciences*, 50, 311-7.

[98]. Briggs, I.C. (1974). Machine contouring using minimum curvature. *Geophysics*, 39, 39-48.

[99]. Parsa, M., Maghsoudi, A., Yousefi, M., & Sadeghi, M. (2016a). Prospectivity modeling of porphyry-Cu deposits by identification and integration of efficient mono-elemental geochemical signatures. *Journal of African Earth Sciences*, 114, 228–241.

[100]. Parsa, M., Maghsoudi, A., Carranza, E.J.M., & Yousefi, M. (2017c). Enhancement and mapping of weak multivariate stream sediment geochemical anomalies in Ahar Area, NW Iran. *Natural Resources Research*, 26, 443–455.

[101]. Harris, J.R., Grunsky, E., Behnia, P., & Corrigan, D. (2015). Data-and knowledge-driven mineral prospectivity maps for Canada's North. *Ore Geology Reviews*, 71, 788–803.

[102]. Sadeghi, M., Morris, G.A., Carranza, E.J.M., Ladenberger, A., & Andersson, M. (2013). Rare earth element distribution and mineralization in Sweden: an application of principal component analysis to FOREGS soil geochemistry. *Journal of Geochemical Exploration*, 133, 160–175.

[103]. Aitchison, J. (1986). The statistical analysis of compositional data, *John Wiley*.

[104]. Mahmoudabadi H., Izadi M., & Menhaj M.B. (2009). A hybrid method for grade estimation using genetic algorithm and neural networks. *Computers & Geosciences*, 13, 91-101.

[105]. Li X-l., Li, L-h,, Zhang, B-l., & Guo, Q-j. (2013). Hybrid self-adaptive learning based particle swarm optimization and support vector regression model for grade estimation. *Neurocomputing*, 118,179-90.

[106]. Chaturvedi, D., Satsangi, P., & Kalra, P. (1996). Effect of different mappings and normalization of neural network models. *Proceedings of the National Power Systems Conference: Indian Institute of Technology*, 377-86.

[107]. Sola, J., & Sevilla, J. (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Transactions on Nuclear Science*, 44, 1464-8.

[108]. Demuth, H.B., Beale, M.H., De Jess, O., & Hagan, M.T. (2014). Neural network design. *Martin Hagan*.

[109]. Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An artificial neural network model for flood simulation

using GIS: Johor River Basin, Malaysia. *Environmental Earth Sciences*, 67, 251–264.

[110]. Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20, 557-585.

[111]. Singla, P., Duhan, M., & Saroha, S. (2022). Different normalization techniques as data preprocessing for one step ahead forecasting of solar global horizontal irradiance. In Artificial Intelligence for Renewable Energy Systems. *Woodhead Publishing*, 209-230.

[112]. Smith, De., Goodchild, M.J., & Longley, P.A. (2007). Geospatial analysis - a comprehensive guide to principles, techniques and software tools, third edition, *Troubador Publishing.*

[113]. Parsa, M., Carranza, E. J. M., & Ahmadi, B. (2021). Deep GMDH neural networks for predictive mapping of mineral prospectivity in terrains hosting few but large mineral deposits. *Natural Resources Research*, 1-14.

[114]. Chen, G., Huang, N., Wu, G., Luo, L., Wang, D., & Cheng, Q. (2022). Mineral prospectivity mapping based on wavelet neural network and Monte Carlo simulations in the Nanling W-Sn metallogenic province. *Ore Geology Reviews*, 143, 104765.

[115]. Ferreira, C. (2006b). Gene expression programing: mathematical modeling by an artificial intelligence. *Springer*.

[116]. Goldberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning. *Kluwer*.

[117]. Koza, J. R. (1992). Genetic programming: On the programming of computers by means of natural selection. *MIT Press*.

[118]. Dey, P., & Das, A. K. (2016). A utilization of GEP (gene expression programming) metamodel and PSO (particle swarm optimization) tool to predict and optimize the forced convection around a cylinder. *Energy*, 95, 447–458.

[119]. Goharzay, M., Noorzad, A., Ardakani, A. M., & Jalal, M. (2017). A worldwide SPT-based soil liquefaction triggering analysis utilizing gene expression programming and Bayesian probabilistic method. *Journal of Rock Mechanics and Geotechnical Engineering*, 9, 683–693.

[120]. Parsa, M., Maghsoudi. A., Yousefi, M., and Sadeghi, M. (2016). Prospectivity modeling of porphyry-Cu deposits by identification and integration of efficient mono-elemental geochemical signatures. *Journal of African Earth Science*s,114, 228–241.

[121]. Yousefi, M., & Carranza, E.J.M. (2015c). Prediction-area (P-A) plot and C-A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling. *Computers & Geosciences*, 79, 69–81.

آنالیز مقایسه‌ای مدل‌های داده‌محور شبکه عصبی مصنوعی و برنامه‌نویسی بیان ژن برای پتانسیل‌های نواحی امیدبخش کانی‌زایی مس؛ مطالعه موردی: محدوده شهربابک، استان کرمان، جنوب شرق ایران

بشیر شکوه سلجوقی*[1] و اردشیر هزارخانی[2]

۱. بخش مهندسی معدن، دانشگاه امیرکبیر تهران، ایران
۲. بخش مهندسی معدن، دانشگاه امیرکبیر تهران، ایران

ارسال ۲۰۲۳/۱۱/۱۶، پذیرش ۲۰۲۳/۱۲/۱۴

* نویسنده مسئول مکاتبات: bashir.shokouh@gmail.com

چکیده:

ناحیه پرپتانسیل مس پورفیری مورد مطالعه در این مقاله در بخش جنوبی کمپلکس آتشفشانی-رسوبی ایران مرکزی واقع شده است و حاوی تعداد زیادی نهشته و رخدادهای معدنی است که در حال حاضر بت کمبود منابع معدنی مواجه شده است. بنابراین، پی‌جویی نواحی پرپتانسیل در نواحی عمیق‌تر و پیرامونی دارای اهمیت بالایی در این منطقه است. روش‌های مستقیم و غیرمستقیم تلاش دارند تا نواحی امیدبخش را برای اکتشافات آتی برآورد نمایند که اغلب آنها زمانبر و پرهزینه هستند. هدف اصلی روش‌های پی‌جویی مواد معدنی، بکاربردن روش‌های صریح و معتبر برای شناسایی محدوده‌های پرپتانسیل کشف شده در آینده است. مطالعه حاضر، دو روش را برای ایجاد نقشه‌های امیدبخش مس پورفیری ارائه می‌نماید. این مطالعه نتایج تکنیک‌های شبکه عصبی مصنوعی و برنامه‌نویسی بیان ژن را مقایسه می‌کند. پایگاه داده بکاربرده شده برای ایجاد نقشه‌های شاهد کانی‌زایی مس پورفیری شامل نقشه‌های زمین‌شناسی، دگرسانی، گسل‌ها، دایک‌ها، شدت میدان مغناطیسی کل هوابرد، داده‌های اسپکترومتری اشعه گاما هوابرد (اورانیوم، توریوم و محتوای کل) و رخدادهای شناخته شده مس است. براساس این مطالعه، روش شبکه عصبی مصنوعی (با ده نرون در لایه مخفی و الگوریتم یادگیری لونبرگ- مارکوادت) تخمین‌گر بهتری از کانی‌زایی مس پورفیری را در مقایسه با روش برنامه‌نویسی بیان ژن فراهم می‌آورد. نتایج بدست آمده از نمودار P-A نشان می‌دهد که مدل شبکه عصبی مصنوعی ۸۰ درصد (در مقابل ۷۰ درصد برای الگوریتم برنامه‌نویسی بیان ژن) رخدادهای مس پورفیری شناسایی شده را تنها در ۲۰ درصد (۳۰ درصد برای الگوریتم برنامه‌نویسی بیان ژن) محدوده بررسی شده ارائه می‌نماید. تکنیک شبکه عصبی مصنوعی با توجه به قابلیت‌هایی مانند طبقه‌بندی، تطابق الگو، بهینه‌سازی و پیش‌بینی در شناسایی آنومالی‌های مرتبط با کانی‌زایی مس سودمند است.

کلمات کلیدی: تهیه نقشه پتانسیل معدنی، شبکه عصبی مصنوعی، برنامه‌ریزی بیان ژن، کانی‌زایی مس.