



## Intelligent Borehole Simulation with python Programming

Hassanreza Ghasemitabar<sup>1\*</sup>, Andisheh Alimoradi<sup>2</sup>, Hamidreza Hemmati Ahoori<sup>2</sup>, Mahdi Fathi<sup>2</sup> and Mahshid Sarookhani<sup>3</sup>

1. Faculty of Mining, Petroleum & Geophysics Eng., Shahrood University of Technology, Shahrood, Iran.

2. Department of Mining Eng., Faculty of Eng., Imam Khomeini International University, Qazvin, Iran.

3. Department of Petroleum and Sedimentary Basins, Faculty of Earth Sci., Shahid Beheshti University, Tehran, Iran.

### Article Info

Received 26 September 2023

Received in Revised form 23 October 2023

Accepted 8 November 2023

Published online 8 November 2023

DOI: [10.22044/jme.2023.13610.2527](https://doi.org/10.22044/jme.2023.13610.2527)

### Keywords

Ore grade estimation (Fe<sub>2</sub>O<sub>3</sub>)

Artificial intelligence

Random forests

Mean shift

Gradient boosting

### Abstract

Drilling of exploratory boreholes is one of the most important and costly steps in mineral exploration, which can provide us with accurate and appropriate information to continue the mining process. There are limitations on drilling the target boreholes, such as high costs, topographical problems in installation of drilling rigs, restrictions caused by previous mining operation etc. The advances in artificial intelligence can help to solve these problems. In this research, we used python as one of the most pervasive and the most powerful programming languages in the field of data analysis and artificial intelligence. In this method mean shift algorithms have been used to cluster data, random forest to estimate clusters, and gradient boosting to estimate iron grade. Finally, in the studied area of Choghart in Central Iran, more than 91% accuracy was achieved in detection of ore blocks. Also, the results of the neural network indicate the mean square error (MSE) and mean absolute error (MAE) in the training data, respectively equal to 0.001 and 0.029, in the test data is 0.002 and 0.03, and in the validation boreholes, we reached a maximum of 0.06 and 0.2.

### 1. Introduction

Exploration boreholes are one of the most important and costly steps in mineral exploration, which can provide us with accurate and appropriate information to continue the mining process. There are limitations on drilling the target boreholes, such as high costs, topographical problems in installation of drilling rigs, restrictions caused by previous mining operation etc. The use of artificial intelligence and its superiority over other methods in solving non-linear problems can be one of the best solutions to reduce exploratory drilling. Due to their structure, neural networks can provide mining engineers with simpler and less expensive ways to achieve more accurate results [1].

AI-based methods can overcome the problems caused by the traditional methods with a more realistic strategy; because these methods can understand the hidden relationships between

different input and output parameters in nonlinear and complex spatial conditions.

The methods based on artificial intelligence are:

1. Artificial neural network methods (operator)
2. Optimization methods (optimizer)
3. Fuzzy logic method (normalizer)
4. Clustering methods (classifier) [2].

Data clustering is an unsupervised method in artificial neural networks. Clustering is used to show the difference, and indeed, we do not anticipate simply dividing the data into different clusters. Understanding clustering models is the key to realize the differences and similarities between different clustering algorithms. Understanding clustering models is the key to knowing the differences and similarities between different clustering algorithms. In the clustering

Corresponding author: [hassanreza1995@gmail.com](mailto:hassanreza1995@gmail.com) (H. Ghasemitabar)

method, the members of each cluster are tried to have the most similarity to each other based on the used variables, and also the members of different clusters have the most differences. The most important of these clusters are partition-based clustering, hierarchical clustering, density-based clustering, model-based clustering, and fuzzy clustering [3].

A good clustering should have features such as scalability, acceptability of data types, extraction of clusters with different shapes, ability to deal with noise and incomplete data, insensitivity to data entry, no need to specify input parameters, and accept big data [4].

Python was designed in the late 1980s by Guido van Rossum, with a dynamic system with an emphasis on readability and rapid prototyping. Python is currently the premier programming language for scientific computing, data science, and machine learning, and it enhances performance and productivity by using low-level libraries and appropriate APIs.

The most important strengths of this programming language are as follows:

1. Ease of use while ensuring computational efficiency
2. Create efficient libraries with lower-level code than other programming languages
3. Parallel processing of operations
4. A free and available programming language
5. Portability between different operating systems [5].

Many researchers have employed different artificial neural networks for grade estimation in the recent years. For example, a Radial Basis Function (RBF) network has been used successfully for grade estimation in an iron deposit in England, and the results obtained have been compared with the geostatistical models [6]. In another case, a four-level perceptron network (4L-MLP) has been used on the modified magnetic data

to estimate the iron grade [7]. Badel has compared one of the newer kriging methods called the median indicator kriging with the artificial neural networks for grade estimation in an iron ore deposit [8]. The grade estimation results of a Choghart iron ore deposit obtained from a back-propagation neural network have been compared with the results of a Support Vector Machine (SVM) [9]. In another research work, the artificial neural networks and geostatistics have been integrated using ANNMG to optimize the mineral reserve evaluation in the SW Sierra Leone [10]. Nezamolhosseini has applied a multi-layer perceptron (MLP) neural network to estimate the storage of Choghart mine using the exploratory boreholes data [11]. In another article, the comparison between the Local Linear Radial Basis Function using Skewed Gaussian activation (LLRBF-SG) and older neural networks such as Differential Evolution (DE), Cuckoo Search (CS), Covariance Matrix Adaptation Evolution (CMAE), Artificial Bee Colony (ABC), Improved Artificial Bee Colony (IABC) were used to estimate the phosphate grade in Bafgh Esfordi [12]. Alimoradi has estimated the silver grade in Zarshuran gold mine by involving the boreholes spatial data and the data obtained from the Induced Polarization (IP) geophysical method with the cuckoo search machine learning algorithm. The results show that grade values can be accurately estimated from geophysical data, especially in areas without drilling operations data [13]. All of these research works have illustrated that artificial neural networks can be used as a reliable approach to obtain the most accurate grade estimations.

The mean shift algorithm is a clustering method for the analysis of complex spaces and non-parametric properties to determine the maximum probability function. Application domains of this algorithm include cluster analysis in machine learning and image processing. This algorithm is based on data density, and can automatically adjust the number of clusters [14].

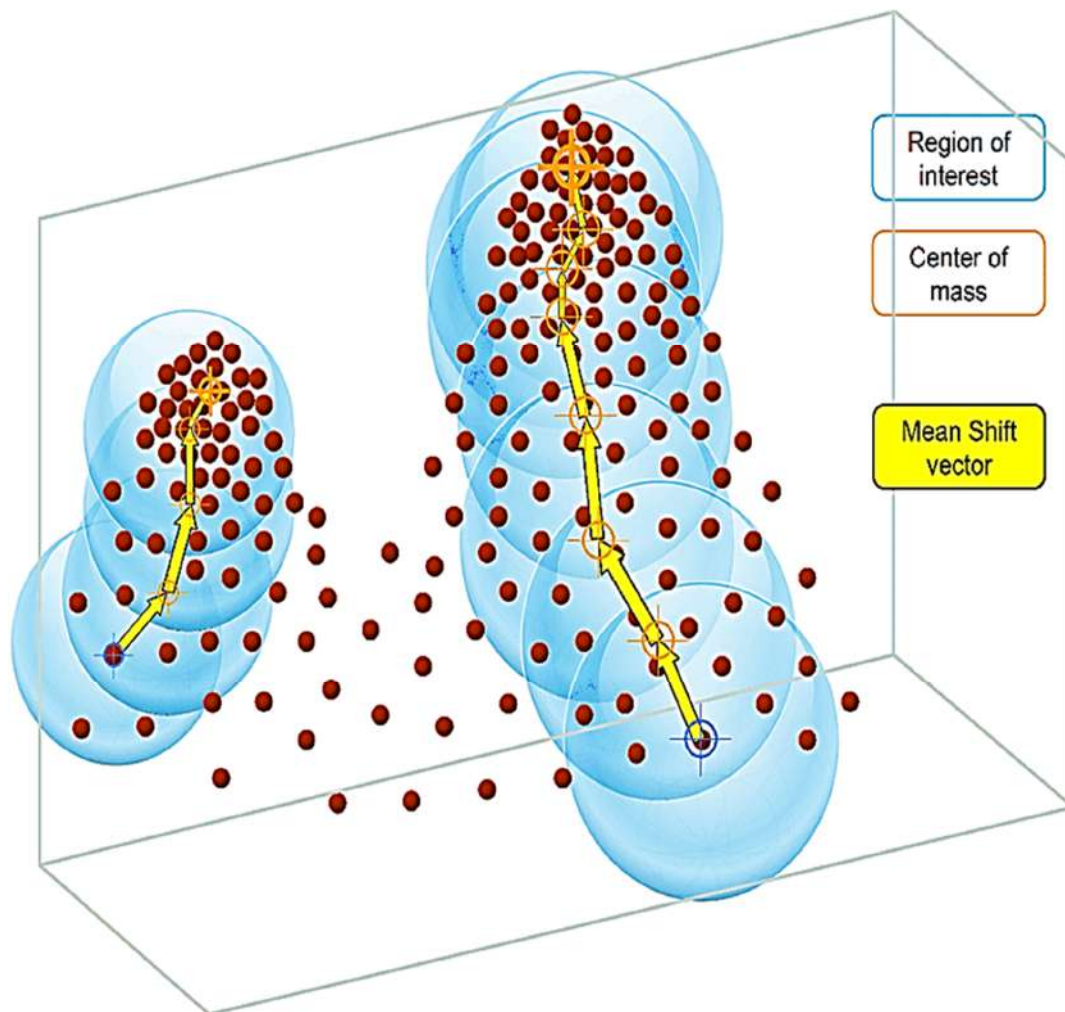


Figure 1. How to select centers in each step in the mean shift algorithm [15].

The problem of this algorithm is in determining the bandwidth, which is a difficult operation. To determine the bandwidth, the results of clustering estimation have been used and the best bandwidth with appropriate estimation has been selected by repeating different numbers [16]. This algorithm is available in the sci-kit-learn library in python programming language.

Random forest or random decision forests is a combined learning method for classification, regression, which is based on a structure consisting of a large number of decision trees, on the training time and output of classes (classification), or for the average predictions of each tree. They work separately [17]. Random forests are suitable for decision trees that are pre-fitted in the training complex. Also, this algorithm is very user-friendly,

and has only two network input parameters, which are the number of trees and subset variables, which usually the network response is not very sensitive to the value of these parameters [18].

The decision tree is an algorithm that is easy to understand and interpret, but a single tree may not be sufficient to learn the properties of the model. Random forest algorithm, on the other hand, is a tree-based algorithm that uses the properties of several trees to make decisions. Also, the decision tree algorithm is extremely vulnerable in terms of over-training and over-fitting, but this problem can be easily overcome by performing random forest regression. Another important feature of this algorithm is its low variance in regression. This algorithm uses averaging to improve performance and control the overfitting [19].

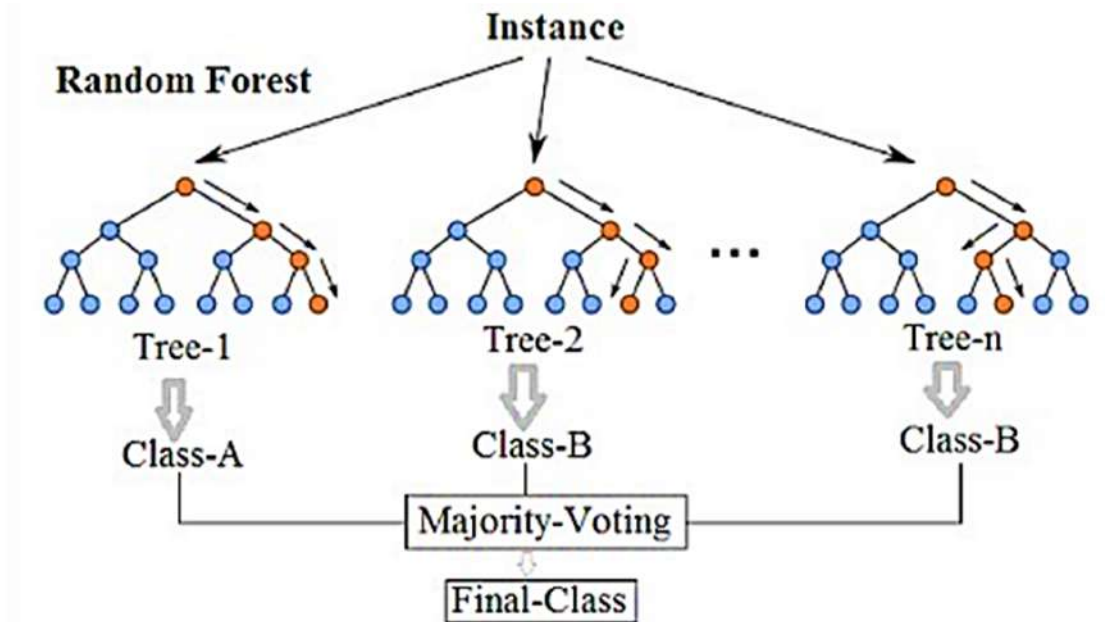


Figure 2. Random forest performance procedure and integration of results of several decision trees [19].

This algorithm is also available in the sci-kit-learn library and the ensemble sub-library in the python programming language.

The gradient boosting algorithm is a machine learning method for regression and classification problems that are typically generated from a prediction model in the form of a set of weak prediction models, typically a decision tree. It builds this model step by step, like other amplifier methods, and generalizes the variable performance

of the decision tree by allowing arbitrary optimization. The gradient boosting algorithm is an integrated, high-performance, stable algorithm; it can control noise data well, and has a high estimation ability to predict non-linear data [20].

Figure 3 shows the training process and progress of the gradient boosting algorithm based on the error function and iteration of the training process:

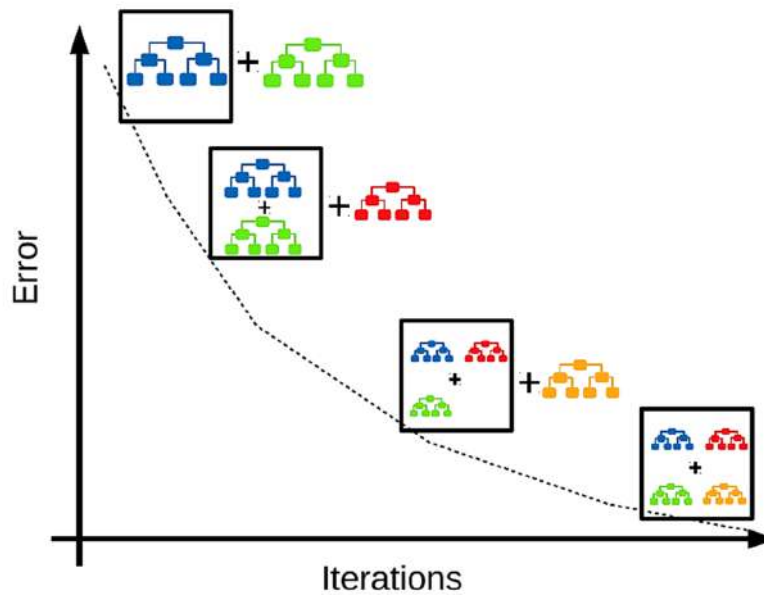


Figure 3. Gradient boosting training process [21].

Gradient boosting is a decision tree-based algorithm developed with the Ada Boost method. The correct understanding of the error function depends on the parameter should be optimized. One of the most important features of this method is that it allows the user to specify the error function according to its needs [22].

The main difference between this method and random forest is that in this method the decision trees form a network in the direction of each other, while in random forest the trees form a network together. It is completely observable in Figure 4. [22].

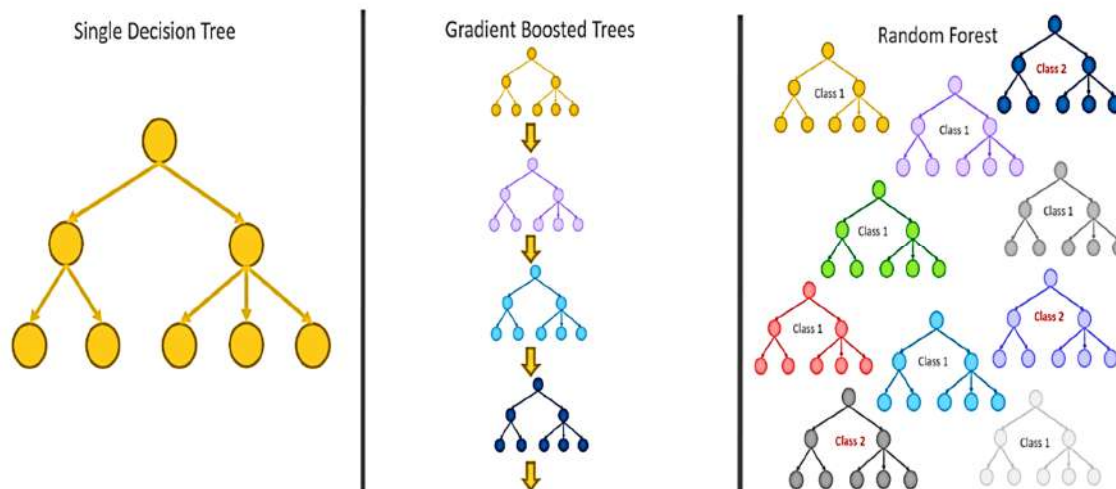


Figure 4. Difference between random forest algorithm and gradient boosting algorithm [22].

In this study, the gradient boosting algorithm is used to estimate the iron grade in the last step. This algorithm can also be used in the sci-kit-learn library in python programming language.

In the beginning of the study, data mine software is used for composite data & then they are reviewed & pre- processed with SQL2014 software; in the main stage, the required code is implemented in python programming language (Python 3.7) and in Spyder environment of the Anaconda package.

**۲. Choghart iron Deposit**

Central Iran zone is located within the Alpine-Himalayan orogenic system, which evolved during the closing of Palettes Ocean. This zone is located in the northeast of Zagros-Makran belt connected to the NeoThetis ocean suture along with the other areas of Cimmerian block of Iran (Alborz and

Sanandaj-Sirjan). The microcontroller separately described the continents of central and eastern Iran with fault boundaries including three crustal areas of Lut, Tabas, and Yazd blocks with a north-south orientation that are adjacent from east to west, respectively. Tabas and Yazd blocks are separated by a long, complex arched structural belt called Kashmar-Kerman tectonic zone (It is also called Posht-Badam block). Located 12 km NE of the city of Bafgh in Iran, Choghart mine is one of the biggest iron ores in this country. The main orebody at Choghart is in the form of a roughly vertical, discordant, pipe-shaped body plunging 73°NNW, and has been explored to a depth of 600 m. Different types of volcanic (intrusive and extrusive alkali rhyolites) and metamorphous rocks occur in the vicinity of the deposit. Syenite, pyroxenite, gabbro, granite, and alkali rhyolites are the major components of the volcanic rocks of Choghart deposit [23].



Figure 5. Satellite photo illustrating location of the drilling points [23].

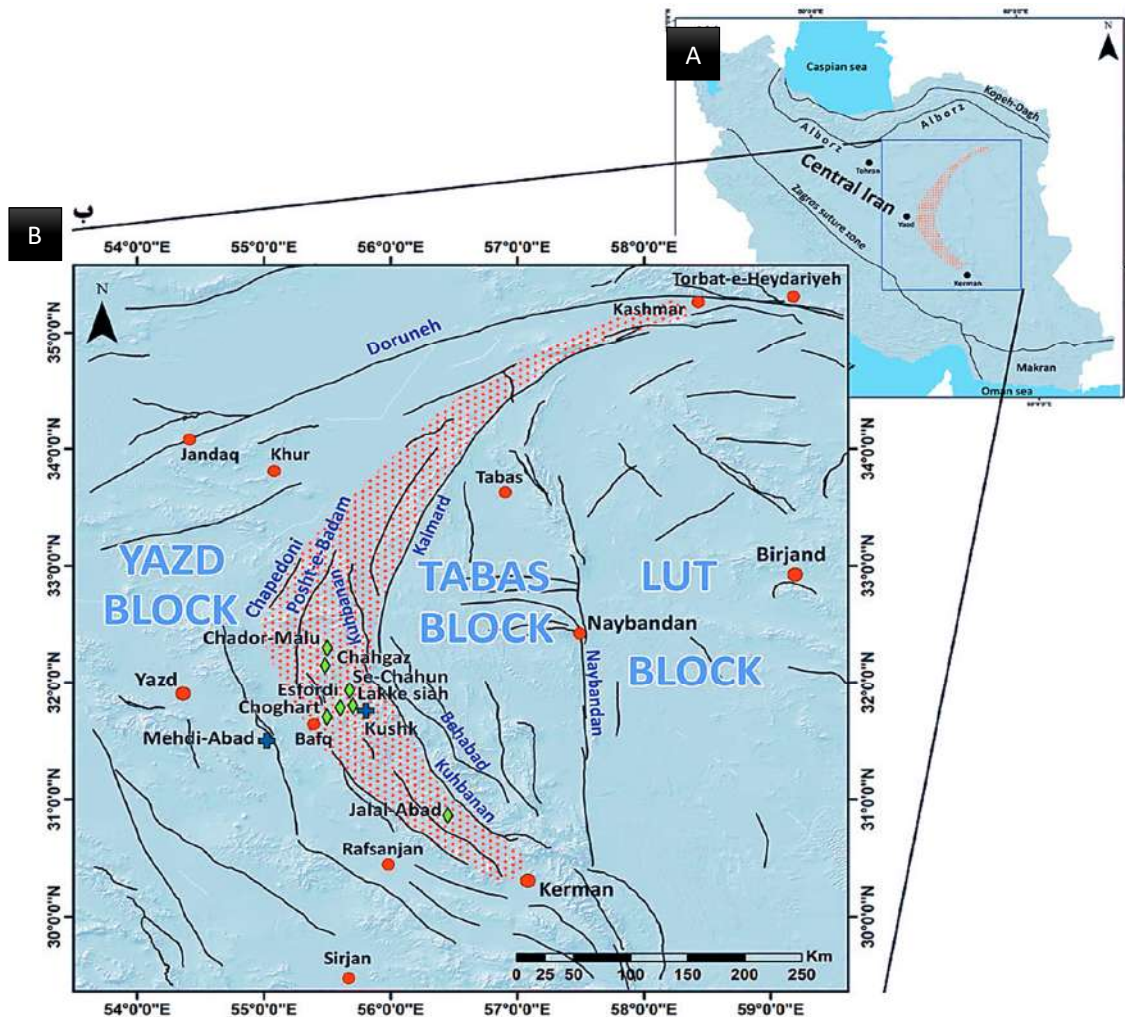


Figure 6. a) The position of Central Iran in relation to the Zagros and Alborz joints; B) Landscape map of Central Iran blocks; C) Geological map of Baigh-Saghand block with the location of iron oxide-apatite, manganese and lead, and zinc deposits [24].

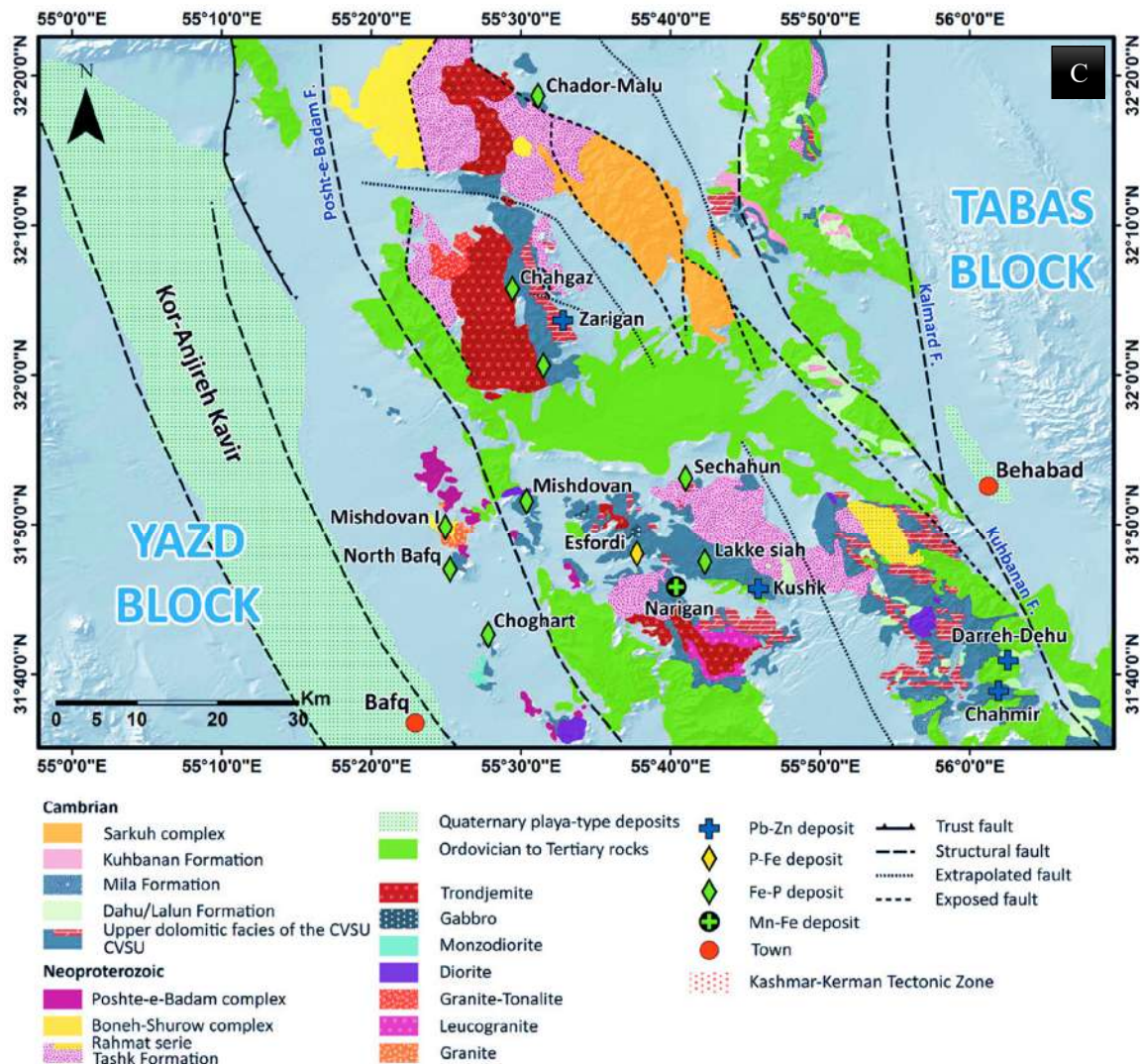


Figure 6. a) The position of Central Iran in relation to the Zagros and Alborz joints; B) Landscape map of Central Iran blocks; C) Geological map of Bafgh-Saghand block with the location of iron oxide-apatite, manganese and lead, and zinc deposits [24].

The oldest rocks at Bafq mining district are composed of Mesoproterozoic metamorphic schists covered by the Neoproterozoic to Cretaceous units. Bafq region is one of the most important mineralized zones of central Iran with the upper Precambrian metamorphic-sedimentary rocks and rift series of Precambrian-Paleozoic [25].

The Cambrian Volcano Sedimentary Unit (CVSU), as the major host of Bafq Fe-P-REE deposits, is composed of Rizu-Desu series and Esfordi Formation. The CVSU is made up of felsic tuff, sandstone and micro-conglomerate, mafic, and felsic volcanic rocks, pyritic siltstone-shale, volcanoclastic beds and tuffaceous shale, dolomite and dolomitic limestone [26].

Potassic, phyllic, argillic, and propylitic are the major alteration types, and they are attended by

the veins to veinlets of quartz, quartz magnetite, and Fe-hydroxides filling [27].

### 3. Methodology

In this section, the background of surveys and observations about the borehole datasets are presented. Then, considering that only coordinates and degrees are available, the proposed method will be as follow: According to have degrees and coordinates in different dimensions, data are clustered and models are created to estimate these categories. Finally, the model proceeds to estimate the grade according to different clusters.

### 3.1. Data preparation

The information used in this paper is obtained in the form of assay, collar, and survey data from the raw data. After entering this information into the data environment, the boreholes were composited to unify the sampled lengths. Composition is a length-dependent operation. Since the mining operations have been carried out in the studied area before our research and the extraction blocks have been considered as 10 meters, the composite has been considered based on the length of the extraction blocks and the limit grade. A total of 2264 data was obtained, of which 2 boreholes in the center and 2 boreholes in the exploration area were excluded from the data set as validation boreholes. The statistical adjustments related to the other data used in the neural network training and testing is as follows, and the data of these four boreholes has been completely excluded from the network process.

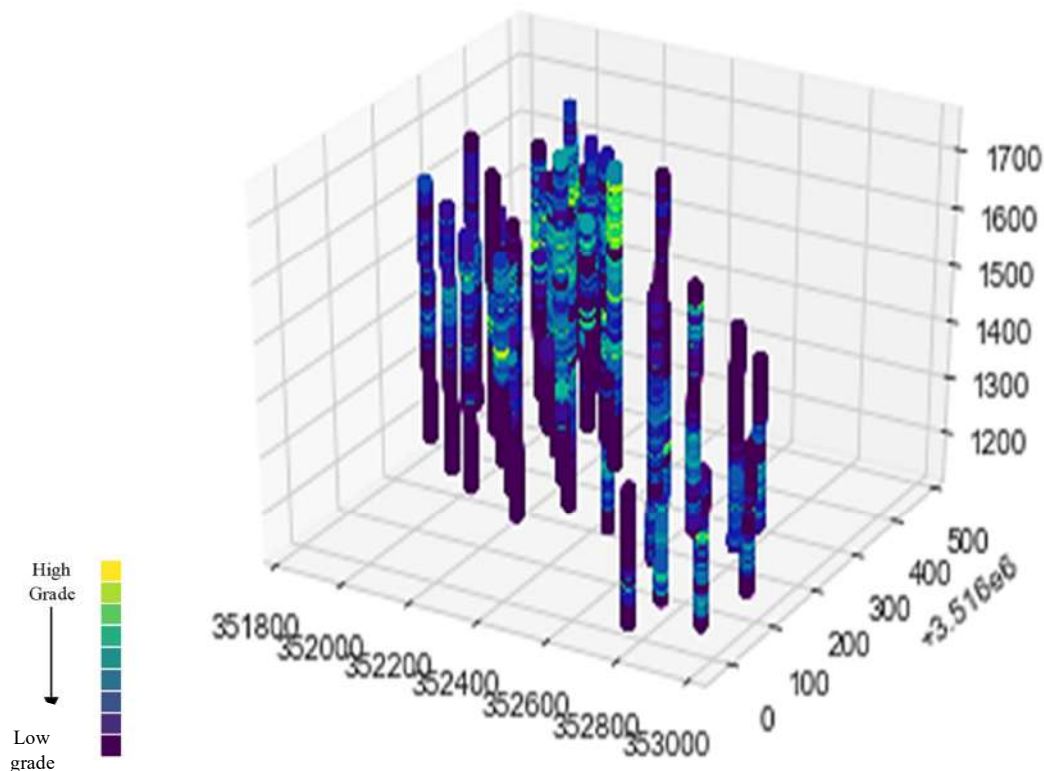
### 3.2. Data statistical studies

The learning data sets from the Pandas library in python environment is extracted as follows:

**Table 1. Statistical parameters of the studied data.**

| Parameter      | Grade value |
|----------------|-------------|
| Mean           | 12.9        |
| Std. deviation | 11.9        |
| Minimum        | 0           |
| 25%            | 0           |
| 50%            | 12.3        |
| 75%            | 20.86       |
| Maximum        | 62.68       |
| Variance       | 142         |

Notable points include the large amount of zero data in the measured values, which can be related to the low accuracy of sampling and borehole analysis as well as human error; this can inevitably lead to some errors.



**Figure 7. Diagram of how the boreholes are positioned according to the measured grade.**

In the picture above, the position of the boreholes can be seen according to the measured grade. Although the distance between the boreholes and the position of the boreholes is not

regular, their distance from each other is approximately 100 meters. Also, the frequency of zero-k data with a dark color has been illustrated, indicating low accuracy in measured grades.



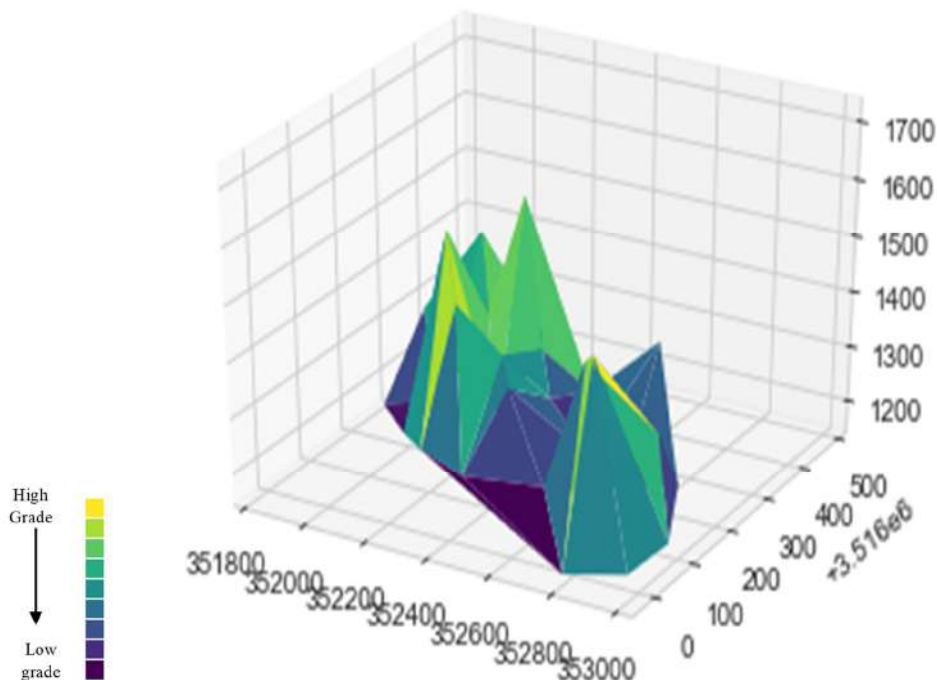


Figure 8. Procedural diagram of available information from boreholes.

According to Figure 8, it can be easily understood that the length of the boreholes is not regular, which is due to the limitations created by the topography of the area. This issue can make difficulty in network training process, especially in

places where there is not enough data for network training in terms of height.

To further investigate this scattering in different directions, diagrams of the degree of change in different directions are given as below:

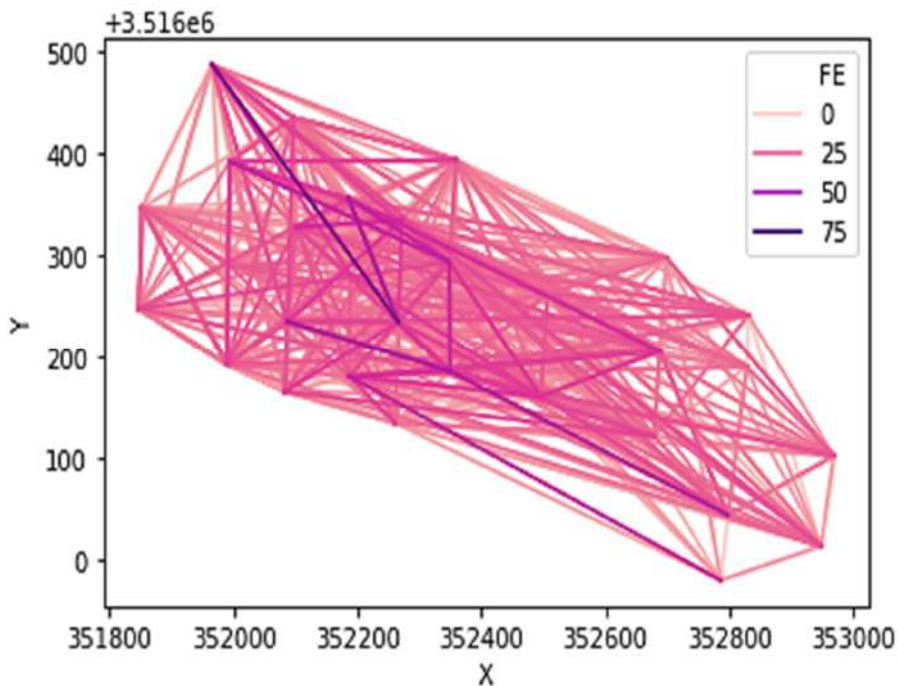


Figure 9. Ore grade changes in the X-Y direction.

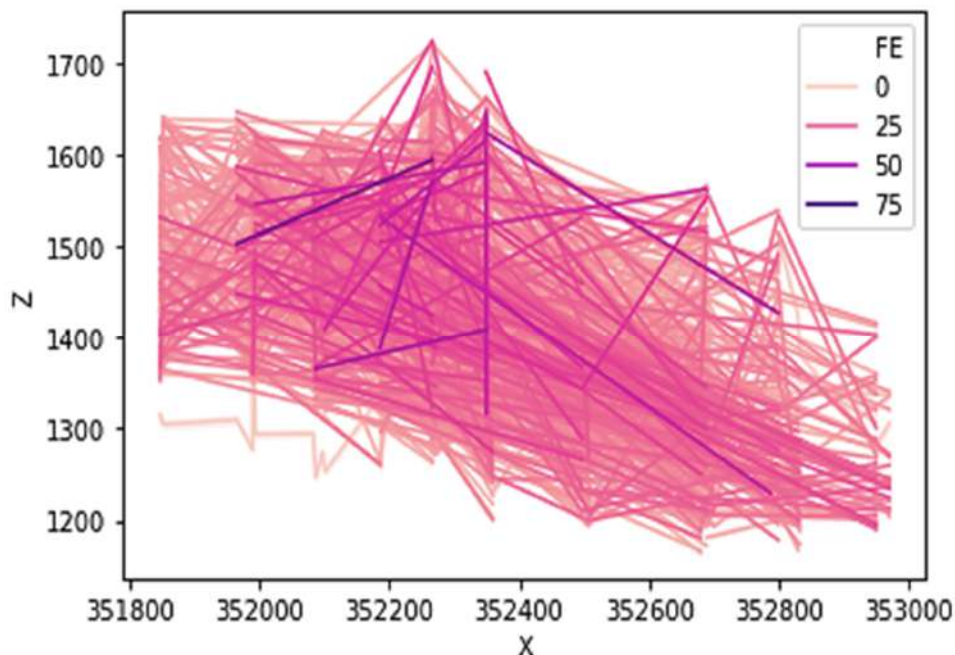


Figure 10. Ore grade changes in the X-Z direction.

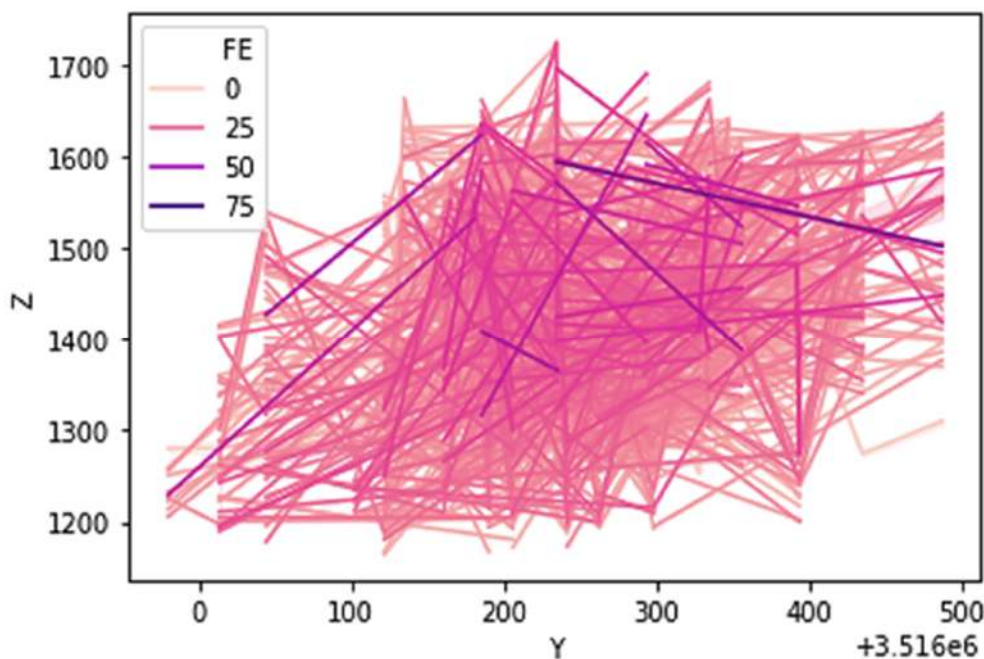


Figure 11. Carat changes in the Y-Z direction.

Looking more closely at the ore grade change diagrams, it can be seen that the changes are more severe along the Z-axis, and the depth parameter will play more important role in predicting the final ore grade. Due to the data scattering, an algorithm which is resistant to the out-of-noise data is needed. This algorithm can cluster data with small amounts. Also, the nonlinear behavior of the data

and their unpredictability are quite evident in these graphs.

### 3.3. Data pre-processing

To prepare the data for use in neural networks, it is necessary to pre-process the data first. At this stage, after examining the data in terms of validity and usability and deleting invalid data, the data is

normalized. This is done to equalize the effect of large and small data, and all data will be within a certain range.

In this paper, the Min-Max model is used to normalize the data and the data is defined in the range of [0,1]. The python programming language uses a preprocessing library to do this. In this method, each set of data is mapped to arbitrary intervals whose minimum and maximum values

$$v' = (v - \min_A) \frac{\text{newMax} - \text{newMin}}{\text{maxA} - \text{minA}} + (\text{newMin})$$

### 3.4. Main shift clustering algorithm

Mean shift is a procedure for locating the maxima of the modes of a density function. This is an iterative method, and we start with an initial estimate. The Kernel function determines the weight of nearby points for re-estimation of the mean. Typically, a Gaussian Kernel on the distance to the current estimate is used:

$$K(x_i - x) = e^{-c\|x_i - x\|^2}.$$

The weighted mean of the density in the window determined by  $K$  is:

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x)x_i}{\sum_{x_i \in N(x)} K(x_i - x)}$$

where  $N(x)$  is the neighborhood of  $x$ , is a set of points for which  $K(x_i) \neq 0$  [28].

The difference of  $m(x) - x$  is called mean shift in Fukunaga and Hostetler. The mean shift algorithm now sets  $m(x) \rightarrow x$ , and repeats the estimation until  $m(x)$  converges [29].

Although the mean shift algorithm has been widely used in many applications, a rigid proof for the convergence of the algorithm using a general Kernel in a high dimensional space is still not known [30]. However, the one-dimensional case has limited real world applications. Also, the convergence of the algorithm in higher dimensions with a finite number of the (or isolated) stationary points has been proved [30,31]. However, sufficient conditions for a general Kernel function to have finite (or isolated) stationary points have not been provided.

### 3.5. Random forests algorithm

#### 3.5.1. Tree learning

Decision trees are a popular method for various machine learning tasks. Tree learning comes closest to meeting the requirements for serving as an off-the-shelf procedure for data mining because

are already known. In this way, any desired interval can be converted to a new interval with just a simple conversion. Suppose that attribute A is to be mapped from the data set between  $\min_A$  to  $\max_A$  to the new range  $\text{newMin}$  to  $\text{newMax}$ . For this purpose, any initial value such as  $v$  in the initial interval will be converted to a new value  $v'$  in the new interval according to the following equation:

it is invariant under scaling and various other transformations of feature values, is robust to inclusion of irrelevant features, and produces inspectable models. However, they are seldom accurate [32].

In particular, trees that are grown very deep tend to learn highly irregular patterns: they over-fit their training sets, i.e. have low bias, but very high variance. Random forests are a way of averaging multiple deep decision trees, trained on different parts of the same training set, with the goal of reducing the variance [32].

This comes at the expense of small increase in the bias and some loss of interpretability, but generally greatly boosts the performance in the final model. Forests are like the pulling together of decision tree algorithm efforts. Taking the teamwork of many trees thus improving the performance of a single random tree. Though not quite similar, forests give the effects of a K-fold cross-validation [33].

#### 3.5.2. Bagging

The training algorithm for random forests applies the general technique of bootstrap aggregating, or bagging, to tree learners. Given a training set  $X = x_1, \dots, x_n$  with responses  $Y = y_1, \dots, y_n$ , bagging repeatedly ( $B$  times) selects a random sample with replacement of the training set and fits trees to these samples:

For  $b = 1, \dots, B$ :

1. Sample, with replacement,  $n$  training examples from  $X, Y$ ; call these  $X_b, Y_b$ .
2. Train a classification or regression tree  $f_b$  on  $X_b, Y_b$ .

After training, predictions for unseen samples  $x'$  can be made by averaging the predictions from all the individual regression trees on  $x'$ :

$$\hat{f} = \frac{1}{B} \sum_{b=1}^B f_b(x')$$

or by taking the majority vote in the case of classification trees.

This bootstrapping procedure leads to better model performance because it decreases the variance of the model, without increasing the bias. This means that while the predictions of a single tree are highly sensitive to noise in its training set, the average of many trees is not, as long as the trees are not correlated. Simply training many trees on a single training set would give strongly correlated trees (or even the same tree many times, if the training algorithm is deterministic); bootstrap sampling is a way of de-correlating the trees by showing them different training sets [34].

### 3.5.3. From bagging to random forests

The above procedure describes the original bagging algorithm for trees. Random forests differ in only one way from this general scheme: they use a modified tree learning algorithm that selects, at each candidate split in the learning process, a random subset of the features. This process is sometimes called "feature bagging". The reason for doing this is the correlation of the trees in an ordinary bootstrap sample: if one or a few features are very strong predictors for the response variable (target output), these features will be selected in many of the B trees, causing them to become correlated. An analysis of how bagging and random subspace projection contribute to accuracy gains under different conditions is given by Ho [35].

Typically, for a classification problem with p features,  $\sqrt{p}$  (rounded down) features are used in

each split. For regression problems the inventors recommend  $\frac{p}{3}$  (rounded down) with a minimum node size of 5 as the default. In practice the best values for these parameters will depend on the problem, and they should be treated as tuning parameters [32].

### 3.6. Gradient boosting algorithm

In many supervised learning problems there is an output variable y and a vector of input variables x, related to each other with some probabilistic distribution. The goal is to find some function  $\hat{F}(x)$  that best approximates the output variable from the values of input variables. This is formalized by introducing some loss function  $L(y, F(x))$  and minimizing it:

$$\hat{F} = \underset{F}{arg \min} E_{x,y} [L(y, F(x))].$$

The gradient boosting method assumes a real-valued y, and seeks an approximation  $\hat{F}(x)$  in the form of a weighted sum of functions  $h_i(x)$  from some class  $\mathbb{H}$ , called base (or weak) learners:

$$\hat{F}(x) = \sum_{i=1}^M \gamma_i h_i(x) + const.$$

We are usually given a training set  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  of known sample values of x and corresponding values of y. In accordance with the empirical risk minimization principle, the method tries to find an approximation  $\hat{F}(x)$  that minimizes the average value of the loss function on the training set, i.e., minimizes the empirical risk. It does so by starting with a model, consisting of a constant function  $F_0(x)$ , and incrementally expands it in a greedy fashion:

---


$$F_0(x) = \underset{\gamma}{arg \min} \sum_{i=1}^n L(y_i, \gamma)$$


---

$$F_m(x) = F_{m-1}(x) + \underset{h_m \in \mathbb{H}}{arg \min} \left[ \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + h_m(x_i)) \right]$$


---

where  $h_m \in \mathbb{H}$  is a base learner function.

Unfortunately, choosing the best function h at each step for an arbitrary loss function L is a computationally infeasible optimization problem in general. Therefore, we restrict our approach to a simplified version of the problem. The idea is to

apply a steepest descent step to this minimization problem (functional gradient descent). If we considered the continuous case, i.e. where  $\mathbb{H}$  is the set of arbitrary differentiable functions on R, we would update the model in accordance with the following equations:

$$F_m(x) = F_{m-1}(x) - \gamma_m \sum_{i=1}^n \nabla_{F_{m-1}} L(y_i, F_{m-1}(x_i))$$

$$\gamma_m = \operatorname{arg\,min}_{\gamma} \sum_{i=1}^n L(y_i, F_{m-1}(x_i) - \gamma \nabla_{F_{m-1}} L(y_i, F_{m-1}(x_i)))$$

where the derivatives are taken with respect to the functions  $F_i$  for  $i \in \{1, 2, \dots, m\}$ , and  $\gamma_m$  is the step length. In the discrete case however, i.e. when the set  $H$  is finite, we choose the candidate function  $h$  closest to the gradient of  $L$  for which the coefficient  $\gamma$  may then be calculated with the aid of line search on the above equations. Note that this approach is a heuristic and therefore doesn't yield an exact solution to the given problem, but rather an approximation [36].

#### 4. Ore grade estimation process

##### 4.1. Apply clustering with mean shift algorithm

The mean shift algorithm has been used to cluster the information in different dimensions. One of the main features of this algorithm is that there is no need to determine the number of clusters. This algorithm is available in the sci-kit-learn library in python. The parameters considered for clustering, which include different categories of data in different dimensions, are listed in Table 2. The main parameter in the average transmission algorithm is bandwidth. The selection of bandwidth for different classifications is based on the best result in terms of the minimum mean absolute error. The results of each type of clustering are shown in Table 3 and Figures 12 to 18.

**Table 2. Clusters and parameters involved in them.**

| Considered parameters | Clustering type                 |
|-----------------------|---------------------------------|
| X & Grade             | 1 <sup>st</sup> Clustering type |
| Y & Grade             | 2 <sup>nd</sup> Clustering type |
| Z & Grade             | 3 <sup>rd</sup> Clustering type |
| X,Y & Grade           | 4 <sup>th</sup> Clustering type |
| X,Z & Grade           | 5 <sup>th</sup> Clustering type |
| Y,Z & Grade           | 6 <sup>th</sup> Clustering type |
| X,Y,Z & Grade         | 7 <sup>th</sup> Clustering type |

**Table 3. Bandwidth and number of clusters for each type of clustering.**

| Number of clusters | Bandwidth | Clustering type                 |
|--------------------|-----------|---------------------------------|
| 5                  | 0.14      | 1 <sup>st</sup> clustering type |
| 10                 | 0.12      | 2 <sup>nd</sup> clustering type |
| 11                 | 0.11      | 3 <sup>rd</sup> clustering type |
| 9                  | 0.17      | 4 <sup>th</sup> clustering type |
| 7                  | 0.19      | 5 <sup>th</sup> clustering type |
| 9                  | 0.19      | 6 <sup>th</sup> clustering type |
| 11                 | 0.2       | 7 <sup>th</sup> clustering type |

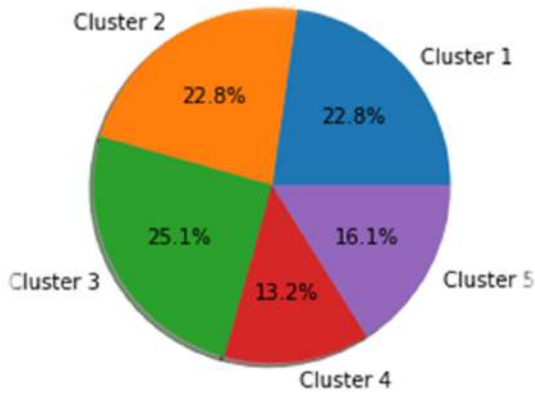


Figure 12. Percentage of data in each cluster of the first type clustering.

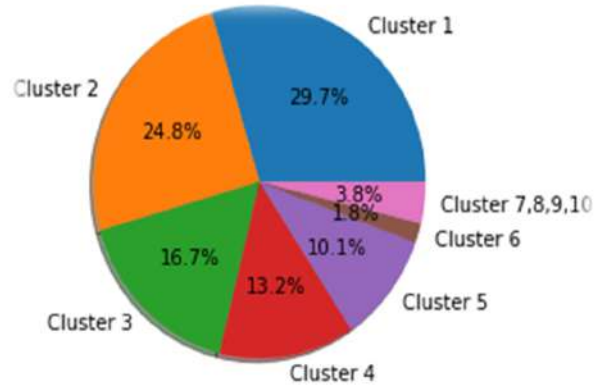


Figure 13. Percentage of data in each cluster of the second type clustering.

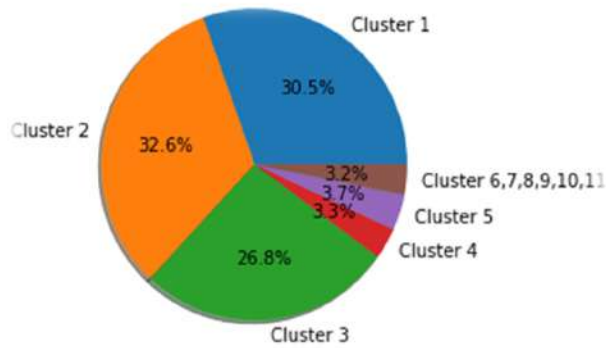


Figure 14. Percentage of data in each cluster of the third type clustering.

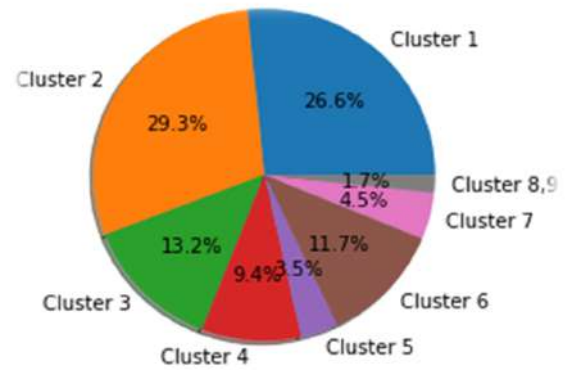


Figure 15. Percentage of data in each cluster of the fourth type clustering.

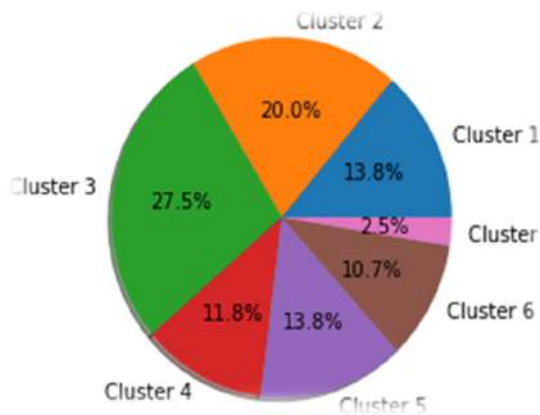


Figure 16. Percentage of data in each cluster of the fifth type clustering.

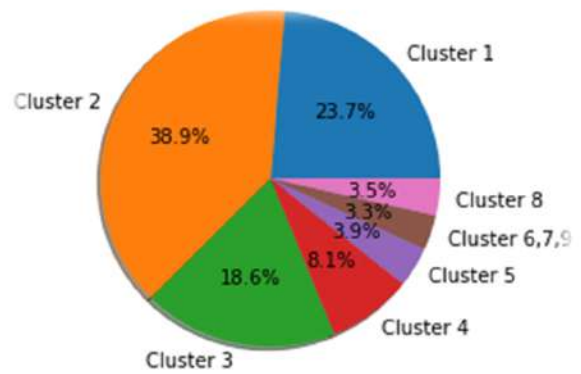


Figure 17. Percentage of data in each cluster of the sixth type clustering.

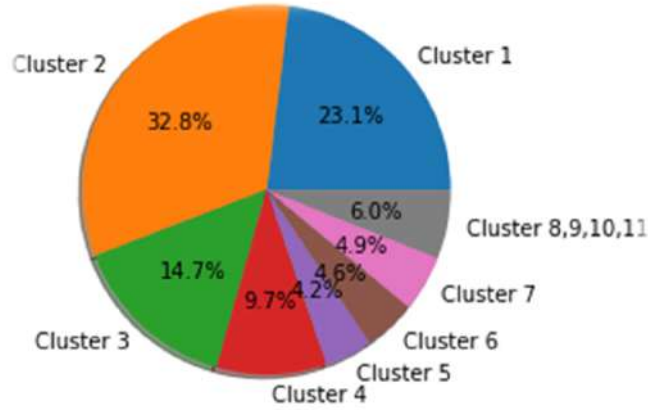


Figure 18. Percentage of data in each cluster of the seventh type clustering.

The following figure (Figure 19) shows the correlation coefficient of the available parameters

for estimating the grade and type of different clusters:

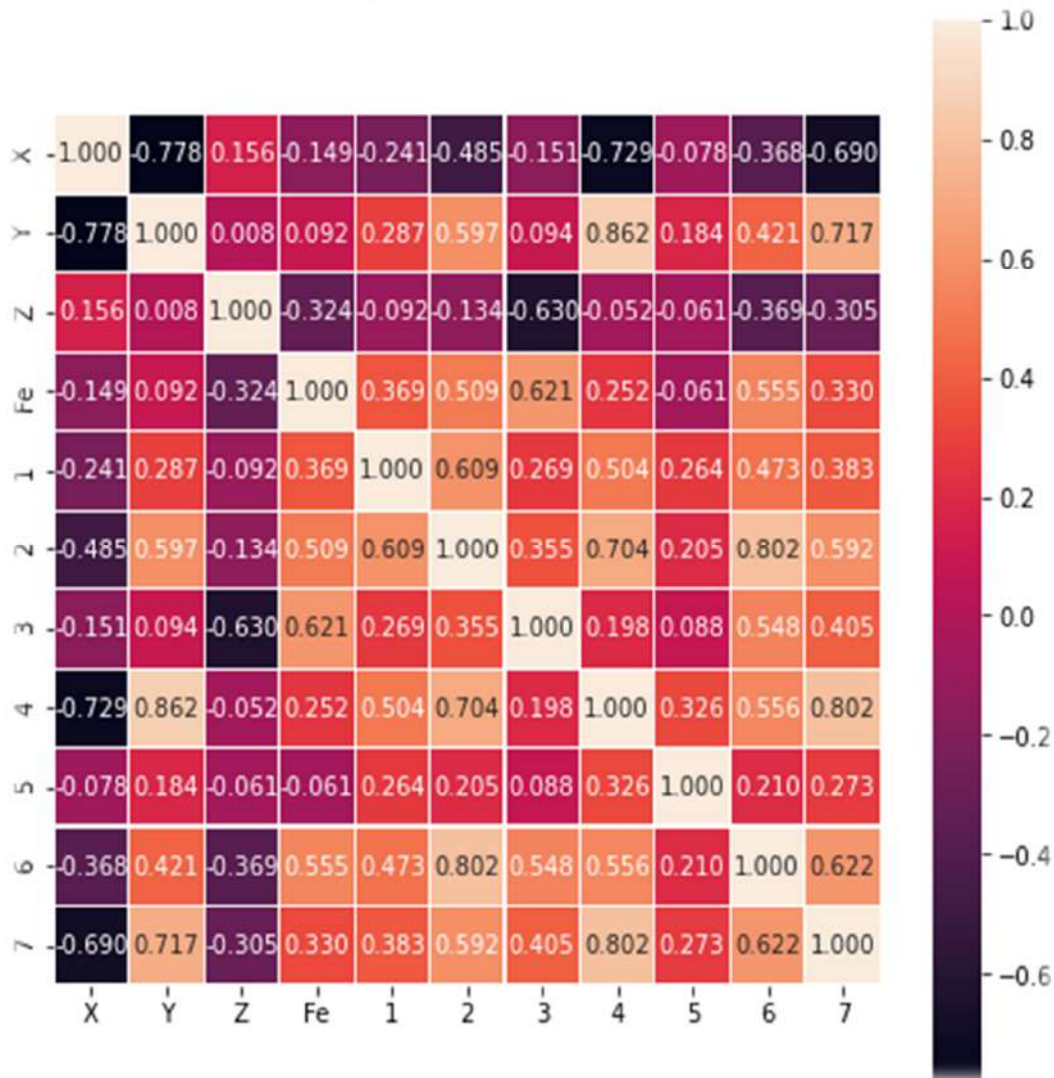


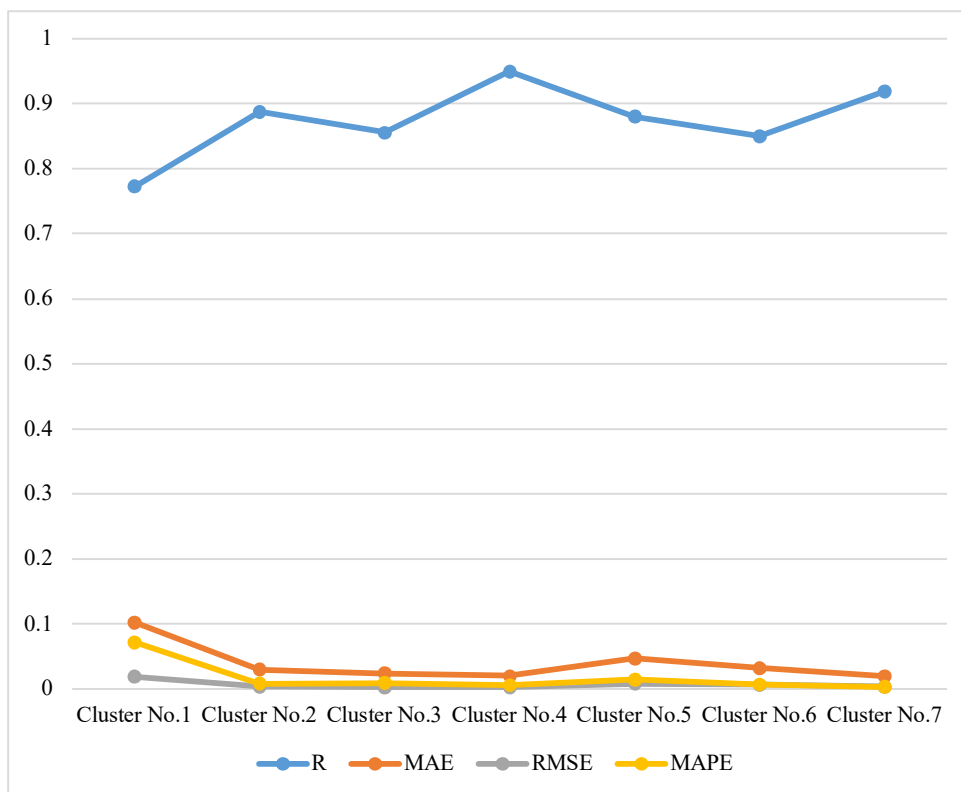
Figure 19. Correlation coefficient of existing parameters for estimating the grade and type of different clusters.

According to the figure above, it can be easily seen that the correlation coefficient of the categories in which the depth was affected has a higher correlation with the iron grade. As a result, it can be concluded that among the data used, depth is the most important parameter in determining the grade.

**4.2. Determining clustering with random forests algorithm**

After clustering the parameters in different directions, it is necessary to predict the numbers

assigned to each category for validation data, which was done by a random forest algorithm. For this purpose, at first, to learn the network for each cluster, the data was divided into training and experimental data in a ratio of 80 to 20. As stated in Section 3.7, the selection of network parameters in the clustering section is based on the selection of the best result in terms of the lowest mean absolute error. You can see the results of comparing the important parameters in the network for each type of clustering in the following figure:



**Figure 20. Comparison of correlation coefficient values and three error parameters (MAE, RMSE, MAPE) in different types of clustering.**

With a little reflection on the accuracy diagrams of the forecast model, it can be seen that, as expected, clusters with fewer data have less accuracy in forecasting; this is due to the lack of training data and consequently poor forecasting in these clusters [38, 39].

**4.3. Ore grade estimation with gradient boosting algorithm**

The ore grade is estimated using the gradient boosting algorithm. The most important parameter

in this algorithm is determining the learning rate. The learning rate of 0.01 has been used to estimate the iron ore grade for this paper according to the use of auxiliary parameters resulting from clustering; for the initial evaluation, the data is divided into train and test data in a ratio of 80 to 20, which are the network input parameters, coordinate specifications and the results of the classification estimation models. In the image below, you can see the results of the grade estimation for the train and test data:



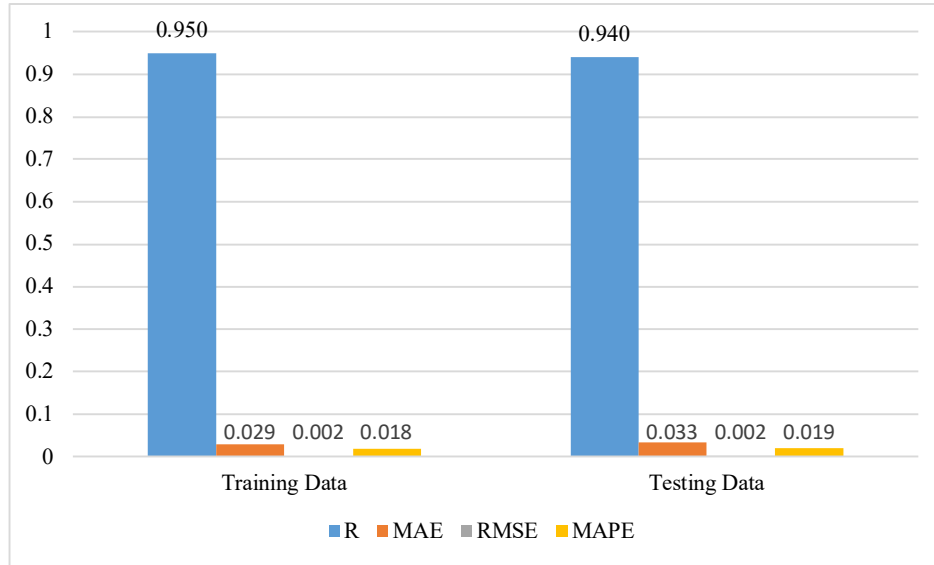


Figure 21. Comparison of correlation coefficient values and three error parameters (MAE, RMSE, MAPE) in the network learning process.

The accuracy of the model in predicting the train and test data can be seen in the above figure, which indicates the appropriateness of the model accuracy and its optimality in estimating the data quality point by point. In addition to the above, it should be noted that all the extraction blocks in terms of waste or ore in the network learning process are correctly predicted and the detection accuracy of the extraction block is equal to 100%.

After comparing the results of training and testing the data and ensuring the low error of the

model, it is time to examine the validation boreholes. As mentioned earlier, four borehole data were completely excluded from the learning process. The stage of reviewing test data can be considered as the stage of point-to-point review of results, and this stage, i.e. validation based on boreholes, can be considered as borehole-to-borehole review. The results of the four validation boreholes are illustrated in Figure 22 and the accuracy of the network prediction in each borehole are shown in Figures 23 to 26.

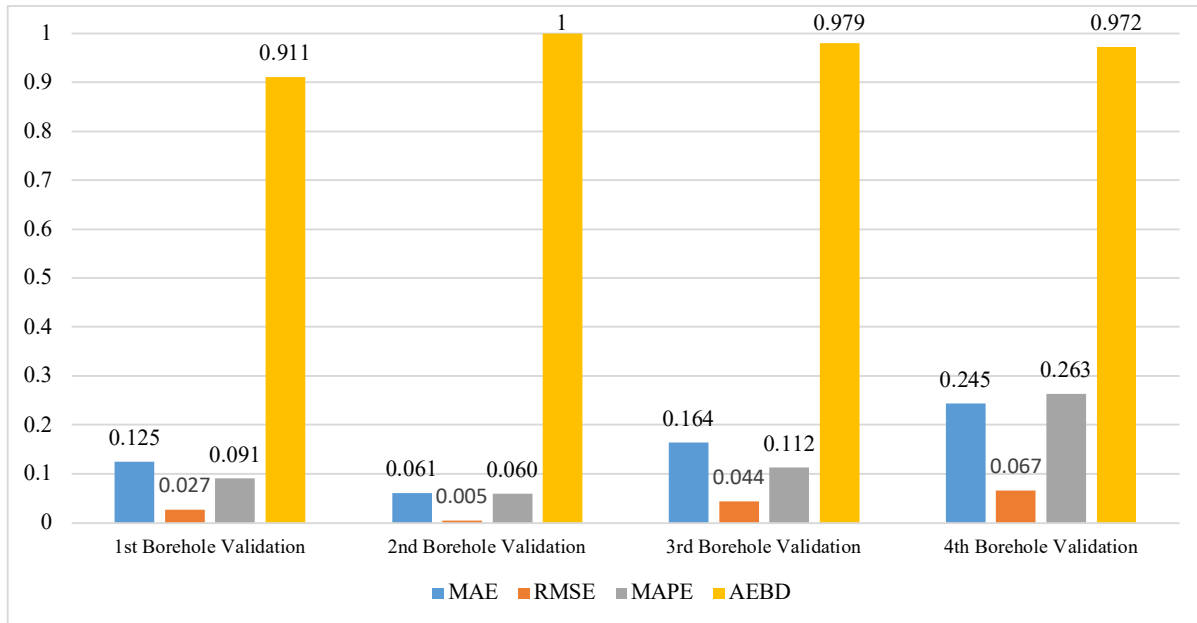


Figure 22. Comparison of accuracy values in extraction block detection (AEBD) and three error parameters (MAE, RMSE, MAPE) in the network validation process.

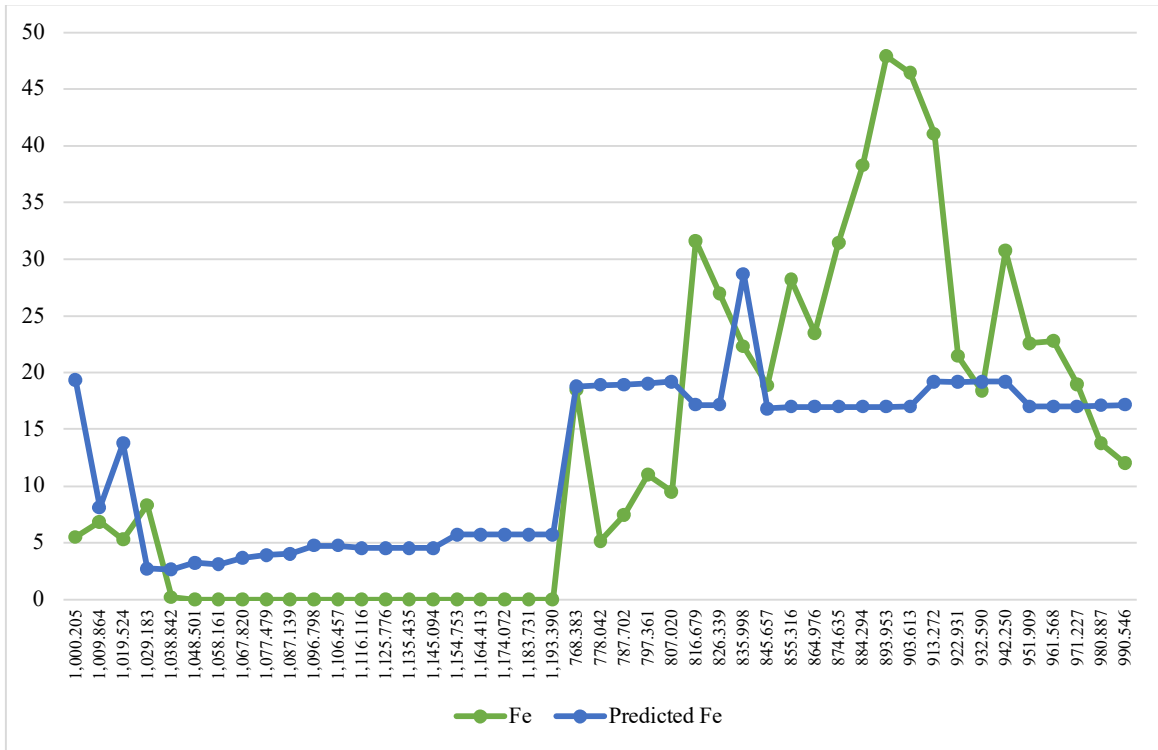


Figure 23. Accuracy of model prediction in the first validation borehole.

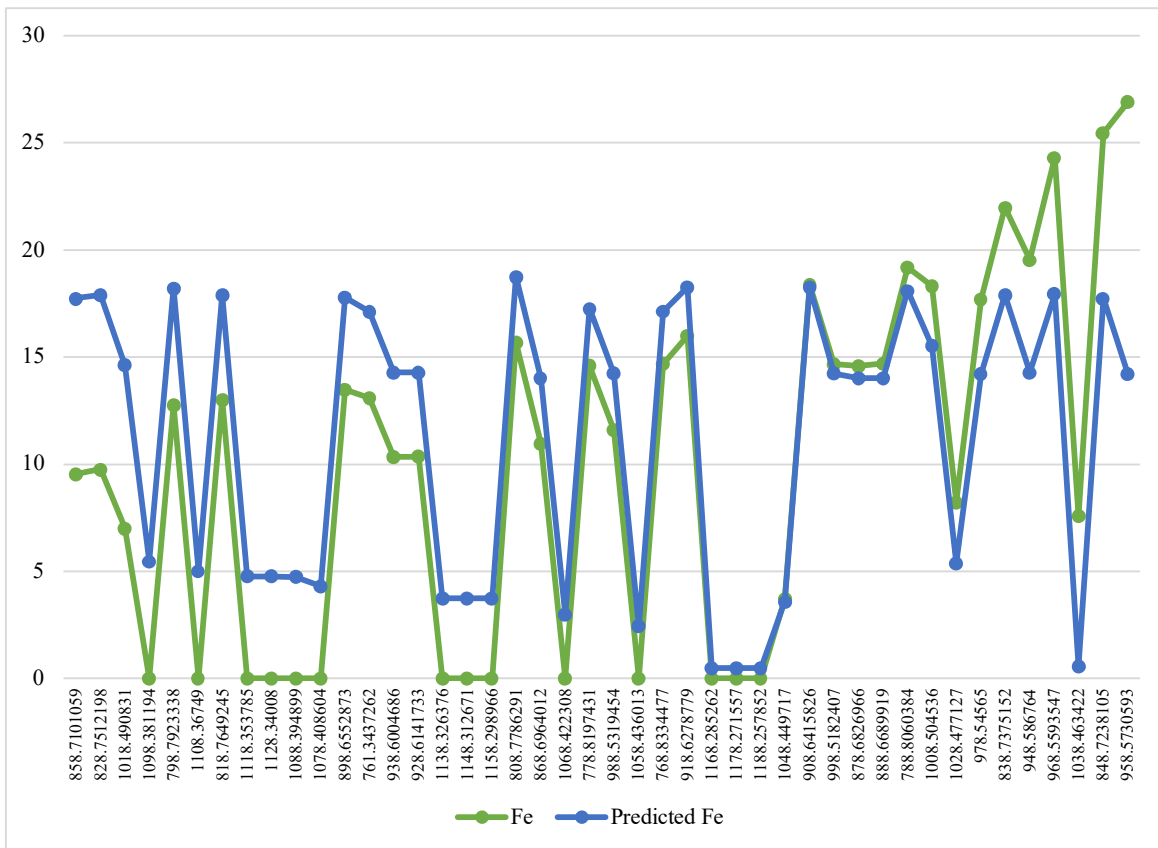


Figure 24. Accuracy of model prediction in the second validation borehole.

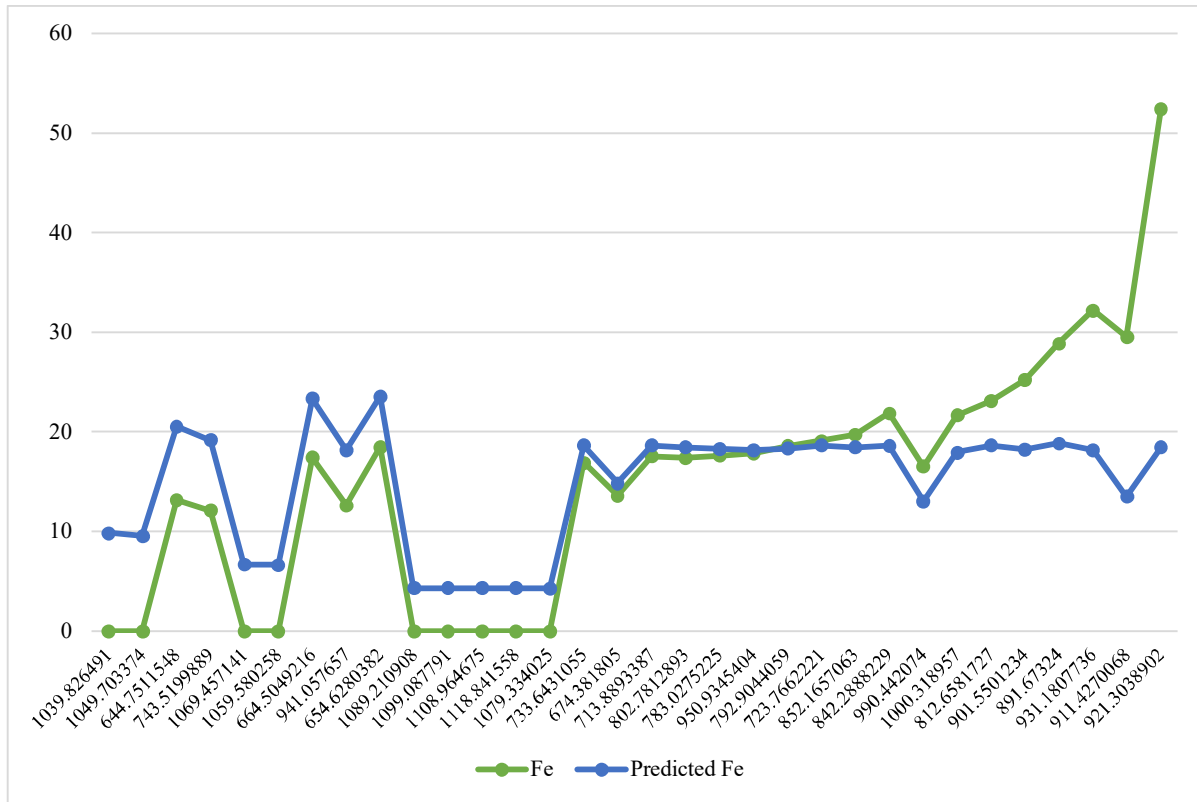


Figure 25. Model prediction accuracy in the third validation borehole.

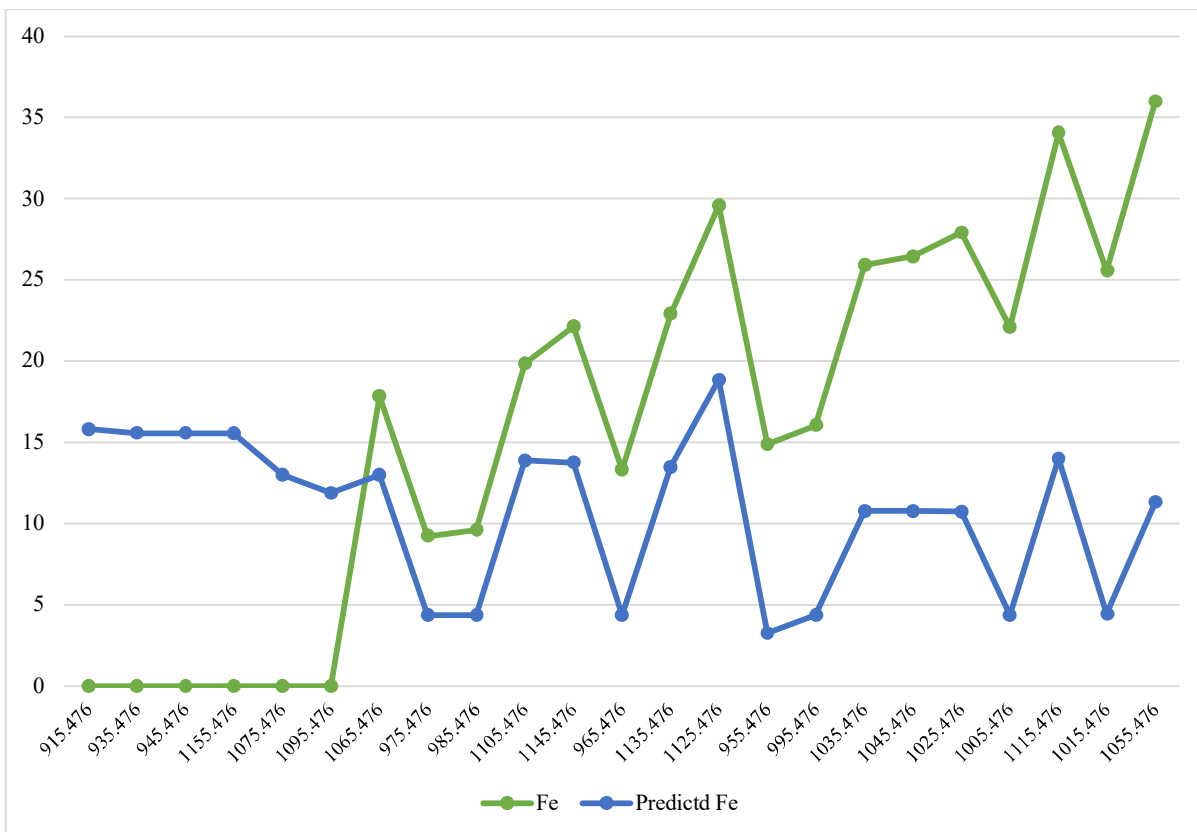


Figure 26. Accuracy of model prediction in the fourth validation borehole.

#### 4.4. Convert mineral ore estimation process to software

In this section, the steps of the grade estimation process in this article were converted into a

software with the help of python programming language, which you can see in the following images related to this software:

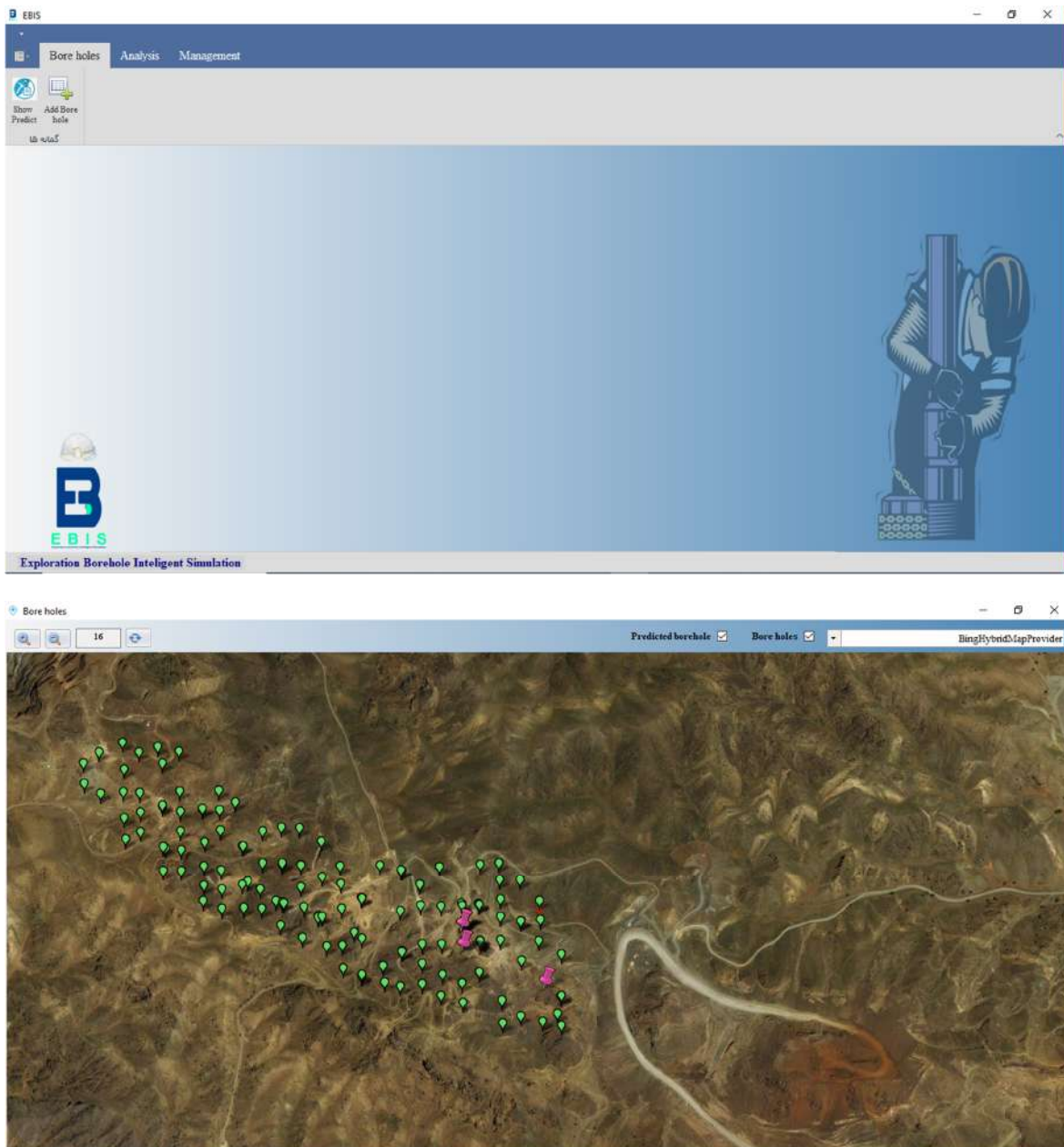


Figure 27. Pictures of the program environment written for the mineral grade estimation process.

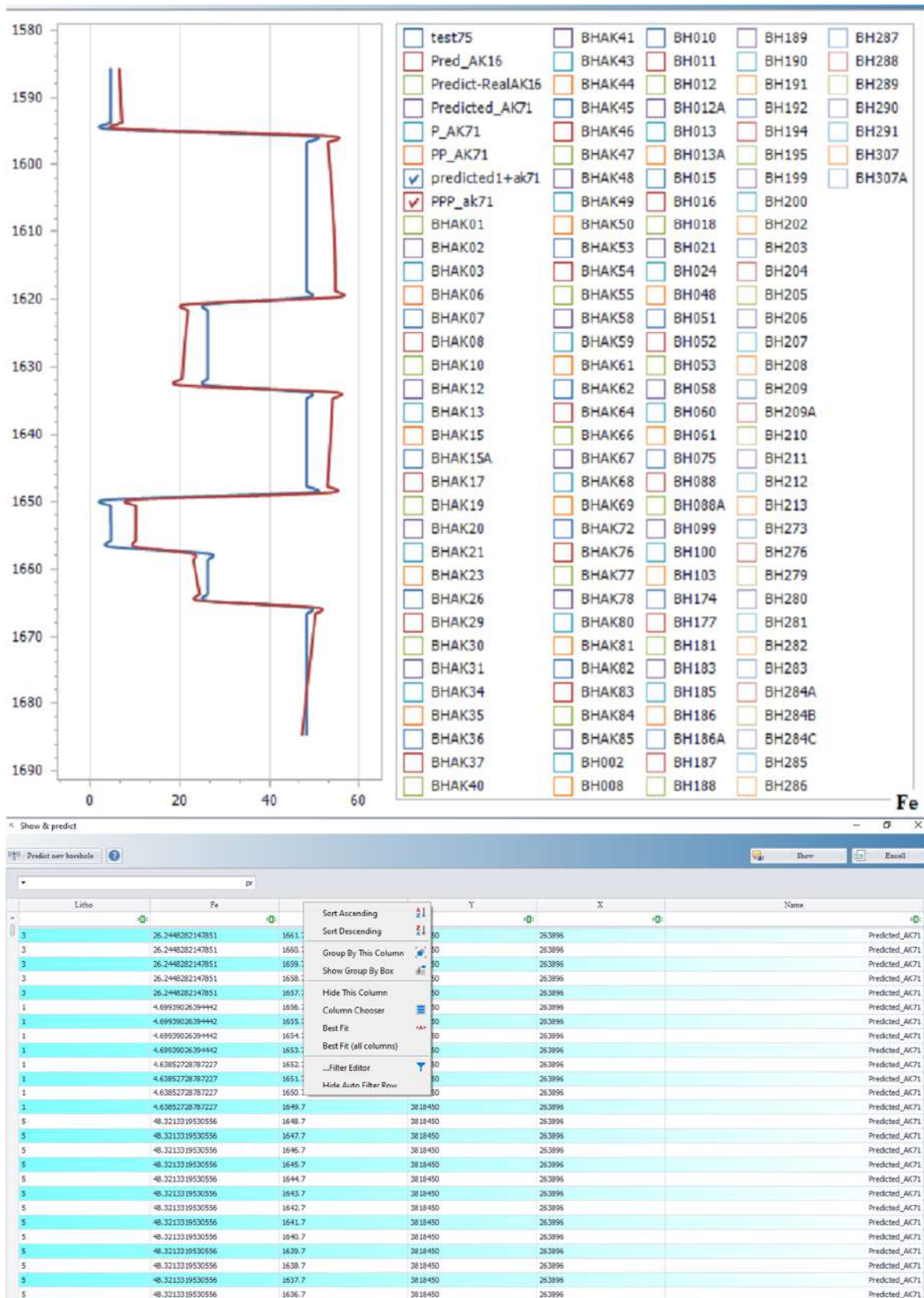


Figure 27. Pictures of the program environment written for the mineral grade estimation process.

## 5. Conclusions

The main problem in the use of machine learning in various sciences is the existence of accurate and appropriate data in terms of quantity and quality. In this paper, the most limiting parameter is the data parameter and its accuracy and dispersion. As seen in the clustering section, the amount of data in some clusters was very small, which in turn disrupts the network learning process and reduces network accuracy.

Combining data such as remote sensing, geochemistry and geophysics can help to increase the accuracy of the model. Accuracy and orderliness in sampling and drilling boreholes under the regular exploration network can also help to increase accuracy.

According to the results, by expanding this method, the cost of mineral exploration can be greatly reduced and a big step can be taken to optimize the exploration drilling network as the costliest part of mine exploration.

## Acknowledgments

The authors would like to express their gratitude to the Central Iran mines industries and the Kavoshgaran Consulting Engineers Company for their collaboration and support in permitting us to use the data that made this research work possible.

## References

- [1]. Hassani Pak, A.A., and Sharafodin, M. (2003). Exploratory Data Analysis, *University of Tehran Publications*.
- [2]. Kapageridis, I.K. et al. (1999). Application of Artificial Neural Network Systems to Grade Estimation from Exploration Data, *University of Nottingham Publications*.
- [3]. Moemeni, M., and Faal Ghyomi, A. (2011). Data Clustering (Cluster Analysis), *Tehran Moalef Publications*.
- [4]. Han, J., Kamber, and M., Pei, J. (2012). Data Mining Concepts and Techniques, *3th Edition, Elsevier Inc.* 243-278
- [5]. Raschka, S., and Patterson, J., Nolet, C. (2020). Machine Learning in Python: Main Developments & Technology Trends in Data Science Machine Learning & Artificial Intelligence, *arxiv*, 45.
- [6]. Kapageridis, I.K., and Denby, B.H. (1999). Ore Grade Estimation with Modular Neural Network Systems—A Case Study.
- [7]. Badel M., Angorani S., and Shariat Panahi, M. (2011). The Application of Median Indicator Kriging and Neural Network in Modeling Mixed Population in

an Iron Ore Deposit, *Computers & Geosciences* 37 (4). 530-540.

- [8]. Maleki, S., Ramazi, H., and Moradi, S. (2014). Estimation of Iron Concentration by using a Support Vector Machine and an Artificial Neural Network—the Case Study of the Choghart Deposit Southeast of Yazd, *Geopersia*, 4(2). 75-86.
- [9]. Abu Bakarr J., Sasakia, K., Yaguba, J., and Karim, B.A. (2016). Integrating Artificial Neural Networks and Geostatistics for Optimum 3D Geological Block Modeling in Mineral Reserve Estimation: A Case Study, *International Journal of Mining Science and Technology*, 26(4), 581-585.
- [10]. Nabilou M., Afzal P., Arian M., Adib A. and Kazemi Mehrnia A., (2018), The relationship between Fe mineralization and the basement linear structures using multi fractal modelling in the Bafgh area, Central Iran, *Geosciences*, 27(108), 181-190.
- [11]. Nezamolhosseini, S.A., Mojhedzadeh, S. H., and Gholamnejad, J. (2017). The Application of Artificial Neural Networks to Ore Reserve Estimation at Choghart Iron Ore Deposit, *Analytical and Numerical Methods in Mining Engineering*, 6(Special Issue), 73-83.
- [12]. Jafrasteh B. and Fathianpour, N. (2017). A Hybrid Simultaneous Perturbation Artificial Bee Colony and Back-propagation Algorithm for Training a Local Linear Radial Basis Neural Network on Ore Grade Estimation, *Neurocomputing*, 235, 217-227.
- [13]. Alimoradi A., Maleki B., Karimi A., Sahafzadeh M. and Abbasi S. (2020). Integrating Geophysical Attributes with New Cuckoo Search Machine-Learning Algorithm to Estimate Silver Grade Values—Case Study: Zarshouran Gold Mine, *Journal of Mining and Environment*, 11(3), 865-879.
- [14]. Fukunaga, K., and Hostetler L. D. (1975). The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition, *IEEE Transactions on Information Theory*, 21 (1). 32–40.
- [15]. Yizong, C. (1995). Mean Shift, Mode Seeking, and Clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17 (8). 790–799.
- [16]. Dorin, C., and Meer, P. (2002). Mean Shift: A Robust Approach Toward Feature Space Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (5). 603–619.
- [17]. Ho, Tin Kam (1995). Random Decision Forests, Proceedings of the 3rd International Conference on Document Analysis and Recognition, *Montreal, QC*, 278–282.
- [18]. Breiman, L. (2001). Random Forests, *Machine Learning*, 45 (1). 5–32.
- [19]. Dietterich, T. (2000). An Experimental Comparison of Three Methods for Constructing

Ensembles of Decision Trees: Bagging, Boosting, and Randomization, *Machine Learning*, 40 (2). 139–157.

[20]. Boehmke, B., and Greenwell, B. (2019). Gradient Boosting, Hands-on Machine Learning with R, Chapman & Hall, 221–245.

[21]. Elith, J. (2008). A Working Guide to Boosted Regression Trees, *Journal of Animal Ecology*, 77 (4). 802–813.

[22]. Friedman, J. (2003). Multiple Additive Regression Trees with Application in Epidemiology, *Statistics in Medicine*, 22 (9). 1365–1381.

[23]. Fathi M., Alimoradi A. and Hemati Ahoori H., (2021), Optimizing The Extreme Learning Machine Algorithm using Particle Swarm Optimization to Estimate Iron Ore Grade, *Journal of Mining and Environment*, 12(2), 397-411.

[24]. Majidi S. A., Lotfi, M., Emami, M. H., and Nezafati, N. (2017). The Genesis of Iron Oxide-Apatite (IOA) Deposits: Evidence from the Geochemistry of Apatite in Bafq-Saghand District, Central Iran, *Scientific Quarterly Journal*, Geosciences, 27(105), 233-244.

[25]. Nabilou M., Afzal P., Arian M., Adib A., Kheyrollahi H., Foudazi M. and Ansarirad P. (2022). The relationship between Fe mineralization and the magnetic basement structures using multifractal modeling in the Esfordi and Behabad Areas (BMD), Central Iran, *Acta Geologica Sinica – English Edition*, 96(2), 591-606.

[26]. Nabilou M., Afzal P., Arian M., Adib A., Kazemi Mehrnia A., Jami M., Kheyrollahi H., Akhavan Aghdam M.R., Ameri A. and Daneshvar Saein L. (2022), Determination of the relationship between Rare Earth Element (REE) mineralization and major faults using fractal modelling in the Gazestan deposit, Central Iran, *Bulletin of Geophysics and Oceanography*, 63(3), 495-518.

[27]. Yasrebi A.B., Hezarkhani A., Afzal P., Karami R., Eskandarnejad Tehrani M. and Borumandnia A. (2020). Application of an Ordinary Kriging-Artificial Neural Networks for Elemental Distribution in Kanag Porphyry Deposit, Central Iran. *Arabian Journal of Geosciences*, 13(748), 1-14.

[28]. Aliyari Ghassabeh, Youness (2013). On The Convergence of the Mean Shift Algorithm in the One-Dimensional Space, *Pattern Recognition Letters*, 34 (12). 1423–1427.

[29]. Carreira-Perpinan, and Miguel A. (2007). Gaussian Mean-Shift Is an EM Algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (5). 767–776.

[30]. Aliyari Ghassabeh, Youness (2015). A Sufficient Condition for the Convergence of the Mean Shift Algorithm with Gaussian Kernel, *Journal of Multivariate Analysis*, 135, 1–10.

[31]. Li, Xiangru; Hu, Zhanyi and Wu, Fuchao (2007). A Note on The Convergence of the Mean Shift, *Pattern Recognition*, 40 (6). 1756–1762.

[32]. Hastie, T.; Tibshirani, R.; and Friedman, J. (2008). The Elements of Statistical Learning (2nd ed.), Springer, 352-592.

[33]. Painsky A., and Rosset S. (2017). Cross-validated Variable Selection in Tree-based Methods Improves Predictive Performance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39 (11). 2142–2153.

[34]. Gareth James; Daniela Witten; Trevor Hastie and Robert Tibshirani (2013). An Introduction to Statistical Learning, Springer, 316–321.

[35]. Ho, Tin Kam (2002). A Data Complexity Analysis of Comparative Advantages of Decision Forest Constructors, *Pattern Analysis and Applications*, 5 (2). 102–112.

[36]. Friedman, J. (2003). Multiple Additive Regression Trees with Application in Epidemiology, *Statistics in Medicine*, 22 (9). 1365–1381.

[37]. E Bahri, A Alimoradi, M Yousefi (2023). Investigating The Performance of Continuous Weighting Functions in The Integration of Exploration Data for Mineral Potential Modeling Using Artificial Neural Networks, Geometric Average And Fuzzy Gamma Operators, *International Journal of Mining And Geo-Engineering*.

[38]. Haj Karimian, H., Alimoradi, A., Hemmati Ahoei, H. R., Salsabili, M. (2022). Comparison Between the Performance of Four Metaheuristic Algorithms in Training a Multilayer Perceptron Machine for Gold Grade Estimation, *International Journal of Mining and Geo-Engineering*, 56 (2), 97-105.

[39]. Fathi, S., Arab Amiri, A., Kamkar Rouhani, A., Alimoradi, A. (2019). Remote Sensing Studies and Geophysical Exploration using IP and Resistivity Methods in Zarshuran Gold Mining Area, Northwest of Iran, *Journal of Research on Applied Geophysics*, 5, 85-106.

## تعیین گسترش ناحیه آسیب‌دیده حفاری ناشی از اثر کارگاه استخراج جبهه کار طولانی

حمید محمدی<sup>۱</sup>، محمد علی ابراهیمی فرسنگی<sup>۱\*</sup>، حسین جلالی فر<sup>۱</sup>، علیرضا احمدی<sup>۲</sup> و علی جواهری<sup>۳</sup>

۱. بخش مهندسی معدن، دانشگاه شهید باهنر کرمان، ایران

۲. بخش مهندسی مکانیک، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، ایران

۳. شرکت زغال‌سنگ نگین طبس ایران، طبس، ایران

ارسال ۲۰۲۳/۰۹/۲۶، پذیرش ۲۰۲۳/۱۱/۰۸

\* نویسنده مسئول مکاتبات: maebrahimi@uk.ac.ir

### چکیده:

در روش معدنکاری جبهه کار طولانی پیشرو، ایمنی شبکه معدن، نرخ تولید و متعاقباً شرایط اقتصادی معدن وابسته به شرایط پایداری گالری‌ها است. پایداری گالری‌ها تابعی از دو عامل مهم است که عبارت‌اند از: (۱) خصوصیات ناحیه آسیب‌دیده حفاری در بالای گالری و (۲) اثر بارگذاری ناشی از ناحیه تخریب در بالای کارگاه جبهه کار طولانی که می‌تواند ناحیه آسیب‌دیده حفاری را گسترش دهد. عموماً در اثر شیب لایه زغال استخراجی، امکان وقوع شکست در گالری اصلی (گالری حمل‌ونقل) بیشتر از گالری تهویه است؛ بنابراین هدف از انجام این تحقیق، تعیین اثر کارگاه استخراج جبهه کار طولانی بر روی گسترش ناحیه آسیب‌دیده حفاری در بالای گالری اصلی است. برای رسیدن به این هدف، با در نظر گرفتن سه عامل خصوصیات ناحیه تخریب، خصوصیات لایه زغال (شیب و ضخامت) و خصوصیات ژئومکانیکی کمربالا، یک مدل هندسی جدید توسعه داده شد. سپس بر اساس محاسبات هندسی، یک رابطه جدید برای تعیین ضریب تأثیر کارگاه بدست آمد. همچنین با در نظر گرفتن مدل هندسی جدید، یک الگوریتم برای تحلیل پایداری گالری اصلی پیشنهاد شد. اعتبار سنجی مدل ارائه‌شده به‌وسیله نتایج ابزار بندی و رفتار نگاری یکی از کارگاه‌های استخراج جبهه کار طولانی معدن پروده ۲ طبس انجام شد. نتایج نشان دادند که توافق مناسبی بین نتایج مدل توسعه داده‌شده و مقادیر اندازه‌گیری شده وجود دارد. درنهایت، یک تحلیل حساسیت بر روی اثر عرض پایه، ظرفیت باربری سیستم نگهداری و شیب لایه زغال انجام شد.

**کلمات کلیدی:** گالری اصلی، ناحیه آسیب‌دیده حفاری، کارگاه استخراج جبهه کار طولانی، ناحیه تخریب، مدل هندسی جدید.