Impulse Ground Pressure Prediction Based on Rough Set-Distance Discriminant Model

Jian Ping*
Teaching and Research Department of Social Construction and Ecological Civilization, CPC Liaoning Provincial Party School, Shenyang China

*Corresponding author email: 1187076493@qq.com

Abstract. In this paper, introduce the rough set theory, and the weight coefficient is calculated by analyzing the support degree and attribute importance of the evaluation object. Combine with the distance discriminant model, successfully construct the weighted distance discriminant model of rock burst, and determine the reasonable discriminant factors. Taking 15 groups of data from Yantai Coal Mine as training samples, and establish the rough set distance discriminant model for rock burst prediction. The discriminant results of this model were compared with those obtained by Fisher analysis discriminant model (FDA), R-type factor Fisher discriminator method, etc. Through the comparison of the judgment results, prove the effectiveness of the through-distinction model in the field of mine water source discrimination. The results show that the through set distance discrimination model has the advantages of high prediction jurisdiction and strong rationality, and it is a new way suitable for mine water source discrimination.

Keywords: Rough set theory; Distance discrimination method; Weight.

1. Introduction
With the maturity of technology, the depth of mine mining is getting deeper and deeper, and the problem of impact ground pressure has become more and more prominent. At present, the impact ground pressure is still a research hotspot, Shi et al[1] combined with fuzzy comprehensive evaluation method, establish the model of impact ground pressure risk analysis using principal component analysis method, and proved the effectiveness of the model; Lu et al[2] established the prediction particle (PSO)-optimized least square support vector machine (LSSVM) model of impact ground pressure classification, and the accuracy and simplicity of the model are proved by experiments; Zhang et al[3] by designing the real-time transmission system of impact ground pressure monitoring data, improve the accuracy of monitoring and early warning; Bi et al[4] established a Fisher discriminant prediction model of R-type factor analysis of rock bursts. The advantage of this model lies in weakening the mutual influence between indicators. Zhou et al[5] used Fisher discriminant and combined the corresponding linear discriminant function to use the back-substitution estimation method for back-checking, the results are in good agreement with the actual situation. Most of the above scholars ignore the influence of subjective factors on the experimental results, which is slightly different from the actual situation.

In this paper, the importance of evaluation index is distinguished by establishing rough set-distance discriminant model, and give the weight coefficient[6], This model effectively circumvents subjective factors by determining the index weight coefficients and is more realistic.

2. Basic Ideas
The basic idea of rough set-distance discriminant analysis: Select a reasonable evaluation index for rock
burst, through sample data processing and training, combine with distance discrimination method to get an effective rough set-distance discriminant model, and use the sample to test the validity of the model, compare the discriminant results of the model with the discriminant results of other discriminants and finally verify the accuracy of the discriminant results.

3. Solutions

3.1. Rough Theory
Rough set is another mathematical tool for analyzing uncertain data after probability theory, fuzzy set, and evidence theory. It has the advantages of no prior knowledge, processing incomplete information and simplifying data[7], effectively avoiding the influence of subjective factors.

Let the quaternion \( M=(U,A,V,F) \) be an information system, where \( U=(x_1,x_2…x_n) \) is a non-empty finite set of objects, which is called the universe of discourse; \( A \) is a set with \((n+q)\) attributes or characteristics; \( V \) is a set of attributes and functions from \( U-A \) to \( V \); The attribute satisfies \( A=C \cup D \), where \( C \) is an \( n \)-dimensional conditional attribute, and \( D \) is an \( m \)-dimensional decision attribute. For each element \( a \) in \( A \), the corresponding value can be found in \( V \), that is, the combination of attributes is also the value range of \( A \).

3.2. Attribute Reduction
Attribute reduction is to remove the redundant parts which are not important for decision-making, reduce the amount of computation and save the training time, and improve the decision-making ability while maintaining the classification ability[8]. There are two kinds of general reduction and relative reduction. This paper adopts relative reduction, that is, the reduction of conditional attribute relative to decision attribute. The concept of relative reduction is as follows:

Let the domain \( U \) contain two equivalence relations \( P \) and \( Q \), define the relative certainty of \( Q \) with respect to \( P \), Denoted as \( Pos_p(Q) \). It is a set of all those objects in the \( U \) domain, which can be correctly classified into the equivalence class under the guidance of the knowledge of \( U/P \) classification, as follows:

\[
Pos_p(Q) = \cup P_-(X) \quad (X \in U/Q)
\]

\( P_-(X) \) is the lower approximation of the set \( X \).

Let the domain \( U \) contain two equivalence relations \( P \) and \( Q \), and \( r \in P \). If

\[
Pos_p(Q) = Pos_{(P-(\{r\})\cup Q)}
\]

Then it is said that \( r \) can be omitted with respect to \( Q \), otherwise it cannot be omitted. In particular, when \( P-\{r\} \) is an independent subset of \( P \)(that is, each element of it can no longer be omitted, and \( Pos_p(Q) = Pos_{(P-(\{r\})\cup Q)} \), then call \( IndQ(P) \) the relative reduction of \( P \) with respect to \( Q \). The set of all attributes that cannot be omitted in \( P \) is called the core of \( P \), Denoted as \( Core_Q(P) \)

\[
Core_Q(P) = \cap IndQ(P)
\]

3.3. Decision Attribute Support and Attribute Importance
If \( n \)-dimensional conditional attributes \( C = (X_1,X_2,…X_n) \), \( m \)-dimensional decision attributes \( D = (Y_1,Y_2,…Y_m) \), The support of decision attribute \( D \) with regard to condition attribute \( C \) can be defined as:

\[
k = y_c(D) = |Pos_c(D)|/|U| = \frac{1}{u} \sum_{i=1}^{m} |y_c(D)|, \quad Y_i \in U/D
\]

The \( k \) in formula (1) indicates the degree of support of conditional attribute \( C \) to decision attribute \( D \), and the specific relationship can be expressed as[9]:

- is entirely up to \( C \), \( k=1 \)
- part is decided by the \( C \), \( 0 < k < 1 \)
- is unaffected by \( C \), \( k=0 \)

Since the importance of each condition attribute subset \( C_i \) in the \( n \)-dimensional condition attribute \( C \) is not the same, the importance of each \( C_i \) related to the decision attribute \( D \) is also different. Therefore, this importance can be defined as:
It can be seen from the above formula, \( \theta_{CD}(C_i) \) can directly reflect the importance of each conditional attribute subset \( C_i \) in the n-dimensional conditional attribute \( C \). The greater the \( \theta_{CD}(C_i) \) the greater the above importance is also true\(^9\).

3.4. Weight Factor

The relative importance of each indicator is displayed by weight. After establish the data model, the condition attribute set \( C \) and the decision attribute set \( D \) are determined. Finally, form the initial decision table, and reduce attributes, and successfully obtain the optimal index combination. Under the premise of using formulas (1)-(3) to calculate the support of decision attributes and attribute importance, the weight coefficient can be defined as:

\[
\varphi_i = \frac{y_{c-c_i}(D)}{\sum_{j=1}^{m} y_{c-c_j}(D)} \quad i = (1, 2, \ldots, n)
\]  

4. Rough Set-Distance Discrimination Model Establishment

The Markov distance discrimination method relies on the classification of the distance between the sample and the population. Multiple overall discriminants can be defined as\(^{10}\):

\[
d^2(X, G) = (X - \mu)^T \Sigma^{-1}(X - \mu)
\]  

This method is easy to ignore the difference of influence factors of each prediction index, but in fact the function of each influence factor in a certain population is not exactly the same. In this paper, the index weight coefficient matrix of sample \( X \) is calculated by rough set theory \( W \), and the \( W \) is embedded in the distance discriminant analysis model, which highlights the importance of different indexes. The rough set-distance discriminant model is established:

\[
d_w^2(X, G) = (X - \mu)^T W \Sigma_w^{-1}(X - \mu)
\]  

In the formula, \( W = \begin{pmatrix} w_1 & \ldots & w_n \end{pmatrix} \) is the weight factor of each evaluation index in the distance discriminant function, taking the example in document 11 as the research object, and combining the model to select the data, Finally, 15 groups of data are selected as training samples (Table 1). The other data can be considered as engineering examples.

\[
d^2(X, G_i) = \min_{1 \leq i \leq k} \left[ d^2(X, G_i) \right] \quad x \in G_L
\]  

Use rough set-distance discriminant model. The calculation steps and process diagram of the evaluation index weight coefficient for the theoretical discrimination of rock burst (1)
Figure 1. Flow chart of identifying rock burst

Table 1. Rock burst training samples and prediction results

| Serial number | $X_1$ (Coal thickness) | $X_2$ (Inclination) | $X_3$ (buried depth) | $X_4$ (Fracture) | $X_5$ (buried depth) | $X_6$ (Gas mass concentration) | $X_7$ (Roof management) | Actual type | Forecast results |
|---------------|------------------------|---------------------|----------------------|-----------------|---------------------|-------------------------------|------------------------|-------------|-----------------|
| 1             | 2.5                    | 65                  | 450                  | 4               | 3                   | 4                             | 0.67                   | 2           | 2               | Strong          |
| 2             | 1.5                    | 65                  | 570                  | 2               | 4                   | 2                             | 0.28                   | 2           | 3               | Middle          |
| 3             | 1.7                    | 60                  | 314                  | 4               | 4                   | 2                             | 1.3                    | 1           | 1               | Strong          |
| 4             | 3                      | 24                  | 573                  | 1               | 1                   | 2                             | 0.36                   | 3           | 4               | Small           |
| 5             | 3                      | 35                  | 612                  | 3               | 1                   | 3                             | 0.56                   | 3           | 1               | Middle          |
| 6             | 1.1                    | 47                  | 485                  | 4               | 4                   | 4                             | 0.43                   | 2           | 1               | Strong          |
| 7             | 1.6                    | 35                  | 583                  | 3               | 1                   | 4                             | 1.5                    | 4           | 4               | Weak            |
| 8             | 1.5                    | 35                  | 530                  | 1               | 1                   | 1                             | 0.56                   | 4           | 4               | Small           |
| 9             | 1.6                    | 62                  | 307                  | 4               | 3                   | 3                             | 1                      | 1           | 1               | Strong          |
| 10            | 1.9                    | 59                  | 542                  | 2               | 3                   | 4                             | 0.25                   | 1           | 1               | Middle          |
| 11            | 1.8                    | 62                  | 283                  | 4               | 3                   | 4                             | 1                      | 1           | 1               | Strong          |
| 12            | 1.3                    | 44                  | 570                  | 1               | 1                   | 1                             | 0.66                   | 4           | 4               | Small           |
| 13            | 2.2                    | 54                  | 290                  | 4               | 3                   | 3                             | 1                      | 1           | 1               | Strong          |
| 14            | 3                      | 34                  | 475                  | 3               | 3                   | 2                             | 0.42                   | 1           | 1               | Middle          |
| 15            | 1.8                    | 54                  | 238                  | 4               | 2                   | 4                             | 1                      | 1           | 1               | Strong          |
5. Engineering Example Application

Five valid data (table2) are selected as the engineering case data in this paper according to the relevant data provided in reference 11. the discriminant results are shown in table2. the obtained discriminant results are completely consistent with the actual situation. in addition, the prediction results are compared with the prediction results of Fisher analysis discriminant model (FDA), r-factor Fisher discriminant method. it is found that the results analyzed by rough set-distance discriminant method are more appropriate to the actual situation, which shows that the rough set-distance discriminant model has good applicability and effectiveness.

| Serial number | Location     | Location     | Actual type | Comparison of Discriminant Results | FDA | r-factor Fisher | This article |
|---------------|--------------|--------------|-------------|-----------------------------------|-----|----------------|--------------|
| 1             | 3601N section moulding | 6#          | Weak        | Weak                             | Weak | Weak          | Weak         |
| 2             | 4603N cutting   | 6#          | Middle      | Middle                           | Middle | Middle     | Middle       |
| 3             | 4401N segment-238 cut | 4#          | Strong      | Strong                           | Strong | Strong     | Strong       |
| 4             | 4402S -60 Coal Lane | 4# Middle East | Middle      | Middle                           | Middle | Middle     | Middle       |
| 5             | 3404N cutting   | 4#          | Middle East | Middle                           | Middle | Middle     | Middle       |

6. Concluding Remarks

By using the Rosetta software to reduce the attribute of the selected data, the interference of redundant attributes is effectively reduced, and the subjective factors are reasonably avoided by combining the weight coefficient to make the whole process more suitable for the actual situation. The 10 discriminant factors can reflect the stability of mine water source and provide a new theoretical basis for judging mine water source. In practical engineering, it can effectively predict the occurrence of rock burst, which is of great significance to the formulation of mine local or regional prevention measures.

References

[1] Shimo, Shi Jiadong. FCA-PCA Model of Impact Ground Pressure Hazard Assessment and its Application [J].]1 Coal, 2021,30(04):44-46.
[2] Lu Pengfei, Qiu Lin. Study on Prediction of Mine Impact Ground Pressure Grading Based on PSO-LSSVM J]. Mining Safety and Environmental Protection, 2021,48(01):120-125.
[3] Zhang Shuancai, Han Zepeng, Chen Runhe, Wang Fuqiang, Gong Siyuan. GIS cloud platform for real-time monitoring and warning of impact ground pressure Coal Mine Safety, 2020,51(10):243-247.
[4] Bi Juan, Li Xijian, Chen Liu Yu. Discriminating [J.]. R Factors for Prediction of Impact Ground Pressure Hazard Grade Chinese Journal of Safety Sciences, 2019,29(12):103-109.
[5] Zhou Jian, Shi Xiuzhi. Fisher discriminant analysis method [J.] for prediction of impact ground pressure risk grade Coal Journal 2010 35(S1):22-27.
[6] Yan Changbin. rough set-distance discriminant model for slope stability prediction and its application [J.]; and Journal of Engineering Geology 24(02):204-210.
[7] Qian Feng. Discrete Method and Reduction Algorithm Based on Rough Set Knowledge [D].] Study Xiamen University, 2007.
[8] Wu Yangyang, Tang Jianguo. Advances in attribute reduction of rough set in big data [J.]. background Computer Engineering and Applications ,2019,55(06):31-38+177.
[9] Tao Zhi, Xu Baodong, Wang Dingwei. A Knowledge Reduction Method [J.] Based on Decision Attribute Support Journal of Northeastern University, 2002(11):1025-1028.
[10] Huang Deyong, Geng Yuanling, Guo Qi, du Zhijin, Yang Yang. A Study on the Evaluation
Method of Slope Stability Based on the Combined Weighted Distance Discriminating Model [J.]; and Nonferrous Metals Engineering, 2021,11(01):101-106+136.

[11] Yan he. A Study on Impulse Ground Pressure Prediction Based on Genetic Algorithm and Artificial Neural Network [D.]; and Chongqing University, 2002.