

Article



Bagged Ensemble of Gaussian Process Classifiers for Assessing Rockburst Damage Potential with an Imbalanced Dataset

Ying Chen ^{1,2}, Qi Da ¹, Weizhang Liang ^{3,*}, Peng Xiao ³, Bing Dai ¹, and Guoyan Zhao ³

- ¹ School of Resource Environment and Safety Engineering, University of South China, Hengyang 421001, China
- ² China Tin Group Co., Ltd., Liuzhou 545026, China
- ³ School of Resources and Safety Engineering, Central South University, Changsha 410083, China
- Correspondence: wzlian@csu.edu.cn

Abstract: The evaluation of rockburst damage potential plays a significant role in managing rockburst risk and guaranteeing the safety of personnel. However, it is still a challenging problem because of its complex mechanisms and numerous influencing factors. In this study, a bagged ensemble of Gaussian process classifiers (GPCs) is proposed to assess rockburst damage potential with an imbalanced dataset. First, a rockburst dataset including seven indicators and four levels is collected. To address classification problems with an imbalanced dataset, a novel model that integrates the under-sampling technique, Gaussian process classifier (GPC) and bagging method is constructed. Afterwards, the comprehensive performance of the proposed model is evaluated using the values of accuracy, precision, recall, and F_1 . Finally, the methodology is applied to assess rockburst damage potential in the Perseverance nickel mine. Results show that the performance of the proposed bagged ensemble of GPCs is acceptable, and the integration of data preprocessing, under-sampling technique, GPC, and bagging method can improve the model performance. The proposed methodology can provide an effective reference for the risk management of rockburst.

Keywords: rockburst; damage potential; Gaussian process classifier (GPC); bagging method; imbalanced dataset

MSC: 90B50; 94D05

1. Introduction

With the increase in mining depth, rockbursting has become an increasingly prominent issue [1–3]. It is induced by the instantaneous release of elastic strain energy, and often is accompanied by ejection and collapse of massive rock [4–6]. Many mines have suffered rockburst disasters, causing serious economic losses and casualties. For example, a rockburst with a magnitude of 3.5 happened in Falconbridge nickel mine, resulting in four deaths [7]; a rockburst with a magnitude of 2.47 occurred in Junde coal mine, causing five deaths and the destruction of a shearer and scraper conveyor [8]; and a rockburst with a magnitude of 5.2 appeared in the Klerksdorp district of South Africa, leading to two deaths and fifty-eight injuries [9]. Due to such serious consequences, assessing rockburst damage potential is necessary and significant.

According to the differences between the locations of damage and seismic event, rockbursts can be classified into self-initiated rockburst and remotely triggered rockburst [10]. For the former, the locations of damage and seismic event are consistent. While for the latter, rockburst is triggered by remote and relatively large magnitude seismic events. For different types of rockburst, the influencing factors are different, resulting in the disparity of the rockburst damage potential evaluation. This study aims to assess the damage potential of remotely triggered rockburst. Because the location of the damage is not consistent with that of the microseismic event, it is difficult to evaluate the damage potential only based



Citation: Chen, Y.; Da, Q.; Liang, W.; Xiao, P.; Dai, B.; Zhao, G. Bagged Ensemble of Gaussian Process Classifiers for Assessing Rockburst Damage Potential with an Imbalanced Dataset. *Mathematics* 2022, 10, 3382. https://doi.org/ 10.3390/math10183382

Academic Editor: Maria C. Mariani

Received: 6 August 2022 Accepted: 14 September 2022 Published: 17 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). on the microseismic event. The microseismic event information, stress wave propagation paths, and rock mass conditions on the excavation face should be considered simultaneously. Due to the complex mechanisms and numerous influencing factors, the evaluation of rockburst damage potential is still a difficult issue.

Scholars have proposed some methods to assess rockburst damage potential. Kaiser et al. [11] developed a rockburst damage assessment procedure. It mainly included four steps: propose rock and support damage scales, put forward an initial condition index, calculate the scaled distance, and establish relationships among the initial condition index, scaled distance, and rock and support damage scales. Durrheim et al. [12] summarized the influencing factors of rockburst damage according to the investigations of rockbursts in South African gold mines, which was valuable for the evaluation of rockburst damage potential. Brink et al. [13] proposed an approach for seismic risk evaluation, which can be summarized in four steps: determine an evaluation indicator system, score each subcategory according to the risk rating, calculate the score of each category, and determine the risk levels. Albrecht and Sharrock [14] investigated ten indicators that affect rockburst damage, and established the relationship between them and rockburst damage based on field rockburst incidences. With the increase of rockburst cases, machine learning (ML) algorithms were used to evaluate rockburst damage potential. Heal et al. [15] proposed the concept of excavation vulnerability potential (EVP), and then adopted logistic regression to assess rockburst damage potential. Zhou et al. [16] employed a stochastic gradient boosting approach for the evaluation of rockburst damage. Li et al. [17] put forward a rockburst damage scale index using rock engineering systems and artificial neural networks to evaluate rockburst damage.

When a large number of rockburst cases accumulate, ML is a possible way to evaluate rockburst damage potential [18,19]. However, due to the fact that most rockburst damage levels are slight, while the strong or even extremely strong type is relatively rare, the distribution of sample data for each level is usually imbalanced [20–22]. Considering the specific characteristics of rockburst data, two key issues need to be solved. The first one is the handling of the imbalanced rockburst dataset. Generally, classical ML algorithms are conceived on the premise of balanced datasets [23]. It is difficult to handle classification problems with an imbalanced dataset, especially for discriminating the minority category cases [24]. Therefore, traditional ML algorithms should be improved to deal with imbalanced datasets. The corresponding strategy can be roughly divided into four groups: data level, algorithm level, cost-sensitive level, and ensemble level [25]. Among them, combining bagging ensemble learning with under-sampling techniques is an effective way to deal with imbalanced datasets [26].

The second one is the selection of algorithms. A large number of ML algorithms have been used to solve classification problems. Although some other statistical algorithms, such as Monte Carlo methods, can also be adopted to solve multidimensional problems and obtain probabilistic results, the probability density functions need to be determined in advance [27–29]. Gaussian process classifier (GPC) is a promising statistical model because it can deal with high-dimensional and nonlinear problems, tune hyperparameters directly based on training data, and obtain probabilistic outputs [30,31]. However, due to the characteristics of imbalance and strong noise in rockburst data, a single GPC is hard to have stable prediction ability. Ensemble learning can overcome this drawback by combining multiple base classifiers to some extent [32–34]. Combining bagging ensemble learning with Gaussian process classifiers (GPCs) may improve the generalization ability and robustness of models.

This study proposes a novel model that integrates the under-sampling technique, GPC, and bagging method to assess rockburst damage potential with an imbalanced dataset. First, the rockburst dataset is collected and preprocessed by the Yeo-Johnson transformation and standardization process. Then, the reliability of the proposed methodology is verified, and the comprehensive performance is evaluated using four metrics. Finally, the

proposed bagged ensemble of GPCs is applied to assess rockburst damage potential in the Perseverance nickel mine.

2. Data Acquisition

According to the original work of Heal [35], a total of 254 rockburst cases were collected from 13 underground metal mines in Canada and Australia. These cases were obtained based on the rock mass failure conditions caused by a single microseismic event. This database contains 83 microseismic events and 254 failure locations. It indicates some failure locations are caused by the same microseismic event. Based on the damage status of rock mass and support, the degree of rockburst damage was divided into five levels: none (L_1) , low (L_2), moderate (L_3), high (L_4), and strong (L_5). Among them, L_1 indicated the rock mass showed no damage or minor loss, and the support was not damaged; L_2 indicated the rock mass was slightly damaged, less than 1 ton of rock was displaced, the support system was loaded, the meshes were loose and the plates were deformed; L_3 indicated 1 ton to 10 tons of rock was displaced and some bolts were broken; L₄ indicated 10 tons to 100 tons of rock was displaced and the support system was severely damaged; L_5 indicated above 100 tons of rock was displaced and the support system was completely destroyed. As the damage locations of L_1 were not reported during investigations, the original database only contained L_2 , L_3 , L_4 , and L_5 . The sample sizes at these levels were 116, 48, 63, and 27, respectively.

The original database included nine indicators: the ratio of total maximum principal stress to uniaxial compressive strength (I_1), the energy capacity of support system (I_2), excavation span (I_3), geology factor (I_4), Richer magnitude of seismic event (I_5), distance between rockburst location and microseismic event (I_6), peak particle velocity (I_7), rock density (I_8) and support types (I_9). The specific meaning of each indicator can be referred to in literature [35]. Different indicator combinations have a significant impact on evaluation results. Heal [35] and Zhou et al. [16] selected I_1 , I_2 , I_3 , I_4 and I_7 ; and Li et al. [17] chose I_1 , I_2 , I_3 , I_4 , I_5 , I_7 and I_8 . Considering I_7 was calculated by I_5 and I_6 based on an empirical formula and I_9 was difficult to be quantified, this study adopted I_1 , I_2 , I_3 , I_4 , I_5 , I_6 and I_8 to evaluate the rockburst damage.

From this dataset, it can be seen that there were some duplicate sample data, and some samples had the same indicator values, but the corresponding rockburst damage levels were different. To improve the prediction accuracy, the duplicate samples were first removed. For the samples with the same indicator values but different levels, only the samples with the highest level were selected to ensure safety. Consequently, the number of samples in the updated dataset was 236, and the sample sizes at L_2 , L_3 , L_4 and L_5 were 107, 45, 57 and 27, respectively. The corresponding ratio of sample sizes at different levels was 4.0:1.7:2.1:1.0. It shows that the distribution is relatively unbalanced, which may affect the accuracy of evaluation results. The detailed dataset was listed in Appendix A.

To quantitatively analyze the correlations between these seven indicators, the heat map of the correlation coefficient was obtained, as shown in Figure 1. It can be seen that some indicators were positively correlated, such as I_1 and I_2 , whereas some indicators were negatively correlated, such as I_1 and I_3 . Overall, the correlations between these indicators were generally small. Although I_5 and I_6 had the largest correlation coefficient of 0.48, they were distinctly based on their physical meanings. Therefore, these indicators were relatively independent, which verified the rationality of the selected indicators.



Figure 1. Heatmap of the correlation coefficient.

The box plot of all indicators for each rockburst damage level was shown in Figure 2. It can be seen that all indicators had some outliers. Especially for I_2 , I_3 , I_6 and I_8 , outliers were more obvious. Some overlapping parts existed in the range of indicator values for various levels. As a result, it was difficult to differentiate the level of rockburst damage only using one indicator. Second, there was no obvious correlation between rockburst damage level and each indicator. In addition, the distribution of indicator values was uneven. All these characteristics illustrated the complexity of rockburst damage evaluation.





Figure 2. Box plot of indicators for each level.

3. Methodology

3.1. Gaussian Process Classifier

GPC is a statistical learning algorithm based on the Gaussian process and Bayesian theory, which has a solid mathematical foundation. By assuming the implicit function obeys the prior distribution of a Gaussian process, the posterior distribution can be obtained according to Bayesian inference [36]. Then, the probability of different classes can be determined. The main calculation steps are as follows.

Suppose the training set is:

$$D = (X, Y) = \{(x_i, y_i) | y_i = \pm 1, i = 1, 2, \cdots, m\},$$
(1)

where $x_i = (x_{1i}, x_{2i}, \dots, x_{Bi})$ is the input; y_i is the output; and m is the number of samples in the training set.

To reflect the mapping relationship between x_i and y_i , the implicit function that obeys the Gaussian process distribution is defined as:

$$f = [f(x_1), f(x_1), \cdots, f(x_i), \cdots, f(x_m)]^T,$$
(2)

Suppose *f* satisfies a Gaussian process distribution with a zero mean and covariance matrix *K*, then:

$$p(f|X) \sim N(f|0,K), \tag{3}$$

where *K* can be calculated by a covariance function k(x, x'), and is specifically defined according to the actual situation.

In general, the radial basis function is selected as the covariance function:

$$k(x, x') = \theta_1 e^{-\frac{\|x - x'\|^2}{\theta_2}},$$
(4)

where θ_1 and θ_2 are hyperparameters.

Based on Equations (3) and (4), the prior probability can be determined as:

$$p(f|X) = \frac{1}{(2\pi)^{0.5} |K|^{0.5}} \exp(-0.5f^T K^{-1} f).$$
(5)

Then, to obtain the probability of the predicted category, a likelihood function is used to map the output value of the implicit function to the interval [0, 1]. The logistic function is generally used as the likelihood function:

$$p(Y|f) = \psi(z) = \frac{1}{1 + \exp(-z)}.$$
 (6)

Based on Bayes' theorem, the posterior probability of the implicit function is:

$$p(\hat{f}|X,Y) = \frac{p(Y|f)p(f|X))}{p(Y|X)},$$
(7)

where p(Y|X) is the marginal likelihood function, which indicates the probability distribution of a training set.

p(Y|X) can be calculated by:

$$p(Y|X) = \int p(D|f)p(f)df.$$
(8)

Suppose the sample to be predicted is (\tilde{x}, \tilde{y}) , the probability of $\tilde{y} = +1$ can be determined by:

$$p(\tilde{y}=+1|X,Y,\tilde{x}) = \int p(\tilde{y}|\hat{f})p(\hat{f}|X,Y,\tilde{x})d\hat{f},$$
(9)

where \hat{f} indicates the implicit function of \tilde{x} .

Since there is no analytical solution in Equations (7)–(9), Laplace approximation algorithm is often to obtain the solutions. Namely, the posterior probability distribution $p(\hat{f}|X, Y)$ is first obtained, then the implicit function \hat{f} can be determined.

Finally, the probability of $\tilde{y} = +1$ can be calculated by

$$p(\tilde{y}=+1|X,Y,\tilde{x}) = \int \psi(\hat{f}) p(\hat{f}|X,Y,\tilde{x}) d\hat{f}.$$
(10)

If $p(\tilde{y}=+1|X, Y, \tilde{x}) \ge 0.5$, then the prediction result is a positive class, otherwise it is a negative class.

For multi-classification issues, the binary Gaussian process classifier can be extended with a "one-vs-rest" or "one-vs-one" strategy [37]. For the "one-vs-rest" strategy, the binary Gaussian process classifier classifies one of the classes and the remaining classes respectively. In this case, the class with the highest probability is selected as the final result. For the "one-vs-one" strategy, the binary Gaussian process classifier classifies the two classes respectively. In this case, each classification is equivalent to one vote, and the class with the highest votes is selected as the final result.

3.2. Bagged Ensemble of Gaussian Process Classifiers

A bagged ensemble of GPCs is proposed to handle classification problems with imbalanced datasets, as shown in Figure 3. This model integrates the under-sampling technique, GPC and bagging method. The under-sampling technique is used to make the training samples balanced. The samples of classes except the minority class are resampled and combined with the minority class samples into a new dataset. GPC has a strong probabilistic prediction ability for unknown data by learning from the existing dataset. By integrating multiple classifiers, the bagging method can avoid over-fitting to a certain extent, and has better anti-noise ability and robustness. In addition, the defect of data loss from a single under-sampling can be overcome through multiple under-samplings with replacement. The specific steps are as follows.



Figure 3. Diagram of a bagged ensemble of Gaussian process classifiers.

First, the under-sampling technique with replacement is used to generate the balanced sample sets from the original dataset.

Second, the GPCs are independently trained based on the generated training sets.

Last, the final result is obtained by integrating the evaluation results of each GPC based on a voting classifier.

3.3. Establishment of Rockburst Damage Evaluation Model

The proposed ensemble model is used to evaluate rockburst damage. The detailed procedure is shown in Figure 4, which is described as follows.



Figure 4. The procedure of the ensemble model for the evaluation of rockburst damage.

First, the original rockburst damage dataset is preprocessed based on the Yeo-Johnson transformation and standardized processing. In many modeling scenarios, data needs to be normalized to improve predictive performance. Power transformation maps sample data from an arbitrary distribution to a Gaussian distribution as close as possible. It builds a set of monotonic functions to stabilize variance and minimize skewness. There are two transformation methods: the Yeo-Johnson and Box-Cox transformation. Since the Box-Cox transformation only works for positive data, the Yeo-Johnson transformation is adopted in this study. The calculation formula is [38]:

$$x_{i}^{(\lambda)} = \begin{cases} \left\lfloor (x_{i}+1)^{\lambda}-1 \right\rfloor \lambda & \text{if } \lambda \neq 0, x_{i} \geq 0\\ \ln(x_{i}+1) & \text{if } \lambda = 0, x_{i} \geq 0\\ -\left[(-x_{i}+1)^{2-\lambda}-1 \right] / (2-\lambda) & \text{if } \lambda \neq 2, x_{i} < 0'\\ -\ln(-x_{i}+1) & \text{if } \lambda = 2, x_{i} < 0 \end{cases}$$
(11)

where x_i is the data to be transformed; and λ is a parameter, which can be estimated by the maximum likelihood method.

In addition, there is a large gap between some indicator values. For example, I_8 is three orders of magnitude larger than I_5 . In this case, it may lead to the dominance of this indicator, while the roles of other indicators are ignored. Therefore, the initial indicator values need to be firstly standardized. In this study, they are converted into a standard

normal distribution with a mean of zero and a standard deviation of one. The conversion formula is:

$$x_i^{(\lambda)} = (x_i^{(\lambda)} - \mu) / \sigma,$$
 (12)

where μ is the mean value and σ is the standard deviation of sample data.

Second, the preprocessed dataset is randomly divided into training and test sets with a ratio of 4:1. Furthermore, the ratio of sample size for different levels in these two sets is kept consistent to make the results more stable.

Third, the hyperparameter of the bagged ensemble of GPCs is optimized using fivefold cross-validation. The number of GPC is adopted as the hyperparameter to be optimized, and both the hyperparameters θ_1 and θ_2 in kernel function of GPC are selected as 1.0. Then, the optimal hyperparameter value is determined based on the average accuracy of five-fold cross-validation.

Fourth, the model with the optimal hyperparameter value is fitted based on the training set, and then the optimal training model is obtained.

Fifth, the comprehensive performance of the proposed methodology is evaluated based on the test set. The accuracy, precision, recall, and F_1 are chosen as the evaluation metrics, which can be calculated by a confusion matrix.

Suppose the confusion matrix is:

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1q} \\ s_{21} & s_{22} & \cdots & s_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ s_{q1} & s_{q2} & \cdots & s_{qq} \end{bmatrix},$$
(13)

where *q* is the number of levels.

Then, the accuracy can be calculated by

Accuracy =
$$\frac{1}{\sum_{j=1}^{q} \sum_{k=1}^{q} s_{jk}} \sum_{j=1}^{q} s_{jj};$$
 (14)

The precision can be calculated by

Precision
$$= \frac{s_{jj}}{\sum\limits_{j=1}^{q} s_{jk}};$$
 (15)

The recall can be calculated by

$$\operatorname{Recall} = \frac{s_{jj}}{\sum\limits_{k=1}^{q} s_{jk}};$$
(16)

The F_1 can be calculated by

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$
(17)

Finally, if the prediction performance is reliable, the entire preprocessed dataset can be used as the training set to fit the model. Then, the rockburst damage level in actual engineering can be evaluated. Conversely, if the prediction performance is unreliable, improvements can be made in terms of database quality, data preprocessing, and evaluation models. Moreover, new cases can be adopted to update the original rockburst dataset, and the evaluation process in the next stage can be conducted.

4. Validity Verification

The collected rockburst dataset was used to verify the feasibility of the proposed methodology. Based on Equations (11) and (12), all indicator values were preprocessed. The distribution of indicator values before and after preprocessing was shown in Figure 5. The values of all indicators after preprocessing followed the standard normal distribution with a mean of zero and a standard deviation of one.

To make the model performance more reliable, the number of GPC was optimized by using the five-fold cross validation based on the training set. The average accuracy of five-fold cross validation corresponding to different numbers of GPC was shown in Figure 6. It can be seen that the average accuracy does not increase with the number of GPC, and the optimal number of GPC was 12 because of its maximum average accuracy.



Figure 5. Cont.



Figure 5. Distribution of indicator values before and after preprocessing.



Figure 6. Average accuracy of five-fold cross validation under different numbers of GPC.

After the bagged ensemble of GPCs with the optimal hyperparameter value was fitted on the training set, it was used to evaluate the rockburst damage on the test set. The evaluation results were expressed by the confusion matrix defined by Equation (13), which were indicated as:

	17	4	0	1	
c _	2	6	1	0	
5 —	2	0	6	4	
	0	1	3	2	

Based on Equation (14), the value of accuracy was 63.27%. If the levels L_2 and L_3 were merged into the low-risk group, and the levels L_4 and L_5 were merged into the high-risk group, then the accuracy value of low and high risk was 93.55% and 83.33%, respectively.

According to Equations (15)–(17), the values of precision, recall and F_1 corresponding to different levels were obtained, as shown in Figure 7. It can be seen that the evaluation performance for level L_2 was the best, while that for level L_5 was the worst. After comprehensively considering the values of precision, recall and F1, the ranking of the evaluation performance for different levels was $L_2 > L_3 > L_4 > L_5$.



Figure 7. Average accuracy of five-fold cross validation under different number of GPC.

5. Case Study

The proposed methodology was applied to evaluate the rockburst damage in the Perseverance nickel mine. The main orebody is hosted in ultramafic rocks, which is mined using the sub-level caving method. The hanging wall is composed of stiff Felsic volcanics and metasediments, which are prone to mining-induced seismicity. The microseismic monitoring system has been established in this mine, making it possible to manage rockburst risk based on microseismic data. Because most of the ramp and infrastructure are located in the hanging wall, it is necessary to evaluate the rockburst damage potential.

Heal [35] recorded twelve rockburst damage cases caused by six microseismic events at the depth of 950 m to 1100 m in this mine. The specific data was shown in Table 1. The Richer magnitude of microseismic events ranged from 1.5 to 2.2, and their real rockburst damage levels were between L_2 and L_4 .

Microseismic Event	<i>I</i> ₁	<i>I</i> ₂	I ₃	I_4	I_5	I ₆	<i>I</i> ₈	Actual Level	Results Using Heal's Method [35]	Evaluation Results in This Study
	57.7	5	12.2	0.5	1.62	14	2700	L ₂	L_5	L_2
#1	57.7	5	12.2	0.5	1.62	22	2700	L_2	L_5	L_2
	47	8	6	0.5	1.62	29	2700	L_2	L_2	L_2
	47	8	10.3	0.5	1.8	10	2700	L ₃	L_4	L_2
#2	47	8	6.6	0.5	1.8	10	2700	L_2	L_2	L_4
	46.9	5	5.9	0.5	1.8	16	2700	L_3	L_3	L_4
#2	47.5	10	4.8	0.5	1.5	10	2700	L ₂	L ₂	L_2
#3	47.5	10	10	1	1.5	10	2700	L_2	L_2	L_2
	39.2	5	5	1	1.8	13	2700	L ₂	L_1	L_2
#4	43.4	8	5	1	1.8	13	2700	L_2	L_1	L_2
#5	58	8	12	0.5	1.6	10	2700	L ₂	L_5	L_2
#6	58.1	8	11	1	2.2	5	2700	L_4	L ₃	L_2

Table 1. Rockburst damage cases at Perseverance nickel mine.

The proposed methodology was used to evaluate the rockburst damage levels for these twelve cases. First, these cases were preprocessed together with the original rockburst dataset. Then, the preprocessed dataset was used as the training set to train the model. Finally, the rockburst damage levels of these twelve cases were obtained using the trained model, as shown in the last column of Table 1.

According to the evaluation results in Table 1, only the rockburst damage levels caused by microseismic events #2 and #6 were not identified. That is, the evaluation results of four cases were inconsistent with the actual situation, and the accuracy is 66.67%. In Heal's method [35], the evaluation results of seven cases did not match the actual situation, and the accuracy is 41.67%. Therefore, the methodology proposed in this study improved the evaluation accuracy of rockburst damage to a certain extent.

6. Discussions

Since the dataset used in this study is the same as that in Heal [35], Zhou et al. [16], and Li et al. [17], the evaluation results using our methodology are compared with theirs. The comparison results are shown in Table 2. Among them, Heal [35] artificially synthesized 277 samples with level L_1 according to the distribution of the existing data. The evaluation criteria of accuracy used in different literatures are dissimilar, which mainly include the following four categories: (1) The evaluation value corresponds to the actual value; (2) the evaluation value corresponds to the actual value after combining L_1 , L_2 and L_3 (or L_2 and L_3) into a group while L_4 and L_5 into another group; and (4) the evaluation value corresponds to these four

evaluation criteria, the accuracy of the proposed method is 63.27%, 91.84%, 89.80% and 75.51%, respectively. From Table 2, it can be seen that the accuracy of the proposed method is higher than that of other methods under these four evaluation criteria. This verifies the effectiveness of the method proposed in this study to a certain extent.

Evaluation Criterion	Level	Method	Accuracy
	L_1, L_2, L_3, L_4, L_5	EVP [35]	28.0%
	L_1, L_2, L_3, L_4, L_5	EVP.PPV [35]	24.4%
The evaluation value corresponds to the actual value	L_2, L_3, L_4, L_5	Stochastic gradient boosting approach [16]	61.22%
	L_2, L_3, L_4, L_5	The proposed method	63.27%
The avaluation value componends to the actual value	L_1, L_2, L_3, L_4, L_5	EVP [35]	66.1%
The evaluation value corresponds to the actual value	L_1, L_2, L_3, L_4, L_5	EVP.PPV [35]	72.4%
or the heighboring value	L_2, L_3, L_4, L_5	The proposed method	91.84%
The evaluation value corresponds to the actual value	L_1, L_2, L_3, L_4, L_5	EVP [35]	71.3%
after combining L_1 , L_2 and L_3 (or L_2 and L_3) into	L_1, L_2, L_3, L_4, L_5	EVP.PPV [35]	78.0%
one group while L_4 and L_5 into another group	L_2, L_3, L_4, L_5	The proposed method	89.80%
		Rock engineering systems	
The evaluation value corresponds to the actual value	L_2, L_3, L_4, L_5	and artificial neural	71%
after combining L_2 and L_3 into one group		network [17]	
	L_2, L_3, L_4, L_5	The proposed method	75.51%

Table 2. Comparison of evaluation results in different literatures.

Moreover, to further illustrate the reliability of the proposed method, it is also compared with the bagged ensemble of GPCs without preprocessing, bagged ensemble of GPCs without under-sampling, and GPC without under-sampling. The evaluation results of different approaches are shown in Table 3. It can be seen that the bagged ensemble of GPCs without data preprocessing has the lowest accuracy of 48.98%, which shows the importance of data preprocessing. Before preprocessing, the distribution of some indicators is skewed, and there is no clear distribution law for each indicator. When using the Yeo-Johnson transformation, the sample data is mapped from an arbitrary distribution to a Gaussian distribution as close as possible to stabilize variance and minimize skewness. In addition, the influence of diverse dimensions and units on the evaluation results can be avoided after standardization. Therefore, the accuracy is improved by using data preprocessing. The bagged ensemble of GPCs and the GPC without under-sampling can identify level L_2 well, but cannot achieve a reliable recognition of level L_5 . Moreover, the value of F_1 is 0 in these methods, which illustrates that under-sampling has an important influence on the evaluation results. Because the distribution of different rockburst damage levels is relatively unbalanced, the prediction results are biased towards the level with a larger number of samples. When using the under-sampling technique, this influence can be avoided to some extent by balancing the training samples. Compared with GPC, the bagged ensemble of GPCs improves the evaluation accuracy, which indicates the bagging method can improve the evaluation performance. By integrating multiple GPCs using the bagging method, the generalization ability and robustness of the model can be increased to a certain extent. Therefore, the integration of data preprocessing, under-sampling technique, GPC, and bagging method improves the comprehensive performance.

Approaches	Confusion Matrix	Accuracy	<i>F</i> ₁
Bagged ensemble of GPCs without data preprocessing	$\begin{bmatrix} 15 & 4 & 2 & 1 \\ 3 & 4 & 1 & 1 \\ 5 & 1 & 4 & 2 \\ 0 & 1 & 4 & 1 \end{bmatrix}$	48.98%	[0.6667, 0.4211, 0.3478, 0.1818]
Bagged ensemble of GPCs without under-sampling	$\begin{bmatrix} 20 & 1 & 0 & 1 \\ 5 & 4 & 0 & 0 \\ 4 & 0 & 6 & 2 \\ 3 & 1 & 2 & 0 \end{bmatrix}$	61.22%	[0.7407, 0.5333, 0.6, 0]
GPC without under-sampling	$\begin{bmatrix} 19 & 2 & 0 & 1 \\ 5 & 3 & 1 & 0 \\ 4 & 0 & 6 & 2 \\ 3 & 1 & 2 & 0 \end{bmatrix}$	57.14%	[0.7170, 0.4, 0.5714, 0]
The proposed method	$\begin{bmatrix} 17 & 4 & 0 & 1 \\ 2 & 6 & 1 & 0 \\ 2 & 0 & 6 & 4 \\ 0 & 1 & 3 & 2 \end{bmatrix}$	63.27%	[0.7907, 0.6, 0.5455, 0.3077]

Table 3. Evaluation results of different approaches.

Although the proposed method can evaluate the rockburst damage to some extent, there are still some shortcomings:

- (1) The bagged ensemble of GPCs has better evaluation performance for level L_2 , but the evaluation performance for level L_5 still needs to be improved. The reason may be that the sample size of L_2 is the largest, while that of L_5 is the least. A large number of samples can make the model fit better, which can improve the evaluation performance in turn. Because the data-driven method is highly dependent on the quality of data, a higher-quality rockburst damage database should be established in the future.
- (2) More indicators for rockburst damage evaluation need to be considered. According to the original rockburst damage database, some samples with the same indicator values have different levels. This shows that some key indicators are ignored, which may be an important reason for restricting the evaluation accuracy of rockburst damage. In the future, some novel evaluation indicators may be proposed from the perspective of focal mechanisms and failure characteristics of rock mass under dynamic and static stress.
- (3) Considering the distribution of some indicators is skewed, the Gaussian process may yield impropriate results. Although the proposed method can obtain relatively good results, the Gaussian process with skewed errors can be further used to investigate the evaluation performance [39,40].

7. Conclusions

To effectively assess rockburst damage potential with an imbalanced dataset, this study proposed a novel model by integrating the under-sampling technique, GPC, and bagging method. Based on the rockburst dataset preprocessed by the Yeo-Johnson transformation and standardization, the reliability of the proposed model was verified. The accuracy values of all samples, low risk (L_2 and L_3) and high risk (L_4 and L_5) were 63.27%, 93.55%, and 83.33%, respectively. According to the values of precision, recall and F_1 , the ranking of evaluation performance at different levels was $L_2 > L_3 > L_4 > L_5$. By using the evaluation criteria in other literature, the accuracy of the proposed method was highest, which further verified the reliability of our method. Based on the evaluation results of the other three methods (the bagged ensemble of GPCs without under-sampling and GPC without under-sampling), the comprehensive performance of the proposed method was better. It indicated the effectiveness of the integration of data preprocessing, under-sampling technique, GPC and bagging method in this study. The proposed methodology was applied to assess rockburst damage

potential in the Perseverance nickel mine. The evaluation accuracy was 66.67%, which was 25% higher than the method in original literature. Due to the improved performance, the evaluation results provided a valuable guidance for the prevention of rockburst disasters.

In the future, a higher-quality database for the evaluation of rockburst damage potential should be established under a unified data acquisition standard. Because of the complex mechanisms of rockburst, some novel evaluation indicators are worth investigating based on focal mechanisms and failure characteristics of rock mass under dynamic and static stress. The Gaussian process with skewed errors can be obtained to investigate the evaluation performance after considering the skewness of indicator distributions. In addition, the proposed methodology can be applied in other mining and geotechnical engineering fields, such as pillar stability prediction and landslide risk analysis.

Author Contributions: This research was jointly performed by Y.C., Q.D., W.L., P.X., B.D. and G.Z. Conceptualization, Y.C. and W.L.; Methodology, W.L.; Validation, Y.C. and W.L.; Formal analysis, Q.D.; Investigation, P.X.; Resources, B.D.; Data curation, W.L. and G.Z.; Writing—original draft preparation, Q.D.; Writing—review and editing, Y.C. and W.L.; Funding acquisition, Y.C. and W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (52004130), the Provincial Natural Science Foundation of Hunan (2022JJ40601, 2022JJ40373), and the China Postdoctoral Science Foundation (2021M693799).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Appendix A

Table A1. Rockburst Damage Dataset.

Number	I_1	<i>I</i> ₂	<i>I</i> ₃ /(m)	I_4	I_5	<i>I</i> ₆ /(m)	<i>I</i> ₈ /(kg/m ³) Level
1	80	5	6.2	1	-0.3	5	2700	4
2	60	5	4.2	0.5	1.7	20	2700	4
3	60	8	4.2	0.5	1.7	25	2700	2
4	80	8	6	0.5	1.8	10	2700	4
5	70	8	4	1	1.8	15	2700	2
6	40	5	3.8	1	0.4	5	2700	2
7	80	8	5.9	1	0.6	5	2700	2
8	90	8	6.8	1	0	5	2700	4
9	80	8	7	1	0	10	2700	2
10	80	8	7	1	2	5	2700	4
11	80	8	4.1	1	2	10	2700	4
12	70	8	9.5	1	2.2	5	2700	4
13	75	8	3.8	1	2.2	10	2700	3
14	75	8	4	1	2.2	10	2700	4
15	60	8	6.2	1	1.6	5	2700	2
16	60	10	10.5	0.5	1.6	5	2700	2
17	65	8	4.3	1	0.3	5	2700	2
18	60	5	5.6	0.5	1.5	5	2700	4
19	45	10	9.1	0.5	1.8	5	2700	5
20	43	5	9.3	0.5	1.8	5	2700	4
21	43	10	9.3	1	1.8	5	2700	2
22	43	10	9.4	0.5	-0.2	5	2700	2
23	54	8	3.5	0.5	1.3	5	2700	2

Table A1. Cont.

Number	I_1	<i>I</i> ₂	<i>I</i> ₃ /(m)	I_4	I_5	<i>I</i> ₆ /(m)	$I_8/(\text{kg/m}^3)$	Level
24	45	8	3.6	1	1.3	10	2700	2
25	80	8	5.4	1	1.3	10	2700	2
26	50	8	7.8	1	1.3	15	2700	2
27	50	5	6.2	1	1	5	2700	3
28	50	5	5.1	0.5	1.2	5	2700	4
29	50	5	5.1	1	1.2	5	2700	2
30	50	8	8.3	1	0.7	5	2700	2
31	50	8	5.5	1	0.7	5	2700	2
32	60	10	8.8	1	2	5	2700	2
33	60	5	6.2	1	2	5	2700	2
34	60	5	5.2	1	2	5	2700	2
35	60	5	8.4	1	2	5	2700	2
36	60	5	6	1	2	5	2700	2
37	60	10	8.4	1	2	5	2700	2
38	60	5	5.3	1	2	5	2700	2
39	60	10	7	1	2	5	2700	2
40	60 75	10	5.4	0.5	2	5	2700	4
41	75 70	5 F	5.1	1	0.6	15	2700	3
42	70	5 F	5.1	1	0.6	10	2700	2
43	70	5 F	5.1	1	0.6	5 10	2700	3
44	75 75	5 E	5.2	1	1.3	10	2700	2
43	75	5	0.7 5 1	1	1.5	15 20	2700	2
40	75	8	5.1	1	1.3	20	2700	2
47	75	8	5	1	1.0	10	2700	2
40	75	0 5	53	1	1.0	5	2700	2
50	35	2	53	1	3.1	15	2700	5
51	35	2	10.6	1	3.1	15	2700	5
52	35	10	10.0	1	3.1	20	2700	2
53	35	5	59	15	3.1	25	2700	4
54	35	2	10.6	1.5	3.1	25	2700	4
55	35	5	11.8	1.5	3.1	20	2700	2
56	35	5	10.6	0.5	3.1	25	2700	3
57	35	8	7.6	0.5	3.1	15	2700	4
58	41.7	5	7.6	1.5	3.1	15	2700	2
59	41.7	2	7	1	3.1	10	2700	4
60	35	10	10.6	0.5	3.1	15	2700	2
61	35	5	7	1	3.1	15	2700	4
62	40	2	14	1	2.8	15	2900	5
63	39	2	7	1	2.8	10	2900	4
64	39	2	6.5	1	2.8	50	2900	4
65	40	2	5.7	1	1.6	25	2900	3
66	42.86	2	6.4	1.5	1.6	20	2900	3
67	35.1	2	6.8	1.5	3.5	50	2900	2
68	38	10	9	0.5	3.5	10	2900	5
69	32.3	2	4.5	1.5	3.5	35	2900	3
70	43.7	8	7.7	0.5	3.5	20	2900	3
71	43.7	10	7	0.5	3.5	20	2900	3
72	43.7	2	4.2	1	3.5	15	2900	4
73	42.8	2	5	0.5	3.5	30	2900	5
74	42.8	10	10	0.5	3.5	30	2900	3
75	39.5	5	5	0.5	3.5	10	2900	4
76	44.4	10	15	1	3.5	15	2900	4
77	47.3	8	8.3	1	3.5	50	2900	3
78	47.3 51.40	8	5.5		3.5	50	2900 2000	3
79	51.43 20.4	5	9.3	1.5	1.9	10	2900 2000	3
8U Q1	39.4 20.4	5	9.5 4 E	1.5	2.1	15	∠900 2000	2
01	39.4	7	4.3	1	2.1	20	2900	3

Table A1. Cont.

Number	I_1	I_2	<i>I</i> ₃ /(m)	I_4	I_5	<i>I</i> ₆ /(m)	$I_8/(\text{kg/m}^3)$	Level
82	39.4	10	11	1	2.1	25	2900	2
83	41.7	2	4.5	0.5	2.1	60	2900	3
84	40.8	10	10	0.5	2.1	30	2900	3
85	40.8	5	8	0.5	2.1	25	2900	3
86	44.1	5	18	1	2.1	50	2900	2
87	36	8	14	1	1.2	10	2900	2
88	36	8	6.5	0.5	1.2	15	2900	2
89	37.7	8	5.2	1	0.4	5	2900	3
90	40.8	8	4.2	1	1.8	5	2900	2
91	30	10	5.3	0.5	1.2	5	2800	3
92	30	8	5.4	0.5	1.2	5	2800	2
93	30	10	5.3	1	1.2	10	2800	2
94	30	10	5.2	1	1.2	10	2800	3
95	74	8	5.9	0.5	1.5	5	2800	5
96	74	8	5.3	1	1.5	10	2800	3
97	74	8	5.9	1	1.5	10	2800	2
98	74	8	6.5	1	1.5	20	2800	2
99	74	8	6.5	0.5	1.5	15	2800	4
100	71	8	8	1	0.9	5	2800	4
101	71	8	5.3	1	0.9	10	2800	2
102	71	10	8	1	0.9	10	2800	2
103	71	8	5.6		0.9	10	2800	2
104	71	10	20	0.5	2.1	5	2800	5
105	40	8	5.9	0.5	2.1	5	2700	4
106	40	8	5.3	0.5	2.1	5	2700	2
107	40	8	5.9	1	2.1	5	2700	2
108	40	8	6.5 E.(1	2.1	10	2700	2
109	40	2	5.6	1	2.1	10	2700	4
110	40	2	5.8 E E	1	2.1	20 20	2700	4
111	40 70	0 5	3.3 8	1	2.1	20	2700	Z 4
112	70	10	8	1	0.0	10	2800	4 2
113	70	10	55	1	0.0	10	2800	2
115	70	5	8	1	0.0	10	2800	2 4
116	70	8	51	1	0.0	10	2800	3
117	70	5	51	1	0.8	5	2800	2
118	70	5	5.1	1	0.8	10	2800	2
119	54	5	5.7	1	0.8	20	2700	2
120	54	10	9.1	0.5	0.8	25	2700	2
121	39	8	5.7	1	0.8	30	2800	2
122	84	5	7.7	1	2.9	10	2900	5
123	45	5	4.8	1	2.9	50	2900	2
124	84	10	7.4	1	2.9	50	2900	3
125	45	10	6.9	1	2.9	10	2900	2
126	45	5	7.4	1	2.9	10	2900	4
127	56	5	4.6	1	2.9	25	2900	2
128	18	5	5.8	0.5	0.4	10	3030	4
129	24	5	8.6	0.5	0.4	10	3030	3
130	95	10	6.9	1	1.5	5	2900	2
131	95	5	6.6	1	0.9	5	2900	5
132	45	10	5.1	1	1.6	5	2900	2
133	21	5	11.2	1.5	0.9	5	3030	2
134	21	5	6.1	1.5	0.9	5	3030	2
135	95	10	8	1	1.6	5	2900	5
136	39	10	5.3	1	1.6	15	2900	3
137	21	5	5.5	1	1.9	20	3030	3
138	24	5	8.7	0.5	1.5	10	3030	5
139	24	5	11	1	1.5	15	3030	2

Number	I_1	I_2	<i>I</i> ₃ /(m)	I_4	I_5	<i>I</i> ₆ /(m)	$I_8/(\text{kg/m}^3)$	Level
140	67	10	5	1	-0.2	5	2900	2
141	21	5	9	0.5	1.8	10	3030	4
142	21	10	9	1	1.8	10	3030	2
143	95	10	6.8	1	1	5	2900	4
144	73	25	6.8	1	1	5	2900	3
145	27	5	11.5	1	3.1	30	3030	4
146	27	5	7.6	1	3.1	40	3030	4
147	35	5	11.5	1	3.1	30	3030	4
148	50	25	4.5	1	3.1	40	2900	2
149	95	25	7.1	1	3.1	20	2900	2
150	73	25	4.7	1	3.1	30	2900	2
151	95	25	6.3	1	3.1	30	2900	2
152	73	25	4.4	1	3.1	40	2900	2
153	73	25	9.6	1	3.1	50	2900	2
154	54	25	4.8	1	3.1	60	2900	2
155	34	5	4.5	1	3.1	70	2900	3
156	25	10	11.6	0.5	1.4	5	3030	4
157	25	5	11.6	0.5	1.4	5	3030	4
158	24	5	12	1	2	5	3030	5
159	39	5	9	1	2	5	2900	4
160	25	5	5.1	1	1.3	5	3080	4
161	25	5	10.5	1	1.3	10	3080	3
162	25	5	7.8	1	1.3	10	3080	2
163	25.97	10	17	0.5	2	5	3080	5
164	25.97	8	6.2	0.5	2	10	3080	3
165	25.97	8	5.7	0.5	2	5	3080	4
166	25.97	5	5.4	1	2	10	3080	3
167	25.97	8	5.3	1	2	10	3080	2
168	25.97	8	5.6	0.5	2	5	3080	4
169	75	5	5.2	1	1.6	10	2800	5
170	75	5	5	1	1.6	5	2800	5
171	65	5	2	1	1.6	5	2800	2
172	50	5	9.1	1	1.4	5	2800	3
173	50	5	4.6	0.5	1.9	30	2800	3
174	70	8	5	0.5	1.6	5	4300	4
175	70	8	8	1	2	5	4300	4
176	70	8	6	0.5	2	5	4300	4
177	70	8	12	1	2	10	4300	3
178	67	10	11	0.5	2.5	15	4300	5
179	67	25	5	1	2.5	15	4300	2
180	67	5	5	1	2.5	15	4300	2
181	76	10	6	0.5	2.7	5	4300	5
182	40	10	9.2	0.5	1.1	5	4300	3
183	40	5	4.8	1	1.1	10	4300	2
184	40	10	11.7	0.5	2.5	5	4300	5
185	50	10	9	1	2.7	20	4300	4
186	50	5	6	1	2.1	5	4300	4
187	55	5	8	1	1.9	5	4300	5
188	55	5	6	1	2.3	5	4300	4
189	65	10	11	1	2.3	10	4300	4
190	55	5	6	0.5	0.9	5	4300	4
191	60	5	5.5	1	2.2	20	4300	4
192	50	5	5.5	0.5	1.4	5	4300	5
193	50	10	8.5	0.5	1.4	20	4300	2
194	50	10	5.5	0.5	1.4	40	4300	2
195	50	5	6	1	1.5	5	4300	3
196	50	10	30	1	1.7	5	4300	5
197	70	5	4.4	1.5	1.7	20	2700	2

Number	I_1	I_2	<i>I</i> ₃ /(m)	I_4	I_5	<i>I</i> ₆ /(m)	$I_8/(\text{kg/m}^3)$	Level
198	70	10	4.6	0.5	2	5	2800	4
199	90	10	4.5	1.5	2	10	2850	2
200	70	5	5.2	1	2.1	5	2700	5
201	56.2	8	10	1	1	5	2870	4
202	56.2	10	10	1	1	10	2870	2
203	56.2	8	6	1	1	20	2870	2
204	57.8	8	6.1	1	1	5	2870	3
205	57.8	8	6.1	1.5	1	10	2870	3
206	57.8	8	6.5	1	1.5	5	2870	4
207	57.8	10	11.3	1	1.5	10	2870	2
208	57.8	10	6.5	1	1.5	10	2870	2
209	57	8	6.7	1	2.2	5	2870	4
210	57	10	9.5	1	2.2	5	2870	4
211	57	10	11.2	1	2.2	10	2870	2
212	57	8	6.4	1	2.2	25	2870	2
213	57	8	6.5	1	2.2	10	2870	2
214	57	10	11.5	0.5	1.7	5	2870	4
215	57	10	11	1	1.7	5	2870	2
216	57	10	11	1	1.7	10	2870	2
217	57	10	11.5	1	1.7	10	2870	2
218	57	10	7.4	1	1.7	15	2870	2
219	57.8	10	6.4	0.5	2.5	5	2870	5
220	57.8	10	11.2	0.5	2.5	5	2870	5
221	57.8	10	6.4	1	2.5	10	2870	2
222	57.8	10	10.6	1	2.5	10	2870	2
223	58.6	10	12.4	0.5	2.2	30	2870	2
224	58.6	10	5.9	1	2.2	30	2870	2
225	58.6	10	6.1	1	2.2	30	2870	2
226	59.3	8	8	1	2.2	5	2870	5
227	59.3	8	5.4	1	2.2	15	2870	3
228	59.3	8	10	1	2.2	10	2870	2
229	59.3	8	8	1	2.2	15	2870	3
230	59.3	8	8.4	1	2.2	15	2870	2
231	59.3	8	5	1	2.2	20	2870	2
232	70.3	10	6.9	0.5	2.3	5	2900	5
233	70.3	10	11	1	2.3	10	2900	2
234	70.3	10	5.5	1	2.3	15	2900	3
235	70.3	10	5.4	1	2.3	15	2900	2
236	72.2	8	4	1	1.6	5	2900	3

Table A1. Cont.

References

- Keneti, A.; Sainsbury, B.A. Review of published rockburst events and their contributing factors. *Eng. Geol.* 2018, 246, 361–373. [CrossRef]
- 2. Gong, F.Q.; Yan, J.Y.; Li, X.B. A new criterion of rock burst proneness based on the linear energy storage law and the residual elastic energy index. *Chin. J. Rock Mech. Eng.* **2018**, *37*, 1993–2014.
- Hudyma, M.; Potvin, Y.H. An engineering approach to seismic risk management in hardrock mines. *Rock Mech. Rock Eng.* 2010, 43, 891–906. [CrossRef]
- 4. Sepehri, M.; Apel, D.B.; Adeeb, S.; Leveille, P.; Hall, R.A. Evaluation of mining-induced energy and rockburst prediction at a diamond mine in Canada using a full 3D elastoplastic finite element model. *Eng. Geol.* **2020**, *266*, 105457. [CrossRef]
- 5. Gong, F.Q.; Yan, J.Y.; Li, X.B.; Luo, S. A peak-strength strain energy storage index for rock burst proneness of rock materials. *Int. J. Rock Mech. Min. Sci.* 2019, 117, 76–89. [CrossRef]
- 6. Ortlepp, W.D.; Stacey, T.R. Rockburst mechanisms in tunnels and shafts. Tunn. Undergr. Sp. Technol. 1994, 9, 59–65. [CrossRef]
- Hedley, D.G.F. A Five-Year Review of the Canada–Ontario Industry Rockburst Project; Special Report SP90, Division Report SP90-064; Canada Centre for Mineral and Energy Technology, Mining Research Laboratory: Ottawa, ON, Canada, 1990.
- 8. Lu, C.P.; Liu, G.J.; Liu, Y.; Zhang, N.; Xue, J.H.; Zhang, L. Microseismic multi-parameter characteristics of rockburst hazard induced by hard roof fall and high stress concentration. *Int. J. Rock Mech. Min.* **2015**, *76*, 18–32. [CrossRef]

- Durrheim, R.J. Mitigating the Risk of Rockbursts in the Deep Hard Rock Mines of South Africa: 100 Years of Research. Extracting the Science: A Century of Mining Research; Brune, J., Ed.; Society for Mining, Metallurgy, and Exploration, Littleton: New York, NY, USA, 2010; pp. 156–171.
- Kaiser, P.K.; McCreath, D.R.; Tannant, D.D. Canadian Rockburst Research Program 1990–1995; Mining Division of the Canadian Mining Industry Research Organization: Sudbury, ON, Canada, 1997.
- 11. Kaiser, P.K.; Tannant, D.D.; McCreath, D.R.; Jesenak, P. Rockburst Damage Assessment Procedure, Rock Support in Mining and Underground Construction; Balkema, K.M., Ed.; CRC Press: Rotterdam, The Netherlands, 1992; pp. 639–647.
- Durrheim, R.J.; Roberts, M.K.C.; Haile, A.T.; Hagan, T.O.; Jager, J.A.; Handley, M.F.; Spottiswoode, S.M.; Ortlepp, W.D. Factors influencing the severity of rockburst damage in South African gold mines. In SARES 97—1st Southern African Rock Engineering Symposium; Gurtunca, R.G., Hagan, T.O., Eds.; SARES: Johannesburg, South Africa, 1997; pp. 17–24.
- 13. Brink, A.; Hagan, T.O.; Spottiswoode, S.M.; Malan, D.F.; Glazer, S.N.; Lasocki, S. *Survey and Assessment of Techniques Used to Quantify the Potential for Rock Mass Instability*; Safety in Mines Research Advisory Committee: Pretoria, South Africa, 2000.
- Albrecht, J.; Sharrock, G. A Model to Forecast Rockburst Damage. Challenges in Deep and High Stress Mining; Australian Centre for Geomechanics: Perth, Australia, 2006; pp. 1–15.
- Heal, D.; Hudyma, M.; Potvin, Y. Evaluating rockburst damage potential in underground mining. In Proceedings of the 41st US Symposium on Rock Mechanics (USRMS), Golden, CO, USA, 17–21 June 2006; pp. 1020–1025.
- Zhou, J.; Shi, X.Z.; Huang, R.D.; Qiu, X.Y.; Chen, C. Feasibility of stochastic gradient boosting approach for predicting rockburst damage in burst-prone mines. *Trans. Nonferr. Met. Soc.* 2016, 26, 1938–1945. [CrossRef]
- 17. Li, N.; Zare Naghadehi, M.; Jimenez, R. Evaluating short-term rock burst damage in underground mines using a systems approach. *Int. J. Min. Reclam. Environ.* 2020, 34, 531–561. [CrossRef]
- 18. Pu, Y.Y.; Apel, D.B.; Liu, V.; Mitri, H. Machine learning methods for rockburst prediction-state-of-the-art review. *Int. J. Min. Sci. Technol.* **2019**, *29*, 565–570. [CrossRef]
- 19. Liang, W.Z.; Sari, Y.A.; Zhao, G.Y.; McKinnon, S.D.; Wu, H. Probability estimates of short-term rockburst risk with ensemble classifiers. *Rock Mech. Rock Eng.* **2021**, *54*, 1799–1814. [CrossRef]
- Yin, X.; Liu, Q.S.; Pan, Y.C.; Huang, X.; Wu, J.; Wang, X.Y. Strength of stacking technique of ensemble learning in rockburst prediction with imbalanced data: Comparison of eight single and ensemble models. *Nat. Resour. Res.* 2021, 30, 1795–1815. [CrossRef]
- 21. Jiang, K.; Lu, J.; Xia, K.L. A novel algorithm for imbalance data classification based on genetic algorithm improved SMOTE. *Arab. J. Sci. Eng.* **2016**, *41*, 3255–3266. [CrossRef]
- 22. Xue, Y.G.; Li, G.K.; Li, Z.Q.; Wang, P.; Gong, H.M.; Kong, F.M. Intelligent prediction of rockburst based on Copula-MC oversampling architecture. *Bull. Eng. Geol. Environ.* 2022, *81*, 209. [CrossRef]
- 23. López, V.; Fernández, A.; Herrera, F. On the importance of the validation technique for classification with imbalanced datasets: Addressing covariate shift when data is skewed. *Inf. Sci.* **2014**, 257, 1–13. [CrossRef]
- Chawla, N.V.; Japkowicz, N.; Kotcz, A. Special issue on learning from imbalanced data sets. SIGKDD Explor. 2004, 6, 1–6. [CrossRef]
- Zhang, Z.; Krawczyk, B.; Garcia, S.; Rosales-Pérez, A.; Herrera, F. Empowering one-vs-one decomposition with ensemble learning for multi-class imbalanced data. *Knowl. -Based Syst.* 2016, 106, 251–263. [CrossRef]
- 26. Sun, B.; Chen, H.Y.; Wang, J.D.; Xie, H. Evolutionary under-sampling based bagging ensemble method for imbalanced data classification. *Front. Comput. Sci.* 2018, 12, 331–350. [CrossRef]
- 27. Todorov, V.; Dimov, I. Innovative digital stochastic methods for multidimensional sensitivity analysis in air pollution modelling. *Mathematics* **2022**, *10*, 2146. [CrossRef]
- Sun, J.J.; Yeh, T.M.; Pai, F.Y. Application of Monte Carlo simulation to study the probability of confidence level under the PFMEA's action priority. *Mathematics* 2022, 10, 2596. [CrossRef]
- 29. Heilmeier, A.; Graf, M.; Betz, J.; Lienkamp, M. Application of Monte Carlo methods to consider probabilistic effects in a race simulation for circuit motorsport. *Appl. Sci.* **2020**, *10*, 4229. [CrossRef]
- 30. Tama, B.A.; Lim, S. A comparative performance evaluation of classification algorithms for clinical decision support systems. *Mathematics* **2020**, *8*, 1814. [CrossRef]
- Rinta-Koski, O.P.; Särkkä, S.; Hollmén, J.; Leskinen, M.; Andersson, S. Gaussian process classification for prediction of in-hospital mortality among preterm infants. *Neurocomputing* 2018, 298, 134–141. [CrossRef]
- 32. Liang, W.Z.; Sari, A.; Zhao, G.Y.; McKinnon, S.D.; Wu, H. Short-term rockburst risk prediction using ensemble learning methods. *Nat. Hazards* **2020**, *104*, 1923–1946. [CrossRef]
- 33. Sagi, O.; Rokach, L. Ensemble learning: A survey. Wires Data Min. Knowl. 2018, 8, e1249. [CrossRef]
- Zhang, J.F.; Wang, Y.H.; Sun, Y.T.; Li, G.C. Strength of ensemble learning in multiclass classification of rockburst intensity. *Int. J. Numer. Anal. Methods Geomech.* 2020, 44, 1833–1853. [CrossRef]
- Heal, D. Observations and Analysis of Incidences of Rockburst Damage in Underground Mines. Ph.D. Thesis, University of Western Australia, Perth, Australia, 2010.
- 36. Williams, C.K.; Rasmussen, C.E. Gaussian Processes for Machine Learning; MIT Press: Cambridge, UK, 2006.
- Santhanam, V.; Morariu, V.I.; Harwood, D.; Davis, L.S. A non-parametric approach to extending generic binary classifiers for multi-classification. *Pattern Recogn.* 2016, 58, 149–158. [CrossRef]

- 38. Yeo, I.K.; Johnson, R.A. A new family of power transformations to improve normality or symmetry. *Biometrika* **2000**, *87*, 954–959. [CrossRef]
- 39. Alodat, M.T.; Shakhatreh, M.K. Gaussian process regression with skewed errors. J. Comput. Appl. Math. 2020, 370, 112665. [CrossRef]
- 40. Benavoli, A.; Azzimonti, D.; Piga, D. A unified framework for closed-form nonparametric regression, classification, preference and mixed problems with Skew Gaussian Processes. *Mach. Learn.* **2021**, *110*, 3095–3133. [CrossRef]