Application of KNN-based Isometric Mapping and Fuzzy C-Means Algorithm to Predict Short-term Rockburst Risk in Deep Underground Projects

Muhammad Kamran (✉ muhdkamran62@gmail.com)  
Department of Mining Engineering, Institute Technology of Bandung, Bandung 40132, Indonesia  
https://orcid.org/0000-0002-2943-9098

Barkat Ullah  
Central South University

Mahmood Ahmad  
University of Engineering and Technology Peshawar

Mohanad Muayad Sabri Sabri  
Peter the Great St. Petersburg Polytechnic University, 195251 St. Petersburg, Russia

Research Article

Keywords: Rockburst, Safety, KNN, ISOMAP algorithm, FCM

Posted Date: October 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2128698/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

The rockburst phenomenon is the major source of the high number of casualties and fatalities during the construction of deep underground projects. Rockburst poses a severe hazard to the safety of employees and equipment in subsurface mining operations. It is a hot topic in recent years to examine and overcome rockburst risks for the safe installation of deep urban engineering designs. Therefore, for a cost-effective and safe underground environment, it is crucial to determine and predict rockburst intensity prior to its occurrence. A novel model is presented in this study that combines unsupervised and supervised machine learning approaches in order to predict rockburst risk. The database for this study was built using authentic microseismic monitoring occurrences from the Jinping-II hydropower project in China, which consists of 93 short-term rockburst occurrences with six influential features. The prediction process was succeeded in three steps. Firstly, the original rockburst database's magnification was reduced using a state-of-the-art method called isometric mapping (ISOMAP) algorithm. Secondly, the dataset acquired from ISOMAP was categorized using the fuzzy c-means algorithm (FCM) to reduce the minor spectral heterogeneity impact in homogenous areas. Thirdly, K-Nearest neighbour (KNN) was employed to anticipate different levels of short-term rockburst datasets. The KNN's classification performance was examined using several performance metrics. The proposed model correctly classified about 96% of the rockbursts events in the testing datasets. Hence, the suggested model is a realistic and effective tool for evaluating rockburst intensity. Therefore, the proposed model can be employed to forecast the rockburst risk in the early stages of underground projects that will help to minimize casualties from rockburst.

1. Introduction

Rockburst is a dynamic phenomenon which occurs in underground excavations when immense amounts of energy are released, rocks are inelastically deformed, and rocks are thrown into the excavations [1]. As defined by the Mine Safety and Health Administration (MSHA), "a rockburst occurs when overstressed rock collapses abruptly, releasing large amounts of energy instantly" [2]. Rockburst occurrence is mainly associated with the geological structure, properties of surrounding rock masses and lithology. It has been demonstrated that rockburst poses a severe hazard to the safety of employees and equipment in underground constructions [3–5, 10]. Therefore, for a cost-effective and safe deep underground construction or mining in burst-prone conditions, it is crucial to determine and predict rockburst intensity prior to its occurrence.

The rockburst intensity is characterized into four different levels [11]. It is a big challenge to predict rockburst due to its complex and nonlinear nature. In the last few decades, several methods have been utilized to assess rockbursts [10]. Monitoring and forecasting the rockburst danger are carried out by utilizing microgravity, microseismic, and geological radar techniques [12]. Based on on-site monitoring of microseismic waves emitted during rock fractures, some precursory features of rockbursts were discovered that could be used to predict the risk of rockbursts. The most commonly used microseismic features to predict rockbursts are the number of events [13], energy features [14], apparent volume [15]
and $b$ value, which is the slope of the commutative hit toward amplitude [16]. In addition to these indexes, researchers have proposed other strategies for predicting the long-term occurrence of rockburst. Rockburst risk can be assessed using tangential stress criterion [20], rock brittleness coefficient [17], strain energy storage index [18], and elastic strain energy density [19]. A burst potential index based on energy has been established to assess burst proneness [21]. The increasing demand for energy and construction has resulted in underground excavations being extended to greater lengths, which has caused severe rockburst disasters. It has therefore been a hot topic in recent years to examine and overcome rockburst risks.

In the previous few decades, rockburst prediction or evaluation approaches have evolved, but there has never been a breakthrough or generally recognised method that has been preferred over others. Since then, the rockburst has been an unsolved and alarming issue that needs to be resolved more precisely. In order to eliminate the threat of rockbursts in the first place, advanced rockburst prediction is crucial to reducing the cost of the damage and preventing major losses from a rockburst. In recent years, state-of-art intelligent techniques have widely been implemented to overcome the severe dynamics hazards of rockburst disasters for the safe installation of underground projects. The researchers have extended their horizons and utilized cutting-edge soft-computing methods to predict the rockburst occurrence intensity successfully. Additionally, the intelligent algorithm is low cost, only focuses on input and output parameters, and has broader applicability [6–9].

The development of artificial intelligence makes the intelligent system more suitable for rockburst prediction. Based on distinct geological conditions, different models are proposed, and these models are so specialized that they cannot be used concurrently to many projects. Rockburst is affected by non-linear factors, and artificial intelligence algorithms are outclassed in non-linear analysis with high-dimensional datasets. The most widely used methods to predict rockbursts are support vector machines (SVMs), artificial neural networks (ANNs), K-nearest neighbors (KNNs), classification and regression trees (CARTs), ensemble learning and random forests (RFs) [23]. The results of a study have shown that ANN models can be used to predict rockburst risks in deep gold mines in South Africa after an improved model was introduced [24]. Zhao et al. [25] constructed a data-driven model using a convolutional neural network (CNN) and compared its performance with that of a traditional neural network. An SVM model was used by Zhou et al. [23] to categorize a long-term rockburst. The four classical single intelligent algorithms, namely k-nearest neighbors (KNN), SVM, deep neural networks (DNN) and recurrent neural networks (RNN), were combined to form four ensemble models (KNN–RNN, SVM–RNN, DNN–RNN and KNN–SVM–DNN–RNN) using stacking ensemble learning [27]. A new probability model for tunnel rockburst prediction was proposed based on Copula theory and the least square support vector machine (LSSVM) method optimized via particle swarm optimization (PSO) [29]. Using kernel principal component analysis (KPCA), adaptive-PSO, and SVM, Li et al. [32] developed a hybrid model (KPCA-APSO-SVM). The short-term rockburst risk was predicted using t-distributed stochastic neighbor embedding (t-SNE), K-means clustering, and extreme gradient boosting (XGBoost) algorithms [26]. On the basis of microseismic monitoring data, Zhao et al. [22] developed a model for prediction of rockbursts that uses a decision tree (DT) model. For assessing the rockburst hazard of an active hard coal mine, Wojtecki et al. [1] used neural networks, decision trees, RF, gradient boosting, and XGBoost. In order to study the
predictability of short-term rockburst, Liang et al. [33] used microseismic data from Jinping-II hydropower project to examine the predictability of short-term rockburst. This study evaluated several ensembles learning algorithms, including RF, adaptive boosting, gradient boosting decision tree, XGBoost, and light gradient boosting machine (LightGBM). RF and GBDT have shown good performance. Li et al. [31] demonstrated the predictability of different ensemble trees in estimating rockburst based on 314 real rockbursts. Sun et al. [30] proposed an RF and firefly algorithm (FA) based ensemble classifier to achieve an optimum rockburst prediction model. A study by Ahmad et al. [28] investigated that J48 and random tree algorithms can successfully predict the rockburst classification ranks based on 165 rockburst cases.

In deep underground projects, a self-organizing map and fuzzy c-mean clustering techniques were used to cluster rockbursts events [34]. Even though several rockburst estimation models have been described and compared by previous researchers [35–41], developing an accurate and reliable predictive model still poses a significant challenge for the ground, which is likely to experience frequent rock bursts. Further, many other models for forecasting rockbursts can be considered valuable and efficient tools for geological and mining engineering applications. Unsupervised machine learning is an approach that researchers have shown increasing interest in adopting, in which data labels are not required to be known in advance to perform the analysis. In order to classify the data, clusters are formed based on their proximity to each other and the distance from each cluster's center is taken into account [22, 42–43]. So far, the available studies have succeeded in predicting and classifying the rockburst dynamic disaster but were never entirely successful. A particular procedure can be appropriate in certain instances but not in others. Therefore, the adaptation of state-of-art data depletion combined with unsupervised and KNN learning has less contribution to rockburst intensity prediction. Table 1 illustrates the summary of previously published literature to predict rockburst
Table 1
Summary of previously published literature to predict rockburst

| Year | Number of Input features | Number of datasets | Approaches | Accuracy (%) | References |
|------|--------------------------|--------------------|------------|--------------|------------|
| 2020 | 6                        | 93                 | XGboost    | 73.33        | [44]       |
| 2020 | 6                        | 93                 | Gradient Boost Decision Tree | 76.67 | [44] |
| 2020 | 6                        | 93                 | Adaboost   | 66.67        | [44]       |
| 2020 | 6                        | 93                 | Random Forest | 80       | [44]       |
| 2021 | 4                        | 165                | J48        | 92.857       | [45]       |
| 2018 | 3                        | 108 and 132        | Decision tree model | 73–93 | [46] |
| 2008 | 4                        | 36                 | AdaBoost   | 87.8–89.9    | [47]       |
| 2008 | 6                        | 45                 | v-support vector regression | 93.75 | [48] |
| 2012 | 6                        | 132                | Heuristic algorithms and support vector machine | 66.67-90 | [49] |
| 2021 | 6                        | 311                | Scorecard methodology | 75 | [50] |
| 2022 | 4                        | 165                | ANN        | 89.286       | [35]       |
| 2022 | 4                        | 165                | J48        | 92.857       | [35]       |

2. Significance Of The Study

The predicting features of rockburst levels vary throughout a wide range of rock engineering, mining and geotechnical engineering projects. The impacts of each uncertainty level are yet unclear. In fact, many different findings are reported in the diverse domains of rockburst prediction, short-term rockburst is a complex and sophisticated phenomenon that cannot yet be accurately predicted.

This study develops short-term rockburst prediction model based on uncertainty incorporating unsupervised and supervised learning in order to apply the model effectively in addressing the rock engineering problems. The following three steps are provided in this study to forecast the short-term rockburst:

1. To begin with, the original rockburst database's magnification was reduced using a state-of-the-art method called isometric mapping.
2. The isometric mapping dataset was categorized using fuzzy c-means clustering as an unsupervised machine learning approach to reduce the minor spectral heterogeneity impact in homogenous areas.
3. In order to anticipate different levels of short-term rockburst datasets, KNN, a supervised machine learning has been designed. The study's flowchart is shown in Fig. 1.
3. Material And Methods

3.1. Data Curation

The database for this study was built using authentic microseismic monitoring occurrences from the Jinping-II hydropower project in China. The Jinping II Hydropower Station is located close to the boundaries of Muli, Yanyuan, and Mianning in the Liangshan Yi Autonomous Region of Sichuan, southwest China. It is an ultra-deep buried long tunnel with an extra-large subterranean water power engineering. The average length of the cave line is around 16,67 kilometers, the width of the excavation hole is 13 meters, the underlying rock mass is normally buried between 1500 and 2000 meters, and the maximum buried depth is roughly 2,525 meters. The excavation tunnel section of the 1# and 3# diversion tunnels is a four-hearted horseshoe with an excavation diameter of 13 m, whereas the tunnel boring machine (TBM) excavation section of the 1# and 3# diversion tunnels is a circular section with an excavation diameter of 12.4 m. The distance between the four diversion tunnels is 60 meters. The auxiliary tunnels A and B are 35 meters apart from the centerline of the construction drainage tunnel, whereas the construction drainage tunnel is 45 meters apart from the 4# diversion tunnel [29]. The database consists of 93 short-term rockburst occurrences with six influential features [51]. The dataset analyzed in this study was obtained from the work of Liang et al. [52] and was built on the dataset readily accessible by Feng et al [52]. Rockburst intensity has been divided into four categories: no rockburst level (represented by Level 0) illustrates that the rock composites have no considerable breakage on the free face; slight rockburst level (represented by Level 1) indicates minor composites with modest fragment movement and kinetic energy transfer; moderate rockburst level (represented by Level 2) illustrates the sample debonding of the rock mass inside the diverticulum and highway structure; while severe rockbursts level (represented by Level 3) involve a significant amount of rock mass cracking that immediately fractures the nearby rock mass. The distribution of different rockburst level in this study is shown in Fig. 2. Table 2 depicts the statistic of input and output features employed in the rockburst database.
Table 2
The statistic of input and output feature in rockburst database

| Rockburst Level | Cumulative Number of Events (Z1, Unit) | Event Rate Z2 (Unit/day) | Logarithm of the Cumulative Release Energy Z3 (J) | Logarithm of the Energy Z4 (J/day) | Logarithm of the Apparent Volume Z5 (m3) | Logarithm of the Apparent Volume Z6 (m3/day) |
|----------------|----------------------------------------|--------------------------|-----------------------------------------------|----------------------------------|----------------------------------------|----------------------------------------|
| 3              | 41                                     | 3.727                    | 4.694                                         | 3.653                            | 4.926                                  | 5.968                                  |
| 2              | 14                                     | 1.556                    | 4.622                                         | 3.668                            | 4.887                                  | 5.841                                  |
| 2              | 17                                     | 1.889                    | 4.397                                         | 3.443                            | 3.8                                    | 4.754                                  |
| 2              | 18                                     | 1.8                      | 4.703                                         | 3.703                            | 4.295                                  | 5.295                                  |
| ..................| .................................| .............................| ...............................................| .....................................| .....................................| .....................................|
| 2              | 36                                     | 2.571                    | 4.336                                         | 3.16                             | 2.583                                  | 4.729                                  |
| 1              | 8                                      | 2.667                    | 3.977                                         | 3.5                              | 4.727                                  | 5.204                                  |
| 1              | 16                                     | 2.667                    | 4.681                                         | 3.903                            | 2.843                                  | 3.621                                  |
| 0              | 6                                      | 1.5                      | 2.735                                         | 2.133                            | 4.698                                  | 5.3                                    |
| Minimum        | 1                                      | 0.11                     | 0.78                                          | 0.178                            | 2.511                                  | 1.66                                   |
| Maximum        | 70                                     | 12.25                    | 7.094                                         | 5.89                             | 5.168                                  | 4.39                                   |
| Mean           | 13.011                                 | 1.735                    | 4.389                                         | 3.562                            | 4.15                                   | 3.334                                  |
| Standard deviation | 13.69                      | 1.738                    | 1.441                                         | 1.332                            | 0.66                                   | 0.558                                  |

3.2 Data Visualization

Table 2 reveals that this study incorporates six significant features. The values of Z1, Z2, Z3, Z4, Z5 and Z6 are adjusted on a logarithmic scale to make the effective implementation more favorable. The log parameter's primary objective is to address the database's skewness toward big data. The rockburst database has been investigated by utilizing Python programming language. The Python programming language offers important assistance for experimental data mining, together with data input classification and statistical learning techniques evaluation. Additionally, the learning outcome for a big amount of data is made apparent by visualizing the input data [53]. Figure 3 depicts the voilen of several features for the four rockburst levels. Figure 3 shows a positive correlation between each feature and the corresponding rockburst level. The higher level of rockburst is indicated by the larger values of the features. Additionally, a few outliers may be seen in all of the short-term rockburst dataset's features for accompanying rockburst intensity, demonstrating the heterogeneity of the rockburst events. Therefore,
this study incorporates the effects of all the features to increase the authenticity of the rockburst data structure.

### 3.3 Isometric Mapping (ISOMAP)

Isometric Mapping (ISOMAP) is a stochastic technique for reducing dimensionality that maintains geodesic adjacency using a non-Euclidean measure. As a result, it preserves nonlinear characteristics of the original data that are lost in conventional analysis [54]. Although the ISOMAP represents nonlinear fluctuations in the broader domain, it preserves linearity in small domains [55]. Alternatively, this is referred to as multifaceted learning or consideration. It follows that the manifold's small local area is a conversation of metric space[56]. Figure 4 depicts the mechanism of ISOMAP. The relevant datasets are denoted by the letters $a–f$. They have star points as neighbors. The measured distances of these data are represented by the green segments. It works by mapping the original dataset into a predetermined low-dimensional embedded space and makes the assumption that the high dimensional data is uniformly sampled from a uniform manifold. This method then attempts to identify the underlying manifold [57].

The standard Euclidean geometry and domain-specific metrics can both be used to approximate the geodesic distance. A series of brief hops between nearby points can be added up to approximate domain-specific distance, and the input space's standard Euclidean metric gives a decent approximation of geodesic distance [58]. With the ability to learn a wide range of nonlinear manifolds, ISOMAP combines the key algorithmic characteristics for computational effectiveness, global optimality, and guarantees the asymptotic convergence [59].

### 3.4 Fuzzy C-Means Algorithm (FCM)

The theory of the fuzzy set was developed for the purpose of precisely resolve the difficulties of certainty and uncertainty in the field of optimization and artificial intelligence [60–61]. The FCM was developed based on clustering analysis concept which allows every level to classify in many categories knows as fuzzy sets. This algorithm often has significant advantages over more conventional methods. Several researchers have proposed innovative strategies to FCM in order to solve the problems related to different fields [62–65] Fig. 4 depicts the flowchart of FCM algorithm.

### 3.5 K-Nearest neighbor (KNN)

The K-Nearest neighbor (KNN) algorithm is a supervised technique that is often utilized in a wide range of situations due to its efficiency and simplicity. The most modern development highlights KNN's potential for reducing distortion in a dataset [66–69]. Big data analysis with a KNN classifier demands powerful computing resources. According to the classification approach, a test sample's class label is established using the $k$ nearby samples from the training dataset [70]. The connection between all training instances and the testing data ought to be determined in order to perceive the $k$ nearest neighbours. Each test instance is allocated using KNN contingent on its $k$ closest neighbours. The separation between all
training instances and the test instances ought to be computed in order to find the \( k \) nearest neighbours [69].

Equation 1 is employed in the KNN to determine the spectral intervals (Euclidean distances) across each unknown value and the samples plots. The \( k \) sample plots that are closest to the estimated level are chosen based on the aforementioned particular node.

\[
\gamma = \sqrt{\sum_{k=1}^{m} (u_k - v_k)^2} \quad \text{Eq. 1}
\]

whereas \( \gamma \) stands for the spectral distance between the \( u \) and \( v \) in the n-dimensional space, \( u_k \) and \( v_k \) are the spectral values of unknown variables and \( v \) in the \( kth \) selected spectral variable, respectively. By weighing the reciprocals of their spectral distances from the ideal \( k \) closest plots, the rockburst level was predicted using Eq. (2).

\[
\text{Level}_p = \frac{\sum_{z=1}^{y} \frac{1}{d_{pz}} * v_z}{\sum_{z=1}^{y} \frac{1}{d_{pz}}} \quad \text{Eq. 2}
\]

whereas \( \text{Level}_p \) represents the appropriate forecasted level at the \( p \), \( v_z \) depicts the rockburst observation associated with to the \( zth \) sample, \( d_{pz} \) represent the spectral interval from the \( P \) to the \( zth \) samples, and \( k \) illustrates the ideal number of samples.

Finally, the suggested KNN algorithm with the rockburst data has been computed on Algorithm 3. The massive training data is divided into \( m \) distinct sections first. The size and dispersion of each cluster throughout the various axes of the data space (\( d_u^t \)) should be estimated once the clusters have been identified. The most appropriate cluster of rockburst dataset may be chosen using Algorithm 1. Selecting the appropriate data cluster can substantially affect the outcome of the classification process because the KNN process' output is dependent on the training data. In order to determine the \( k \) nearest neighbours and determine the rockburst level of the test sample, the KNN algorithm is applied to the chosen portion of the training data [71]. A simple 3D multi-class KNN model for the rockburst dataset is shown in Fig. 6.

4. Result And Discussion

Large-scale theoretical and mathematical analyses of datasets are performed using the powerful programming language Python. The clustering and classifying focused data mining techniques and algorithms are supported by the Python programming language. This option is one of the best platforms for designing scalable applications since it has so many advantageous characteristics. Therefore, it may be used to the framework of big data analysis in large rockburst datasets to produce reliable findings.

The ISOMAP technique has been implemented as a method for learning a nonlinear manifold from a collection of unstructured high-dimensional datasets. Its foundation is an expansion of the conventional multidimensional scaling approach to data reduction. The ISOMAP enhances the data identification and make it simple to quantify differences between data points and the reconstructed space. In comparison
to the distances in the other conventional data dimensionality reduction rebuilt spaces, the separation far or near in the ISOMAP-reconstructed matrix give the dimension of the similarity intervening the data point. The ISOMAP effectively separates the points by calculating the geodesic distance. Several ISOMAP iterations can assist in separating the data points, allowing them to be grouped into various clusters. The benefit of ISOMAP is that it employs the Dijkstra's algorithm, which determine the optimum route through neighbours and terminate at every point, to provide realistic distance estimates between the interconnected points distance, which accurately depicts the change in actual space distance between the data points. ISOMAP tool is used to visualize the original rockburst database from high-resolution matrix to low-resolution matrix. The original rockburst dataset having six influential feature was employed in this study. The proposed ISOMAP technique is implemented in Jupyter notebooks. The ISOMAP's three leading factors were utilized to depict the rockburst points that had been reduced after dimensionality reduction. The original rockburst data points in the ISOMAP reconstructed 3D structure showed greater spatial variation, which were evenly spaced apart. Table 3 represent three influential factors rockburst acquired from ISOMAP. The 3D structure of the isometric reconstructed rockburst dataset is shown in Fig. 7.

| Pattern No | Factor 1    | Factor 2    | Factor 3    |
|------------|-------------|-------------|-------------|
| 1          | 29.63107    | 1.098796    | -0.24922    |
| 2          | 1.649354    | -1.04534    | 1.085205    |
| 3          | 4.826978    | -0.45574    | 0.016694    |
| 4          | 5.97017     | -0.5952     | -0.01854    |
| 5          | -2.48045    | -0.55883    | 0.118498    |
| ...        | ...         | ...         | ...         |
| 89         | 17.77873    | -1.29822    | -0.75349    |
| 90         | 24.13729    | 1.433746    | 0.084803    |
| 91         | -4.68471    | -1.88809    | 2.039045    |
| 92         | 3.84182     | 1.570014    | 1.10249     |
| 93         | -7.98256    | -2.45747    | 2.358197    |

The application of FCM could be used to solve a wide range of geostatistical data analysis issues. Each type of numerical data can be employed to generate prototypes and fuzzy partitions with this programmed. These partitions are helpful for supporting established substructures or pointing to substructure in undiscovered data. The FCM clustering with ISOMAP reconstructed points has been employed to determine whether it is feasible to categorize the rockburst level. The FCM clustering is
implemented in Jupyter notebook. Researchers have established the cluster monitoring's generalizability of the findings [22, 72–73]. A metric used to determine the effectiveness of a clustering method is the silhouette coefficient, often known as the silhouette score. Its value is between −1 and 1. The score 1 clusters are clearly distinguishable and spaced far apart, score 0 indicates that clusters are undifferentiated or that there is no statistically significant difference across clusters, whereas score −1 implies that the clusters are assigned incorrectly. Eq. 3 depicts the silhouette score.

\[ \text{Silhouette score} = \frac{(y-x)}{\max(y,x)} \]  

Eq. 3

Whereas \( y \) is the average distance between all clusters and \( x \) is the average intra-cluster distance, or the average distance between each point inside a cluster.

The silhouette score can demonstrate that the ISOMAP acquired data is correctly categorized, representing the arrangement of the features into the categories to which they belong. This index measures the effectiveness of the clustering's authentication in choosing the best \( k \) cluster members. We suppose that there will be four FCM clusters, which corresponds to the four distinct rockburst intensities. This study calculated many iterations phases, as illustrated in Fig. 8. The yellow, aqua, green and red colour were selected to identify the rockburst level 0, 1, 2 and 3 respectively. A suitable model for clustering has been demonstrated in several studies to have a silhouette score of higher than 0.5 [74–76]. Following the tenth iteration in the Isometric mapping derived short-term rockburst dataset, the silhouette score of 0.53 demonstrates that the clusters were consistent and authentic. The center of the circle of the four clusters were (0.740, -0.08, 0.035), (-0.139, -0.543, 0.864), (-0.581, -0.419, -1.136), (-0.776755, 1.755, -0.212). Based on the performance evaluation, the FCM performed well to categorize four different levels of rockburst.

The KNN is a supervised machine learning method that is appropriate for classification and modelling and is reasonably simple to construct. It may be employed in a variety of rock engineering applications. It selects \( k \) sample plots from the training dataset that are most similar to the predicted rockburst level, and it then utilizes the "feature similarity" concept to weight the observations of the \( k \) plots based on how similar their significance is to the unknown rockburst level. The KNN technique for predicting short-term rockburst levels was implemented using the Python programming language. Each unknown rockburst level in this study was given a Euclidean distance to each sample plot, and these distances were used to determine the spectral distances.

The FCM acquired data were randomly divided into training datasets (70%) and testing datasets (30%). The training dataset is used to develop and verify the framework while the testing datasets are utilized in order to assess the framework's capability to estimate the rockburst levels that use previously unobserved data. On the testing dataset, the KNN's forecasting outcomes were obtained. Various performance indices including precision, recall, and F1-score have been employed by the researchers to evaluate the performance of a classification model [77]. In this study, precision, recall, and F1-score have been used to predict the outcomes of the proposed KNN algorithm when associated with ISOMAP and FCM.
The Python programming language was used to generate the classification report for the testing dataset. The classification report provides insight into the effectiveness of the framework on the rockburst events, which is depicted in Table 4. The testing dataset had almost the accuracy of 96% confirming the approximate fitting of the model.

In comparison to level 1, the precision value for level 0, 2 and 3 produced superior results. In terms of precision, the level 0, 1, 2, and 3 have values of 100%, 86%, 100%, and 100%, respectively. The recall values for levels 1, 2, and 3 outperformed level 0 in terms of findings. In comparison to level 0 and 1, the F1-score value for level 2 and 3 attained superior results. The average and weighted recall scores could reach high values of 98% and 96% as a trade-off between precision, recall, and F1-score. The accuracy of the testing dataset was 96%, supporting the proposed model’s approximation of fitting. Hence, the proposed KNN based ISOMAP and FCM algorithm demonstrates favorable classification results for the developed KNN model in successfully identifying the rockburst risk in underground civil structures.

| Precision (%) | Recall (%) | F1-score (%) |
|---------------|------------|--------------|
| Level 0       | 100        | 92           | 96            |
| Level 1       | 86         | 100          | 92            |
| Level 2       | 100        | 100          | 100           |
| Level 3       | 100        | 100          | 100           |
| Accuracy      |             |              | 96            |
| macro avg     | 96         | 98           | 97            |
| weighted avg  | 97         | 96           | 96            |

Figure 9 demonstrates the confusion matrix that’s been generated for the KNN algorithm. The data on the primary diagonal represent the number of samples that the KNN accurately predicted. As can be observed, the KNN successfully classified the majority of rockburst samples. In the whole short-term rockburst dataset, only one rockburst level has been incorrectly predicted. In more detail, one level (0) is incorrectly labelled as level (1). Hence, the KNN algorithm performed well in forecasting the rockburst level in underground civil structures.

**Conclusion**

This study developed ISOMAP + FCM + KNN framework to effectively and accurately anticipate rockburst levels. By examining the results for the proposed model using several performance metrics, the robustness of the generated framework was demonstrated. During this study, three approaches that are frequently used in geotechnical engineering—ISOMAP, FCM, and KNN model—were used to forecast the
rockburst level. More specifically, the data used in this study is collected from several microseismic monitoring occurrences. The statistical performance is used to assess the short-term rockburst level in order to approximation the resilient framework for the best effective model in connection with data prediction. The results of the ISOMAP, FCM, and KNN model demonstrate that it is capable of generating highly precise predictions of the rockburst level.

As a result, it is recommended to employ the ISOMAP + FCM + KNN model developed in this study as a reliable and effective model for predicting the intensity levels of rockbursts. Due to the suggested model's accurate prediction performance in various rock environments, it can be used as a rockburst mitigation and warning system. By keeping several additional geological and rock mechanics data, the model can be made more comprehensive. This model can be utilised as a practical and useful tool for determining rockburst risk. The suggested model can be used to predict the level of rockburst in the early stages of underground mining projects in order to reduce the injuries and fatalities from rockburst.

It is important to take into account the range and volume of trainings because this has an impact on the data-driven models' ability to make logical inferences. The proposed model will be further expanded by developing certain cutting-edge machine learning techniques and contrasting the results of those models with the results of the model attained in this work.

**Declarations**

The authors declare no competing interests.

**Funding**

The research is partially funded by the Ministry of Science and Higher Education of the Russian Federation under the strategic academic leadership program 'Priority 2030' (Agreement 075-15-2021-1333 dated 09/30/2021).

**Author Contributions:**

Conceptualization: Muhammad Kamran, Data Curation: Muhammad Kamran, Writing - original draft: Muhammad Kamran and Barkat Ullah, Validation: Muhammad Kamran, Barkat Ullah., Mahmood Ahmad, Mohanad Muayad Sabri Sabri, Formal analysis: Muhammad Kamran, Barkat Ullah, Mahmood Ahmad, Mohanad Muayad Sabri Sabri, Software: Muhammad Kamran, Visualization: Muhammad Kamran, Supervision: Mahmood Ahmad, and Mohanad Muayad Sabri Sabri, Project administration: Mahmood Ahmad, and Mohanad Muayad Sabri Sabri, Funding acquisition: Mohanad Muayad Sabri Sabri. All authors have read and agreed to the published version of the manuscript.

**References**
[1] Wojtecki, Ł., Iwaszenko, S., Apel, D. B., Bukowska, M., & Makówka, J. (2022). Use of machine learning algorithms to assess the state of rockburst hazard in underground coal mine openings. *Journal of Rock Mechanics and Geotechnical Engineering, 14*(3), 703-713.

[2] Wang, F., & Kaunda, R. (2019). Assessment of rockburst hazard by quantifying the consequence with plastic strain work and released energy in numerical models. *International Journal of Mining Science and Technology, 29*(1), 93-97.

[3] Keneti, A., & Sainsbury, B. A. (2018). Review of published rockburst events and their contributing factors. *Engineering geology, 246*, 361-373.

[4] Zhou, J., Li, X., & Mitri, H. S. (2018). Evaluation method of rockburst: state-of-the-art literature review. *Tunnelling and Underground Space Technology, 81*, 632-659.

[5] Pu, Y., Apel, D. B., Liu, V., & Mitri, H. (2019). Machine learning methods for rockburst prediction-state-of-the-art review. *International Journal of Mining Science and Technology, 29*(4), 565-570.

[6] Jong, S. C., Ong, D. E. L., & Oh, E. (2021). State-of-the-art review of geotechnical-driven artificial intelligence techniques in underground soil-structure interaction. *Tunnelling and Underground Space Technology, 113*, 103946.

[7] Mahmoodzadeh, A., Mohammadi, M., Ghafoor Salim, S., Farid Hama Ali, H., Hashim Ibrahim, H., Nariman Abdulhamid, S., ... & Rashidi, S. (2022). Machine Learning Techniques to Predict Rock Strength Parameters. *Rock Mechanics and Rock Engineering, 55*(3), 1721-1741.

[8] Mahmoodzadeh, A., Mohammadi, M., Farid Hama Ali, H., Hashim Ibrahim, H., Nariman Abdulhamid, S., & Nejati, H. R. (2022). Prediction of safety factors for slope stability: comparison of machine learning techniques. *Natural Hazards, 111*(2), 1771-1799.

[9] Mahmoodzadeh, A., Mohammadi, M., Ali, H. F. H., Salim, S. G., Abdulhamid, S. N., Ibrahim, H. H., & Rashidi, S. (2022). A Markov-based prediction model of tunnel geology, construction time, and construction costs. *Geomechanics and Engineering, 28*(4), 421-435.

[10] Li, D., Liu, Z., Armaghani, D. J., Xiao, P., & Zhou, J. (2022). Novel ensemble intelligence methodologies for rockburst assessment in complex and variable environments. *Scientific reports, 12*(1), 1-23.

[11] Russenes, B. F. (1974). Analysis of rock spalling for tunnels in steep valley sides. *Norwegian Institute of Technology*.

[12] Lu, C. P., Dou, L. M., Liu, B., Xie, Y. S., & Liu, H. S. (2012). Microseismic low-frequency precursor effect of bursting failure of coal and rock. *Journal of Applied Geophysics, 79*, 55-63.

[13] Srinivasan, C., Arora, S. K., & Yaji, R. K. (1997). Use of mining and seismological parameters as premonitors of rockbursts. *International Journal of Rock Mechanics and Mining Sciences, 34*(6), 1001-
1008.

[14] Liu, J. P., Feng, X. T., Li, Y. H., & Sheng, Y. (2013). Studies on temporal and spatial variation of microseismic activities in a deep metal mine. *International Journal of Rock Mechanics and Mining Sciences, 60*, 171-179. https://doi.org/10.1016/j.ijrmms.2012.12.022

[15] Ma, T. H., Tang, C. A., Tang, S. B., Kuang, L., Yu, Q., Kong, D. Q., & Zhu, X. (2018). Rockburst mechanism and prediction based on microseismic monitoring. *International Journal of Rock Mechanics and Mining Sciences, 110*, 177-188.

[16] Ma, X., Westman, E., Slaker, B., Thibodeau, D., & Counter, D. (2018). The b-value evolution of mining-induced seismicity and mainshock occurrences at hard-rock mines. *International Journal of Rock Mechanics and Mining Sciences, 104*, 64-70.

[17] Altindag, R. (2003). Correlation of specific energy with rock brittleness concepts on rock cutting. *Journal of the Southern African Institute of Mining and Metallurgy, 103*(3), 163-171.

[18] Kidybiński, A. (1981, August). Bursting liability indices of coal. In *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* (Vol. 18, No. 4, pp. 295-304). Pergamon.

[19] WATTIMENA, R. K., SIRAIT, B., Widodo, N. P., & MATSUI, K. (2012). Evaluation of rockburst potential in a cut-and-fill mine using energy balance. *International Journal of the JCRM, 8*(1), 19-23.

[20] Wang, J. A., & Park, H. D. (2001). Comprehensive prediction of rockburst based on analysis of strain energy in rocks. *Tunnelling and underground space technology, 16*(1), 49-57.

[21] Mitri, HS*, Tang, B.* & Simon, R. (1999). FE modelling of mining-induced energy release and storage rates. *Journal of the Southern African Institute of Mining and Metallurgy, 99*(2), 103-110.

[22] Ullah, B., Kamran, M., & Rui, Y. (2022). Predictive modeling of short-term rockburst for the stability of subsurface structures using machine learning approaches: T-SNE, K-Means clustering and XGBoost. *Mathematics, 10*(3), 449.

[23] Zhou, J., Li, X., & Shi, X. (2012). Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. *Safety science, 50*(4), 629-644.

[24] Pu, Y., Apel, D. B., & Xu, H. (2019). Rockburst prediction in kimberlite with unsupervised learning method and support vector classifier. *Tunnelling and Underground Space Technology, 90*, 12-18.

[25] Zhao, H., & Chen, B. (2020). Data-driven model for rockburst prediction. *Mathematical Problems in Engineering, 2020*.

[26] Zhao, H., Chen, B., & Zhu, C. (2021). Decision Tree Model for Rockburst Prediction Based on Microseismic Monitoring. *Advances in Civil Engineering, 2021*.
[27] Yin, X., Liu, Q., Pan, Y., Huang, X., Wu, J., & Wang, X. (2021). Strength of stacking technique of ensemble learning in rockburst prediction with imbalanced data: Comparison of eight single and ensemble models. *Natural Resources Research, 30*(2), 1795-1815.

[28] Ahmad, M., Hu, J. L., Hadzima-Nyarko, M., Ahmad, F., Tang, X. W., Rahman, Z. U., ... & Abrar, M. (2021). Rockburst hazard prediction in underground projects using two intelligent classification techniques: a comparative study. *Symmetry, 13*(4), 632.

[29] Wu, S., Wu, Z., & Zhang, C. (2019). Rock burst prediction probability model based on case analysis. *Tunnelling and underground space technology, 93*, 103069.

[30] Sun, Y., Li, G., Zhang, J., & Huang, J. (2021). Rockburst intensity evaluation by a novel systematic and evolved approach: Machine learning booster and application. *Bulletin of Engineering Geology and the Environment, 80*(11), 8385-8395.

[31] Li, D., Liu, Z., Armaghani, D. J., Xiao, P., & Zhou, J. (2022). Novel ensemble intelligence methodologies for rockburst assessment in complex and variable environments. *Scientific reports, 12*(1), 1-23. https://doi.org/10.1038/s41598-022-05594-0

[32] Li, Y.; Wang, C.; Xu, J.; Zhou, Z.; Xu, J.; Cheng, J.(2021). Rockburst Prediction Based on the KPCA-APSO-SVM Model and Its Engineering Application. Shock Vib., 2021, 7968730.

[33] Liang, W., Sari, A., Zhao, G., McKinnon, S. D., & Wu, H. (2020). Short-term rockburst risk prediction using ensemble learning methods. *Natural Hazards, 104*(2), 1923-1946.

[34] Shirani Faradonbeh, R., Shaffiee Haghshenas, S., Taheri, A., & Mikaeil, R. (2020). Application of self-organizing map and fuzzy c-mean techniques for rockburst clustering in deep underground projects. *Neural Computing and Applications, 32*(12), 8545-8559.

[35] Ahmad, M., Katman, H. Y., Al-Mansob, R. A., Ahmad, F., Safdar, M., & Alguno, A. C. (2022). Prediction of Rockburst Intensity Grade in Deep Underground Excavation Using Adaptive Boosting Classifier. Complexity, 2022.

[36] Cai, W., Dou, L., Zhang, M., Cao, W., Shi, J. Q., & Feng, L. (2018). A fuzzy comprehensive evaluation methodology for rock burst forecasting using microseismic monitoring. *Tunnelling and Underground Space Technology, 80*, 232-245.

[37] Kidega, R., Ondiaka, N., Maina, D., Jonah, K., Kamran, M., (2022). Decision Based Uncertainty Model to Predict Rockburst in Underground Engineering Structures Using Gradient Boosting Algorithms. Geomechanics and Engineering. 30 (3). 259-272 https://doi.org/10.12989/gae.2022

[38] Afraei, S., Shahriar, K., & Madani, S. H. (2019). Developing intelligent classification models for rock burst prediction after recognizing significant predictor variables, Section 1: Literature review and data preprocessing procedure. *Tunnelling and Underground Space Technology, 83*, 324-353.
[39] Li, T. Z., Li, Y. X., & Yang, X. L. (2017). Rock burst prediction based on genetic algorithms and extreme learning machine. *Journal of Central South University, 24*(9), 2105-2113.

[40] Roohollah, S. F., & Abbas, T. (2019). Long-term prediction of rockburst hazard in deep underground openings using three robust data mining techniques. *Engineering with Computers, 35*(2), 659–675.

[41] Xue, Y., Bai, C., Qiu, D., Kong, F., & Li, Z. (2020). Predicting rockburst with database using particle swarm optimization and extreme learning machine. *Tunnelling and Underground Space Technology, 98*, 103287.

[42] Pu Y, Apel DB, Xu H (2019b) Rockburst prediction in kimberlite with unsupervised learning method and support vector classifier. *Tunn Undergr Space Technol* 90:12–18. https://doi.org/10.1016/j.tust.2019.04.019

[43] Gao W (2015) Forecasting of rockbursts in deep underground engineering based on abstraction ant colony clustering algorithm. *Nat Hazards* 76(3):1625–1649

[44] Liang, W., Sari, A., Zhao, G., McKinnon, S. D., & Wu, H. (2020). Short-term rockburst risk prediction using ensemble learning methods. *Natural Hazards, 104*(2), 1923-1946. https://doi.org/10.1007/s11069-020-04255-7

[45] Ahmad, M., Hu, J. L., Hadzima-Nyarko, M., Ahmad, F., Tang, X. W., Rahman, Z. U., ... & Abrar, M. (2021). Rockburst hazard prediction in underground projects using two intelligent classification techniques: a comparative study. *Symmetry, 13*(4), 632. https://doi.org/10.3390/sym13040632

[46] Pu, Y., Apel, D.B. and Lingga, B. (2018), “Rockburst prediction in kimberlite using decision tree with incomplete data”, *J. Sustain. Min.*, 17(3), 158-165. https://doi.org/10.1016/j.jsm.2018.07.004.

[47] Zhu, Y. H., Liu, X. R., & Zhou, J. P. (2008). Rockburst prediction analysis based on v-SVR algorithm. *J. China Coal Soc*, 33(3), 277-281.

[48] Zhou, J., Li, X., & Shi, X. (2012). Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. *Safety science, 50*(4), 629-644. https://doi.org/10.1016/j.ssci.2011.08.065

[49] Wang, H., Li, Z., Song, D., He, X., Sobolev, A., & Khan, M. (2021). An intelligent rockburst prediction model based on scorecard methodology. *Minerals, 11*(11), 1294.

[50] Ahmad, M., Katman, H. Y., Al-Mansob, R. A., Ahmad, F., Safdar, M., & Alguno, A. C. (2022). Prediction of Rockburst Intensity Grade in Deep Underground Excavation Using Adaptive Boosting Classifier. *Complexity, 2022*. https://doi.org/10.1155/2022/6156210
[51] Feng, XT, Chen, BR, Zhang, CQ, Li, SJ, Wu, SY. (2013) Mechanism, Warning and Dynamic Control of Rockburst Development Processes; Science Press: Beijing, China, 2013. (In Chinese)

[52] Liang, W., Sari, A., Zhao, G., McKinnon, S. D., & Wu, H. (2020). Short-term rockburst risk prediction using ensemble learning methods. Natural Hazards, 104(2), 1923-1946. https://doi.org/10.1007/s11069-020-04255-7

[53] Abeyratne, D., & Halgamuge, M. N. (2020). Applying Big Data Analytics on Motor Vehicle Collision Predictions in New York City. Intelligent Data Analysis: From Data Gathering to Data Comprehension, 219-239.

[54] Hannachi, A., & Turner, A. G. (2013). Isomap nonlinear dimensionality reduction and bimodality of Asian monsoon convection. Geophysical Research Letters, 40(8), 1653-1658.

[55] Tseng, J. C. H. (2022). An ISOMAP Analysis of Sea Surface Temperature for the Classification and Detection of El Niño & La Niña Events. Atmosphere, 13(6), 919.

[56] Krivov, E., & Belyaev, M. (2016, February). Dimensionality reduction with isomap algorithm for EEG covariance matrices. In 2016 4th International Winter Conference on Brain-Computer Interface (BCI) (pp. 1-4). IEEE.

[57] Sun, W., Halevy, A., Benedetto, J. J., Czaja, W., Liu, C., Wu, H., ... & Li, W. (2014). UL-Isomap based nonlinear dimensionality reduction for hyperspectral imagery classification. ISPRS Journal of Photogrammetry and Remote Sensing, 89, 25-36.

[58] Mehrbani, E., & Kahaei, M. H. (2022). Low-rank isomap algorithm. IET Signal Processing, 16(5), 528-545.

[59] Cho, M., & Park, H. (2009, June). Nonlinear dimension reduction using ISOMap based on class information. In 2009 International Joint Conference on Neural Networks (pp. 566-570). IEEE.

[60] Dubois, D., Ostasiewicz, W., & Prade, H. (2000). Fuzzy sets: history and basic notions. In Fundamentals of fuzzy sets (pp. 21-124). Springer, Boston, MA.

[61] Bundy, A., & Wallen, L. (1984). Fuzzy Set Theory. In Catalogue of Artificial Intelligence Tools (pp. 41-41). Springer, Berlin, Heidelberg.

[62] Lu, Y., Ma, T., Yin, C., Xie, X., Tian, W., & Zhong, S. (2013). Implementation of the fuzzy c-means clustering algorithm in meteorological data. International Journal of Database Theory and Application, 6(6), 1-18.

[63] Parlina, A., Ramli, K., & Murfi, H. (2021). Exposing emerging trends in smart sustainable city research using deep autoencoders-based fuzzy c-means. Sustainability, 13(5), 2876.
[64] Rout, R., Parida, P., Alotaibi, Y., Alghamdi, S., & Khalaf, O. I. (2021). Skin lesion extraction using multiscale morphological local variance reconstruction based watershed transform and fast fuzzy C-means clustering. Symmetry, 13(11), 2085.

[65] Alam, M. S., Rahman, M. M., Hossain, M. A., Islam, M. K., Ahmed, K. M., Ahmed, K. T., ... & Miah, M. S. (2019). Automatic human brain tumor detection in MRI image using template-based K means and improved fuzzy C means clustering algorithm. Big Data and Cognitive Computing, 3(2), 27.

[66] Sun, L., Du, J., & He, Z. (2016, November). Machine learning for nonlinearity mitigation in CAP modulated optical interconnect system by using K-nearest neighbour algorithm. In Asia Communications and Photonics Conference (pp. AS1B-1). Optica Publishing Group.

[67] Rottondi, C., Barletta, L., Giusti, A., & Tomatore, M. (2018). Machine-learning method for quality of transmission prediction of unestablished lightpaths. Journal of Optical Communications and Networking, 10(2), A286-A297.

[68] Pérez, A. E., Torres, J. J. G., & González, N. G. (2019, July). KNN-based Demodulation in gridless Nyquist-WDM Systems affected by Interchannel Interference. In Signal Processing in Photonic Communications (pp. SpTh1E-3). Optical Society of America.

[69] Marquez-Viloria, D., Castano-Londono, L., & Guerrero-Gonzalez, N. (2021). A modified knn algorithm for high-performance computing on fpga of real-time m-qam demodulators. Electronics, 10(5), 627.

[70] Anchalia, P. P., & Roy, K. (2014, January). The k-nearest neighbor algorithm using MapReduce paradigm. In 2014 5th International Conference on Intelligent Systems, Modelling and Simulation (pp. 513-518). IEEE.

[71] Saadatfar, H., Khosravi, S., Joloudari, J. H., Mosavi, A., & Shamshirband, S. (2020). A new K-nearest neighbors classifier for big data based on efficient data pruning. Mathematics, 8(2), 286.

[72] Kamran, M., & Shahani, N. M. (2022). Decision support system for the prediction of mine fire levels in underground coal mining using machine learning approaches. Mining, Metallurgy & Exploration, 39(2), 591-601.

[73] Kamran, M., Shahani, N. Armaghani, D (2022). Decision Support System for Underground Coal Pillar Stability Using Unsupervised and Supervised Machine Learning Approaches. Geomechanics and Engineering, 30 (2), 107-121. https://doi.org/10.12989/gae.2022.30.2.107

[74] Kim, S. W., & Gil, J. M. (2019). Research paper classification systems based on TF-IDF and LDA schemes. Human-Centric Computing and Information Sciences.

[75] Ma, Y., Peng, M., Xue, W., & Ji, X. (2013, April). A dynamic affinity propagation clustering algorithm for cell outage detection in self-healing networks. In 2013 IEEE Wireless Communications and Networking Conference (WCNC) (pp. 2266-2270). IEEE.
Figures

Figure 1
Flowchart of the proposed study
Figure 2

The distribution of different rockburst level
Figure 3

Voilen chart of significant features employed in the study
Figure 4

The mechanism of Isomap
Figure 5

Flowchart of fuzzy c-means algorithm
Figure 6

A simple 3D multi-class KNN model for the rockburst dataset
**Figure 7**

3D visualization of ISOMAP based rockburst database
Figure 8

Fuzzy C-means algorithm visualization on Isomap database
Figure 9

Confusion matrix of KNN on testing dataset