

# Predicting Dangerous Seismic Events in Active Coal Mines Through Data Mining

**Ligendra Kumar Verma**

*Department of Mining Engineering  
National Institute of Technology Raipur,  
Chhattisgarh 492001, India.*

**D.C. Jhariya**

*Department of Applied Geology  
National Institute of Technology Raipur,  
Chhattisgarh 492001, India.*

**P. Y. Dhekne**

*Department of Mining Engineering  
National Institute of Technology Raipur,  
Chhattisgarh 492001, India.*

**Pankaj Kumar Dewangan**

*Department of Mining Engineering  
National Institute of Technology Raipur,  
Chhattisgarh 492001, India.*

## Abstract

The objective of this paper is to devise a prediction model for detecting periods of increased seismic activity that endangers miners working underground in coal mines. In order to predict the abnormal seismic activity data mining method is employed. Benchmark support vector machine classifier is applied and obtained 99.23% of accuracy in large 30,000 dataset. The data set has been taken from Research and Development Centre EMAG [1] and UCI repository [2].

**Keywords:** Classification; Support Vector Machine; Seismic Activity prediction

## INTRODUCTION

A coal mine bump is a spontaneous, violent activity that occurred in mine, due to the explosive collapse of a wall or one or more support pillars, that is also called a rock burst [3]. These pillars are left in place during room and pillar mining, where an original narrow passage is dug and then substantially widened as ore is removed. Coal mine bumps, sometimes referred to as outbursts, bursts, or bounces, have been recognized as a serious problem in mining for more than 75 years. The analysis of historical seismic data shows a high correlation between observed mining seismicity and geometric parameters at the sites [4]. The past (historical data), bumps has been acknowledged as having a greater likelihood of occurrence at depths greater than 300 m (1,000 ft), in the presence of strong roof and/or floor, and when an unusually strong massive unit exists in the main roof. The Miners working under these conditions need to be constantly aware of the possibility of a coal bump to prevent from extreme conditions.

These problems necessitate to introduce some efficient technique which alarm of this rock burst event. Accuracy of so far achieved is not sufficient for industry norms. Complexity of seismic processes and big disproportion between the number of low-energy seismic events and the number of high-energy phenomena (e.g.  $> 10^4$ J) causes the statistical techniques to be introduced to predict seismic hazard. In this

paper, we discuss some data mining technique that may give a better solution to solve this challenging problem. The objective of this paper is to devise a reliable prediction model for detecting periods of increased seismic activity that alarm the endangers miners working in underground coal mines.

## Dataset :

In this study, we have taken two dataset. First dataset has been taken from UCI repository [1]. The dataset contains 2584 instances with 18 attribute and one class label. The class label denotes two state hazardous an non- hazardous state.

In second data set, it is provided by Research and Development Centre EMAG. It consists of hourly aggregated readings from seismic sensors that count the number of seismic bumps perceived at longwalls and measure their total energy. Data records are composed of 24 consecutive hours of such readings coupled with the most recent assessments of the conditions at the longwalls made by mining experts. The target attribute IND the data corresponds to information, whether in a following period of 8 hours the total energy of seismic bumps exceeds a warning level, i.e.  $5 \times 10^4$  Joules. In total, the training file contains 30,893 records, each corresponding to 24 hours of measurements. There is a total number of 541 attributes in the dataset. The example dataset is shown in Table 1. The table consist of A1-A540 attribute and one class label. The attributes are main\_working\_id total\_bumps\_energy, total\_tremors\_energy, total\_destressing\_blasts\_energy, total\_seismic\_energy, latest\_progress\_estimation\_l, latest\_progress\_estimation\_r, latest\_seismic\_assessment, latest\_seismoacoustic\_assessment, latest\_comprehensive\_assessment, latest\_hazards\_assessment, latest\_maximum\_yield, latest\_maximum\_meter, count\_e2.1-count\_e2.24, count\_e3.1-count 23.24, count\_e4.1to count\_e4.24, count\_e5.1to count\_e5.24, count\_e6plus.1-count\_e6plus.24, sum\_e2.1-sum\_e2.24 similarly for e3,e4,e5, sum\_e6plus.1- sum\_e6plus.24, total\_number\_of\_bumps.1-total\_number\_of\_bumps.24, number\_of\_rock\_bursts.1-number\_of\_rock\_bursts.24, number\_of\_destressing\_blasts.1-number\_of\_destressing\_blasts.24, highest\_bump\_energy.1-highest\_bump\_energy.24, max\_gactivity.1- max\_gactivity.24,

max\_genenergy.1- max\_genenergy.24, avg\_gactivity.1- avg\_difference\_in\_gactivity.1- avg\_difference\_in\_gactivity.24,  
 avg\_gactivity.24,avg\_genenergy.1-avg\_genenergy.24, max\_difference\_in\_gactivity.1- max\_difference\_in\_gactivity.24,  
 max\_difference\_in\_genenergy.1- max\_difference\_in\_genenergy.24, avg\_difference\_in\_genenergy.1- avg\_difference\_in\_genenergy.24  
 and Class label.

**Table 1:** Dataset example

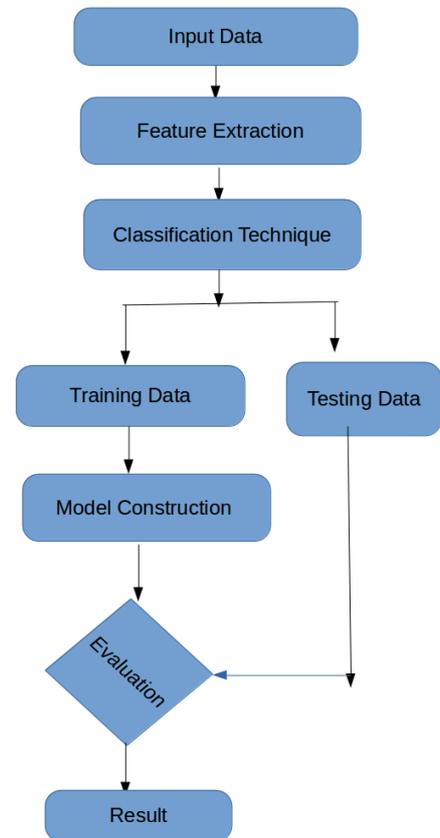
A529	A530	A531	A532	A533	A534	A535	A536	A537	A539	A540	Class
1.92	-58.63	91.54	28.39	77.40	-17.82	136.19	75.75	68.38	40.73	128.50	normal
-58.63	91.54	28.39	77.40	-17.82	136.19	75.75	68.38	112.92	128.50	18.92	normal
91.54	28.39	77.40	-17.82	136.19	75.75	68.38	112.92	40.73	18.92	310.29	normal
28.39	77.40	-17.82	136.19	75.75	68.38	112.92	40.73	128.50	310.29	-45.44	normal
77.40	-17.82	136.19	75.75	68.38	112.92	40.73	128.50	18.92	-45.44	32.68	Extreme
-17.82	136.19	75.75	68.38	112.92	40.73	128.50	18.92	310.29	32.68	0.29	normal
136.19	75.75	68.38	112.92	40.73	128.50	18.92	310.29	-45.44	0.29	12.72	Extreme

Section 2 explores the related work of different classifier. Section 3 presents the proposed methodology adopted. The results of the application of the model and the discussion of results are presented in Section 4. Paper is concluded in Section .

**RELATED WORK**

Classification is an important problem in data mining. Under the guise of supervised learning, classification has been studied extensively by the Machine learning and data mining community as a possible solution to the “knowledge acquisition” or “knowledge extraction” problem. The input to the classifier construction is a training set of records, each of which is tagged with a class label. A set of attribute values defines each record. It is a supervised learning problem which has many applications. Data mining techniques propose promising ways to uncover hidden patterns within large amounts of data. These hidden patterns can potentially be used to predict future behavior from complex dataset. In this section we discuss some of the classification techniques for normal and extreme condition.

**Definition 1 (Classification).** Classification is the task of learning a target function  $f$  that maps each attribute set  $x$  to one of the predefined class labels  $y$ . The target function is also known informally as a classification model. The process of classification process is shown in Figure 1.



**Figure 1.** Taxonomy of classification process

*A. Decision Tree*

Decision tree is well known recursive tree like classification technique used in several fields like medical data, digital image processing, machine learning. A decision tree classifier is built in two phases growing phase followed by pruning

phase [9]. In growth phase the tree is built by recursively partitioning the data until each partition is either pure or very small. In grown up phase the constructed tree may be overfitted. To remove overfitting In pruning phase the tree

**Tree Building phase :**

The algorithm for building tree

Procedure buildTree (S)

- step(1). Initialize root node using dataset S
- step(2). Compute the The total Information of data eq.(1)
- step(3). Compute the entropy for each attribute using categorized data
- step(4). Compute the information gain using  $I(p,n) - \text{Entropy}$
- step(5). If  $a\_best$  be the attribute with the highest normalized information gain
- step(6). Create a decision *node* that splits on  $a\_best$
- step(7). Recur on the sublists obtained by splitting on  $a\_best$ , and add those node as child

In above algorithm the split is based on Information gain and entropy defined in equation 1 and 2 respectively. Information gain check the importance of feature vector in order to decide the class label. Entropy is used to check the impurity in the dataset.

$$I(p, n) = -\frac{p}{P+n} \log_2 \frac{p}{P+n} - \frac{n}{P+n} \log_2 \frac{n}{P+n} \text{ ----- eq(1)}$$

$$\text{Entropy} = -\sum_i p_i \log_2(p_i) \text{ ----- eq(2)}$$

**B. Neural Network**

Neural Network is made up of a structure of network with interconnected units. Each of these units consists of input/output characteristics of data. It consists of units (neurons), arranged in layers that convert an input vector into some output. Each unit takes an input as a feature vector, applies a (often nonlinear) function to it and then passes the output on to the next layer. The input is set of feature vector and output is class label of the dataset. The function is a summation of input feature vectors. It is based on the concept of regression model.

$$f(x) = \varphi_0 + \varphi_1 x_1 + \varphi_2 x_1 x_2 + \varphi_3 x_1 x_2 x_3 + \dots \text{---eq(3)}$$

The algorithm iterates till the mean square error is not minimized.

**C. Support Vector Machine**

Support Vector Machine(SVM) has proven one of the powerful technique of machine learning for classifications. SVM is a modern mechanism for two-class classification, regression, and clustering problems. This method identifies the support vector in order to obtain the boundary between different classes. SVMs [5,6]) try to map the original training data into a higher dimensional space by a kernel function  $\Phi$ . The kernel function can be linear, radial or sigmoid. The linear kernel is applicale if the objects are linearly seperable, similarly the sigmoid and radial kernel are applicable for non-linear boundary. The remove the word aim is then to identify a

linear separating hyper-plane, with the maximal margin between negative and positive samples, in this higher dimensional space. From the mathematical point of view, given a training set of instance label pairs  $(x_i, y_i); i = 1, \dots, l$ , where  $x_i \in R^n$  and  $y \in \{-1, +1\}$  search for a solution to the optimization defined in equation (4).

$$\min_{w,b} \xi \frac{1}{2} w^T w + C \sum_{i=1} \xi_i \text{----- eq (4)}$$

Subject to the training vector  $x_i$  are mapped into the higher dimensional space by the function.

**D. Naïve Bays Classifier [8]**

Naive Bayes is one of the simplest technique for constructing classifiers based on bayes theorem. It assigns class labels to problem instances where the input vectors are represented as vectors of feature values and class labels are drawn from some finite set. All naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable.

$$P(y | x_1, \dots, x_n) = \frac{P(y)P(x_1, \dots, x_n | y)}{P(x_1, \dots, x_n)}$$

**EXPERIMENTAL RESULTS**

In this section we discuss the experimental result which we performed in both the dataset All the experimental studies were performed in Windows operating system with 2.8Ghz processor, 500 GB harddisk in Statistical software R and weka open source software. We applied Random Forest, Naïve Bays Classifier, Logistic regression model, Classification via Regression, Bagging, One R, Zero R, Simple CART, REP Tree, SVM with kernel Radial, Polynomial, Linear and sigmoid kernels model.

**Performance Evaluation**

**Accuracy**

Accuracy is the total percentage of correctly classified divided by the total number of classifications [7]. The confusion matrix to compute the accuracy is shown in table 2.

**Table 2:** Confusion Matrix

	Normal	Warning
Normal	TP (159)	FP (1)
Warning	FN (14)	TN (289)

$$\text{Accuracy}(\%) = \frac{TP + TN}{TP + FN + TN + FP} \times 100$$

**Precision**

Precision is a measure of the accuracy provided that a specific class has been predicted. It can be computed using following equation.

$$\text{Precision} = \frac{TP}{(TP + FP)}$$

**Recall**

Recall is a measure of the ability of a predictive model to select instances of a certain class from a data set. It is commonly also called sensitivity, and corresponds to the true positive rate.

$$Recall = Sensitivity = \frac{TP}{(TP + FN)}$$

**Specificity**

It is also called true negative rate. It is defined as the percentage of benign samples correctly classified as benign. It is calculated using-

$$Specificity = \frac{TN}{(TN + FP)}$$

**F-measure**

A measure that combines precision and recall is the harmonic mean of precision and recall, the traditional F-measure or balanced F-score:

$$F - Measure = \frac{2 * Precision * Recall}{Precision + Recall}$$

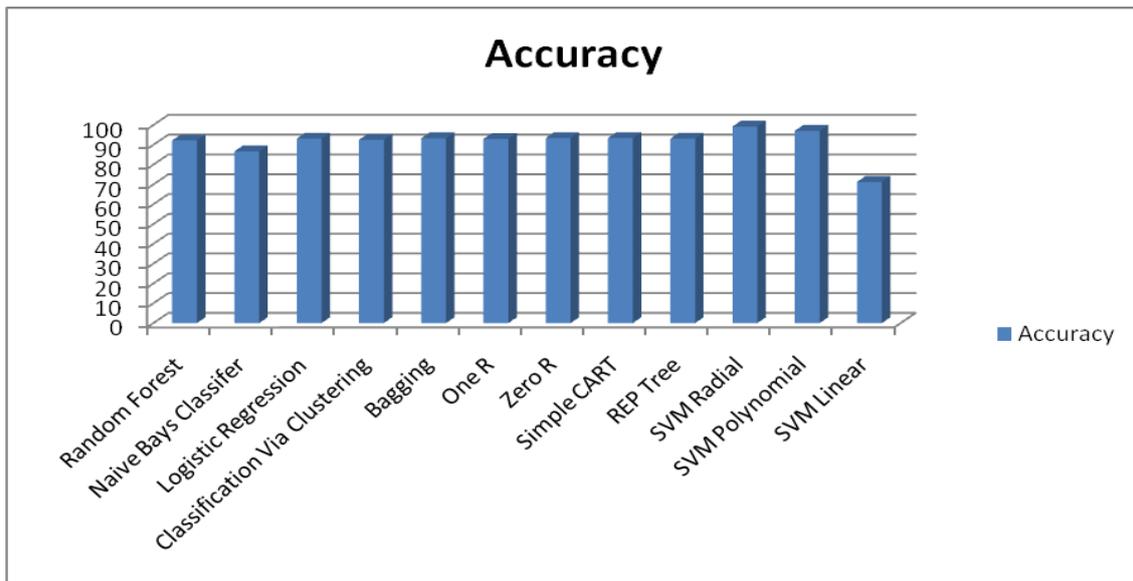
The experimental results with Seismic Dataset of UCI Repository and dataset 2 are shown in Table 3 and 4 respectively. The overall accuracy of different classifier model is shown in Figure 2.

**Table 3 :** Expermental Result with Seismic Dataset of UCI Repository

Classifier	Accuracy	True Positive Rate	False Positive Rate	Precision	Recall	F- Measure	ROC Area
Random Forest	92.2	0.921	0.875	0.888	0.921	0.902	0.692
Naive Bays Classifier	86.7	0.867	0.567	0.907	0.867	0.885	0.755
Logistic Regression	93.1	0.931	0.913	0.89	0.931	0.904	0.749
Classification Via Clustering	92.5	0.925	0.88	0.89	0.925	0.904	0.672
Bagging	93.3	0.933	0.912	0.898	0.933	0.905	0.751
One R	93.0	0.93	0.929	0.877	0.93	0.901	0.5
Zero R	93.4	0.934	0.934	0.873	0.934	0.902	0.5
Simple CART	93.4	0.934	0.934	0.873	0.934	0.902	0.5
REP Tree	93.1	.931	.929	.879	.931	.902	.655
SVM Radial	99.234	.992	.98	.987	.968	.903	.905
SVM Polynomial	97.057	.997	.97	.978	.988	.906	.906
SVM Linear	71.277	.712	.723	.725	.745	.756	.756

**Table 4:** Experimental Result with Dataset II

Classifier	Accuracy	True Positive Rate	False Positive Rate	Precision	Recall	F- Measure	ROC Area
Random Forest	98	.987	.197	.987	.987	.987	.98
Naive Bays Classifier	74.4	.871	.432	.933	.871	.896	.744
One R	66.4	.957	.63	.95	.957	.95	.664
Zero R	49.9	.948	.948	.898	.948	.922	.499



**Figure 2.** Accuracy with different classifiers

## CONCLUSION

In this paper, we devised a prediction model for detecting periods of increased seismic activity that endangers miners working underground in coal mines. Various prediction models like Random forest, Naïve bays classifier, decision and other benchmark classifiers are used to evaluate the performance of classifiers. The experiments are performed Research and Development Centre EMAG [1] and UCI repository [2] dataset. In experiment we found the performance of random forest and SVM is best among other prediction and classification models.

## REFERENCES

- [1] <https://knowledgepit.fedcsis.org/mod/page/view.php?id=889> (Accessed on December, 2015)
- [2] <https://archive.ics.uci.edu/ml/datasets/seismic-bumps#>. (Accessed on November, 2015)
- [3] [[https://en.wikipedia.org/wiki/Coal\\_mine\\_bump](https://en.wikipedia.org/wiki/Coal_mine_bump)] (Accessed on October 2015)
- [4] John L. Ellenberger, Keith A. Heasley. Coal mine seismicity and bumps: historical case studies and current field activity.
- [5] B. E. Boser, I. M. Guyon, and V. N. Vapnik, A training algorithm for optimal margin classifiers, in Proc. Int. Conf. on Computational Learning Theory (COLT92), 144152, 1992.
- [6] Cortes C. and V. Vapnik, Support-vector networks, Machine Learning 20, 273297, 1995.
- [7] Provost, F., Fawcett, T., & Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. Proceeding of the 15th International Conference on Machine Learning (pp. 445 {453}). Morgan Kaufmann, San Francisco, CA.
- [8] [https://en.wikipedia.org/wiki/Naive\\_Bayes\\_classifier](https://en.wikipedia.org/wiki/Naive_Bayes_classifier) (Accessed on October, 2015)
- [9] J. Ross Quinlan. C4.5: Programs for Machine Learning. Morgan Kaufman, 1993.