Journal of Mines, Metals and Fuels, 70(4) : 171-181; 2022. DOI: 10.18311/jmmf/2022/30018

Print ISSN : 0022-2755

Journal of Mines, Metals and Fuels

Contents available at: www.informaticsjournals.com/index.php/jmmf

Artificial Intelligence Model for Prediction of Local and Main FALL in caving Panel of Bord and Pillar Method of Mining

Ram Bilash Prajapati[1*], Rabindra Kumar Sinha[1], R. N. Gupta[2], Sikandar Kumar[1] and Deepti Prajapati[3]

[1]Department of Mining Engineering, Indian Institute of Technology (ISM), Dhanbad – 826004, Jharkhand, India; rbprajapati2010@gmail.com, rksinha@iitism.ac.in, sikandarbit2k11@gmail.com
[2]National Institute of Rock Mechanics, Bangalore – 560070, Karnataka, India; guptarn1942@rediffmail.com
[3]All India Institute of Medical Sciences, Patna – 801507, Bihar, India; deeptiprajapati15@gmail.com

Abstract

Depillaring with caving method of mining is a common practice in Indian coalfields and so is the occurrence of fall in goaf area, which can be considered as a boon in disguise as it allows wining of coal from large reserves but this becomes a curse just because of its unpredicted occurrence. Various empirical and statistical models are developed after idealization of several complicated mechanisms but they are not able to predict roof fall accurately especially in caving panels. Therefore, a new approach based on Artificial Intelligence is used to predict the sequence of local and main fall in caving panel taking into account a host of geotechnical and mining parameters of the mine. Mathematical equations and hidden calculations of artificial neural networks are known to have the capability of learning and analyzing records endlessly. Two different models have been deployed after optimal hyper parameter optimization to predict the occurrence of fall and to characterize the nature of fall (local or main) with considerable and reliable accuracy.

Keywords: *Bord and Pillar, Caving, Deep Learning Algorithm, Deep Neural Network, Hyper Parameter Optimization, Local Fall, Main Fall, Talos*

1. Introduction

Artificial intelligence is competing with human intelligence nowadays, using neural network algorithms. Simon (1957) has predicted way back in 1957 that "Machine will think, learn and create" like humans. Before rock mechanics advancement in India, experienced mining personnel used to predict roof fall when strata started talking to them with admissible accuracy and saved many valuable lives. Artificial Neural Network (ANN) has been used for roof fall hazard assessment (Malkowski and Juszynski, 2021), roof fall hazard detection (Isleyen *et al.*, 2021),

*Author for correspondence

predicting roof fall rate by novel fuzzy interface system (Razani *et al*., 2013), forecasting shield pressures (Deb *et al*., 2006), pillar stress prediction (Monjezi *et al*., 2009) and also being used for weather forecasting, voice and image recognition, stock market prediction, health care, driverless vehicle etc.

Success of bord and pillar mining is dependent on stability of four basic structures i.e. roof strata, bord, pillar and support at various stages of depillaring which are highly dependent upon in situ stresses (Sheorey, 1994), mining induced stresses (Singh *et al*., 1996) and; natural as well as artificial supports provided during extraction.

Jena *et al*. (2016) have observed that loading dynamism produced due to unequal strata movement is proportional to increase in goaf span and area and forms a dynamic loading zone which affects the working area and goaf edges. The loading effect is highest near the goaf edges which decreases farther in dynamic loading zone of working area and requires advance support in caving panel.

For prediction of main fall and periodic caving span, a theoretical model has been given by Obert and Duvall, (1967) based on plate-beam theory and by Majumdar, (1986)"type":"article-journal"},"uris":["http://www.mendeley.com/documents/?uuid=6a2cb130-8df3-4207-9ad0-a29e91420db2",""http://www.mendeley.com/documents/?uuid=306012fb-1898-4992-a434-4f56843228e5"]}],"mendeley":{"formattedCitation":"(Majumdar, 1986 based on bending moment approach. Some researchers have given empirical relation for span of main fall (Pawlowicz, 1967; Peng and Chiang, 1984) and periodic caving span (Peng and Chiang, 1984; Sarkar and Dhar, 1993; Sarkar, 1998). Other researchers have given models for assessment of roof caving and caving span (Ghose and Dutta, 1987; Sarkar, 1998; Sarkar and Dhar, 1993; Sheorey, 1984). Nimaje and Sai (2015) have developed a software for calculation of roof fall risk, based on probability and consequences using parameters responsible for roof fall in underground mines.

McCulloch and Pitts, 1943 have introduced the first ANN, which is a powerful computational model consisting of large number of elemental units called neurons arranged in input, hidden and output layers simulating biological neurons (Fausett, 1993).

The revolutionary developments in 1980s led to the resurgence of ANNs, namely the energy analysis of feedback neural networks by (Hopfield, 1982, 1984). The ANNs can separate non-linear data by assigning certain weights for each input and after multiplying them, sums the product, and passes it through a non-linear transfer function to provide the output (Lee *et al*., 2003). The activation for i^{th} neuron in l^{th} layer is calculated as:

$$z_i^{[l]} = w_i^{[l]T} a^{[l-1]} + b_i^{[l]}$$

$$a_i^{[l]} = f\left(z_i^{[l]}\right)$$

Similarly, all the activations for a given layer can be calculated as:

$$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}$$

$$A^{[l]} = f\left(Z^{[l]}\right)$$

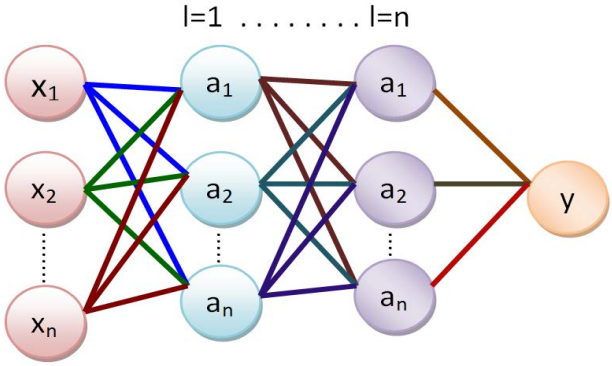


Figure 1. Basic structure of a neural network.

A neural network is an interconnected aggregation of units, its characteristics are determined by the topology and properties of the 'neuron' (Russell and Norvig, 2021). A multilayer perceptron in a neural network can be formed by stacking the neurons to produce a layer (Haykin, 1999) and then cascading these layers together (Nazzal *et al*., 2008) as shown in Figure 1.

2. Proposed Methodology

The roadway to model creation has been summarized in Figure 2. Extensive literature review has been done to analyse the contributory factors and then select the parameters to be included for model creation as discussed in Section 3.0. This is followed by visits to various mines for collection of relevant data (Section 4.0). This is the passed through various pre-processing techniques to enhance the quality of data which is a crucial step for creating a good model (Section 4.2).

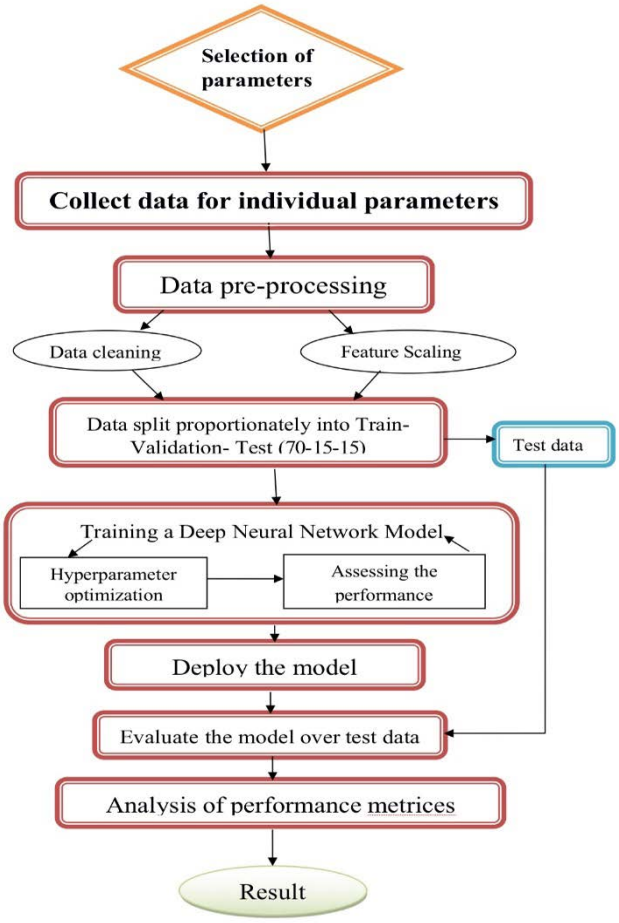


Figure 2. Roadway to model.

Once the raw data gets pre-processed, it is then introduced to Keras, (Chollet *et al*., 2015) deep neural network algorithm, written in Python (Rossum and Drake, 2009) programming language on Google Colab (Bisong, 2019).

Due to the presence of a lot of hyper parameters (Section 4.4), training the neural network and achieving a desirable performance becomes difficult. Thus, a neural network hyper parameter optimization tool, Talos (Talos, 2019) is required to extract the best set of hype parameters (Figure 3), which has been used recently to classify air quality (Parashar and Sonker, 2019).

After achieving a reasonably good model the scan has been stopped and the model is deployed and then it is evaluated on the test dataset which has been kept aside from the training process. The performance of the model has been analysed in terms of confusion matrix, classification report receiver operator characteristic (ROC) curve and area under curve (AUC).

3. Contributory Parameters of Local and Main Fall

Local fall takes place within twenty-four to forty-eight hours after withdrawal of support in the goaf area. It does not extend up to the surface and affects only few meters of the strata above the coal seam whereas main fall affects the surface and takes place long after local fall.

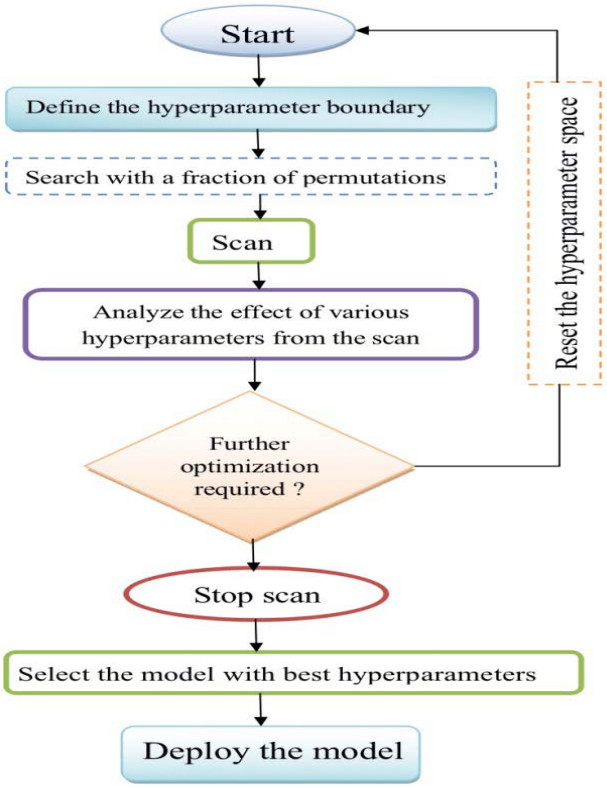


Figure 3. Hyper parameter optimization using Talos.

The factors that have ability to affect the roof-fall directly or indirectly are geological disturbances like joints, folds, faults, dykes, shape of deposit etc, physico-mechanical properties of deposit and surrounding rock, immediate roof, Rock Mass Rating (RMR) or Coal Mine Rock Rating (CMRR), depth of cover, systematic support rule, goaf area, method of excavation, working height etc. (Ghasemi *et al*., 2012; Kumar *et al*., 2019; Mark and Michael, 2017; Mark and Molinda, 2007; Palei and Das, 2009). A parameter is considered to be a variable internal to the model, its value can be approximated from the data (Torres, 2018). For model creation the actual names of the parameters have been used but for ease of depiction the parameters have been renamed as (h1-h36):

h1 Quantity of coal in the panel (t)
h2 Seam thickness (m)
h3 Gradient of dip (degree)
h4 Average thickness of cover/ Depth (m)
h5 Average gallery width (m)
h6 Pillar number
h7 Length of pillar (m)
h8 Width of pillar (m)
h9 Length of depillaring panel (m)
h10 Width of depillaring panel (m)
h11 Sandstone percentage (%)
h12 Coal left in immediate roof (m)
h13 Coal layer thickness value (cm)
h14 Coal SFI value
h15 Coal slack durability index (SDI,%)
h16 Coal UCS (kg/cm^2)
h17 Coal GWS value (mL/min)
h18 Total rating of coal
h19 Thickness of immediate rock (m)
h20 Rock layer thickness value (cm)
h21 Rock SFI
h22 Rock SDI (%)
h23 Rock UCS (kg/cm^2)
h24 Rock GWS value (mL/min)
h25 Total rating of rock
h26 Combined Rock Mass Rating (RMR)
h27 Rock load(t/m^2)
h28 Tensile strength (kg/cm^2)
h29 Poisson's ratio
h30 Young's modulus (GPa)
h31 Area of fall (m^2)
h32 Days
h33 Rock type shale (S)
h34 Rock type shaly sandstone (SH SST)
h35 Rock type sandstone (SST)
h36 Fall (Model 1 – No fall(0) or Fall(1),
Model 2 – Local fall(0) or Main fall(1)).

where, UCS - Uni-axial compressive strength.
SFI - Structural feature Indices
GWS - Ground water seepage

4. Data Acquisition

The data is collected from two different areas, Dhori and Bhurkunda. The Dhori Khas colliery is situated in the Dhori area of East Bokaro coalfields. The mine is bounded by latitudes 23^{0}46'00" to 23^{0}46'53" North and longitudes 86^{0}00'22" to 86^{0}01'22" East. The lithological formations present here are Talcher, Karharbari and Barakar formations of Lower Gondwana, the later forms major part of the minefield area. It comprises of two units, 7 and 8 Incline and 4,5 and 6 Incline. A borehole section (CMED-19) out of 9 boreholes drilled is shown in Table 1 and plan of mine Figure 4. The Karo Special Seam- III is one of them which in crops in the area. The average thickness of the seam is 2.5 m dipping at 1 in 6.5 due S 46^o W, has been developed along floor in full height on bord and pillar pattern. Bhurkunda colliery is located in Barka Sayal area of South Karanpura coalfields. The mine is bounded by latitudes 23^{o}39'00" to 23^{o}41'00" North and longitudes 85^{o}21'00" to 85^{o}23'00" East. The geological formations belong to Lower Gondwana group, except the Raniganj formation all other formations are present here namely, Talcher, Karharbari, Barren measures, Barakar measures.

Table 1. Borehole section – CMED -19 (Dhori)

Bore hole section	Thickness (m)
Soil	3.05
Fine grain sandstone	13.4
Sandy Shale	0.91
Corse grain sandstone	5.73
Sandy Shale	2.23
Gray Shale	2.07
Shally coal	1.07
Fine grain sandstone	13.91
Karo special seam IV	1.22
Sandy shale	1.00
Fine grain sandstone	8.54
Karo special seam III	2.32 (Being worked)
Corse grain sandstone	5.70

The Barakar formation being the major coal bearing formation. There are 11 seams, with thickness of 2.89 to 5.68m and average gradient of 1 in 6, out of which the Upper Semana, Lower Semana and Hathidari seams are developed on bord and pillar pattern.

At Dhori and Bhurkunda collieries, the depillaring operation is being done using semi-mechanized method of mining with side discharge loader, universal drilling machine and full column grouted roof bolting system for strata control. Telltale, load cell and convergence recorders are be used for strata monitoring. Rope haulage system with coal-tub is used for coal evacuation from face to surface.

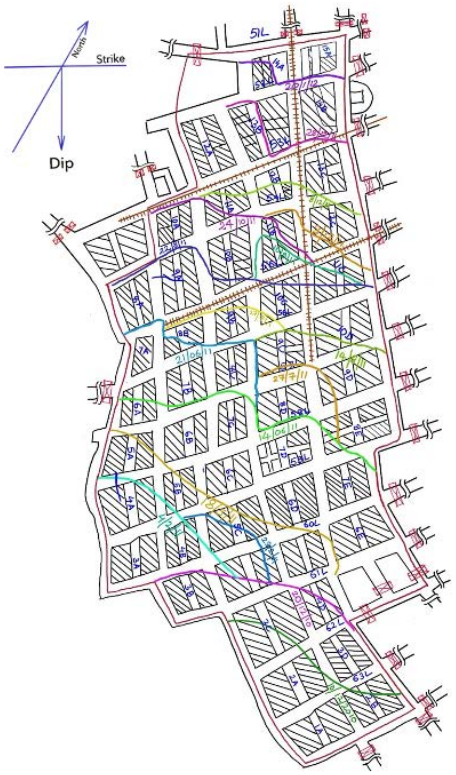


Figure 4. Part hand plan at 4, 5 and 6 incline, Panel A, Dhori Khas Colliery.

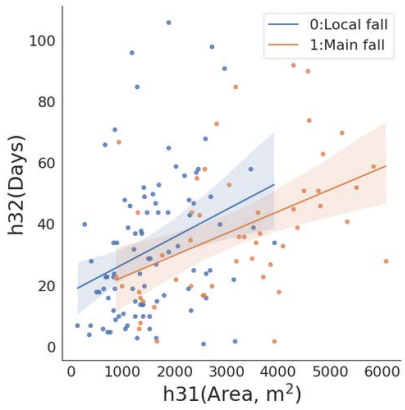


Figure 5. The plot between area and days represents that a fall of larger area takes a longer time to occur, irrespective of it being local or main.

4.1 Data Analysis

To analyze the relationship amongst various parameters, the data has been visualized through scatter (Figures 5–7) and Pearson Correlation plot (Figure 8). A total of 324 records have been collected which has been employed to form Model 1, the training dataset details can be seen in Table 2. Dataset for creating Model 2 contains total of 208 records.

Table 2. Overview of dataset for Model 2

	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10	h11	h12	h13	h14	h15	h16
count	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000
mean	56210.904545	3.273864	0.177245	86.490727	4.942955	55.654545	25.342409	21.605727	245.850000	132.659091	77.248182	0.256045	1.476136	2.322727	32.768636	90.526364
std	26165.908477	0.551671	0.123969	63.463793	0.975643	21.904499	5.013646	4.849960	64.369582	39.689401	17.537895	0.364428	2.101271	3.303579	46.607372	129.368997
min	3400.000000	2.700000	0.120000	27.500000	4.100000	28.000000	20.000000	16.000000	130.000000	76.000000	55.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	38400.000000	2.800000	0.140000	49.500000	4.300000	38.000000	21.000000	18.000000	200.000000	113.750000	55.000000	0.000000	0.000000	0.000000	0.000000	0.000000
50%	54400.000000	3.100000	0.150000	58.000000	4.500000	49.000000	24.000000	20.000000	250.000000	125.000000	82.800000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	80000.000000	3.750000	0.166000	110.625000	5.537500	85.000000	29.417500	24.420000	300.000000	150.000000	94.000000	0.750000	4.307500	7.000000	98.500000	248.200000
max	96000.000000	4.270000	0.666000	277.030000	7.750000	98.000000	36.620000	34.750000	345.000000	250.000000	94.000000	0.800000	4.600000	7.000000	99.600000	295.400000

	h17	h18	h19	h20	h21	h22	h23	h24	h25	h26	h27	h28	h29	h30	h31	h32	h36
count	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000
mean	0.131818	18.813455	1.743955	13.496091	6.572727	98.320455	693.828273	0.450000	69.943409	59.524727	2.341773	69.178918	0.141409	5.143727	1663.290909	29.977273	0.650000
std	0.339064	26.763017	0.364428	4.594357	1.534737	0.661319	196.779826	0.498628	5.486298	9.010425	1.324934	19.859921	0.068547	0.614631	1304.025849	21.338836	0.478057
min	0.000000	0.000000	1.200000	5.100000	5.000000	96.600000	322.700000	0.000000	55.400000	40.400000	0.820000	32.270000	0.060000	4.220000	98.000000	1.000000	0.000000
25%	0.000000	0.000000	1.250000	10.000000	5.000000	97.800000	526.350000	0.000000	64.300000	49.440000	1.430000	52.635000	0.082000	4.600000	687.500000	14.000000	0.000000
50%	0.000000	0.000000	2.000000	14.000000	7.000000	98.500000	748.600000	0.000000	71.720000	64.540000	1.700000	74.860000	0.120000	5.220000	1347.500000	24.000000	1.000000
75%	0.000000	56.350000	2.000000	18.400000	7.000000	98.900000	863.075000	1.000000	73.280000	65.950000	3.400000	86.307500	0.220000	5.400000	2289.500000	40.250000	1.000000
max	1.000000	58.500000	2.000000	20.000000	13.000000	99.000000	984.400000	1.000000	79.820000	71.830000	6.100000	98.440000	0.280000	7.040000	7951.000000	106.000000	1.000000

Table 3. Data characteristics after robust scaling

	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10
count	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000	220.000000
mean	0.043531	0.183014	1.047902	0.466106	0.357943	0.141586	0.159478	0.250113	-0.041500	0.211285
std	0.628988	0.580706	4.768034	1.038262	0.788398	0.466053	0.595622	0.755446	0.643696	1.094880
min	-1.225962	-0.421053	-1.153846	-0.498978	-0.323232	-0.446809	-0.475200	-0.623053	-1.200000	-1.351724
25%	-0.384615	-0.315789	-0.384615	-0.139059	-0.161616	-0.234043	-0.356400	-0.311526	-0.500000	-0.310345
50%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	0.615385	0.684211	0.615385	0.860941	0.838384	0.765957	0.643600	0.688474	0.500000	0.689655
max	1.000000	1.231579	19.846154	3.583313	2.626263	1.042553	1.499257	2.297508	0.950000	3.448276

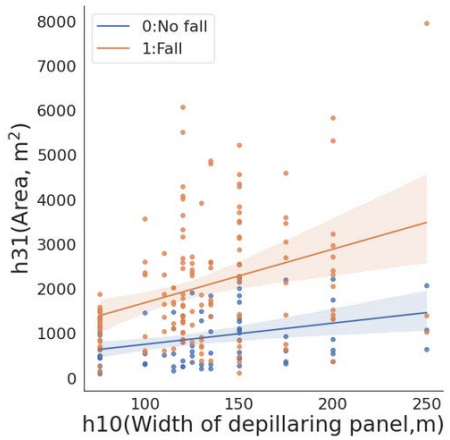


Figure 6. As the width of the depillaring panel increases, there is increased chance of occurrence of falls of larger area.

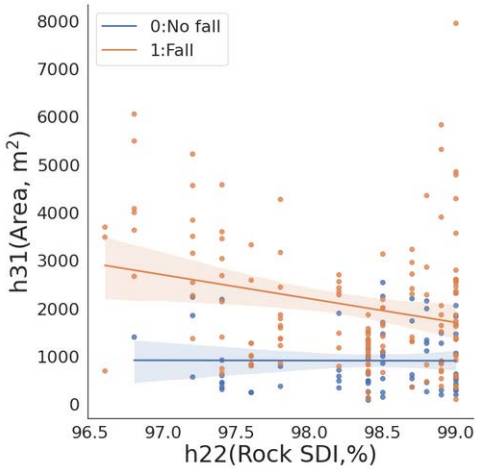


Figure 7. As Rock SDI increases, falls of smaller area usually occurs as high SDI weakens the rock.

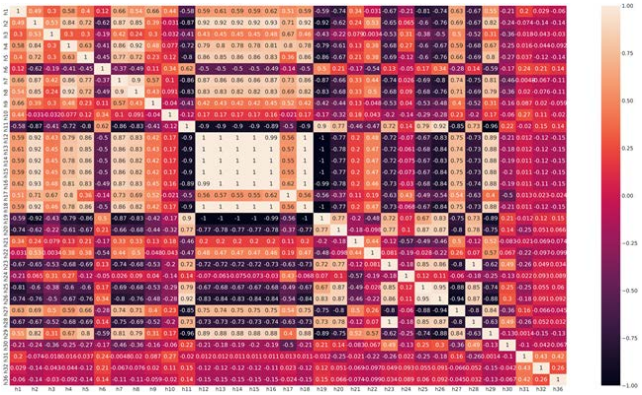


Figure 8. The Pearson correlation plot depicts that there is positive correlation between fall and area, h31 (0.42) and days, h32 (0.26) and negative correlation with seam thickness, h2 (-0.14) and average gallery width, h5 (-0.14).

4.2 Data Pre-processing

Fan *et al.* (2021) have summarized various data pre-processing tasks as: Data cleaning, reduction, partitioning, scaling and transformation. The missing values have been dropped, after data cleaning, the categorical variables are transformed into numerical vectors. The deep learning algorithms which use weighted sum of inputs get biased by variables with large values and the small values are ignored.

To prevent this numerical data is scaled to standard range. However, standardization becomes biased if input variables contain outliers, which can be seen in the training dataset (Figure 9).

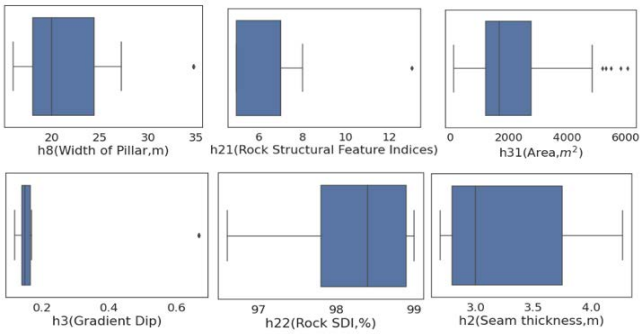


Figure 9. Presence of outliers and skewed data.

Hence, Robust Scaler is applied which scales the data on the basis of median (50th percentile) and inter-quartile range (difference between 75th and 25th percentile), given by:

$$Scaled\ value = \frac{value - median}{P_{75} - P_{25}}$$

The resulting values (Table 3) are not skewed by the outliers however they are still present with same relative relationships (Brownlee, 2020). The data is partitioned in the ratio of 70% for training and 30% for validation and testing. To prevent any data leakage, Robust scale is fit only on the training data, it then transforms the training, validation and test data.

4.3 Model Training

The dataset is randomly shuffled. Various hyper parameters that are tweaked and tuned during the training process are:

- Hyper parameters pertaining to network architecture: Different numbers of first layer neurons, hidden layer

neurons, the number of hidden layers and dropout ratios are tried out during training the model.

- Hyper parameters necessary for model training: Different activation functions like Rectified Linear unit (ReLu), Exponential Linear unit (Elu) and Tanh (Hyperbolic Tan) are used in the hidden layers, whereas sigmoid activation is used for the output layer of the binary classifier model. A variety of optimizers like Stochastic Gradient Descent (SGD), Root Mean Square propagation (RMSprop), Adaptive Moment Estimation (Adam), Nesterov accelerated Adaptive Moment Estimation (Nadam) at different learning rates are experimented to train over the dataset in batches. Callbacks and dropouts are used to prevent over fitting over the training dataset.

5. Result and Analysis

5.1 Hyper Parameter Space in the Scan (Table 4)

Table 4. Initial space and optimal set of hyperparameters (Hp)

Hp	Initial space	Model 1 (No fall (0) Fall (1))	Model 2 (Local fall (0) Main fall (1))
First layer neurons	33,144,256	256	256
Hidden layer neurons	72,144, 256	256	144
Hidden layers	0,1,2	2	1
Activation	ReLu, Elu, Tanh	Relu	Elu
Batch size	16, 32, 64	32	32
Dropout	0.2,0.45,0.5	0.5	0.2
Epochs	200,250,300	300	300
Learning rate	0.007,0.01, 0.04	0.01	0.007
Kernel Initializer	Uniform, Normal	Normal	Normal
Optimizer	SGD, RMSprop, Adam, Nadam	RMSprop	Nadam
Shape	Long funnel, brick	Long funnel	-

The box plots (Figure 10) provides a detailed comparison of all the hyper parameters, a useful tool to assess their interplay and to decide the next set of hyper parameter boundary. SGD does not perform well and dropout of 0.5 provides better results.

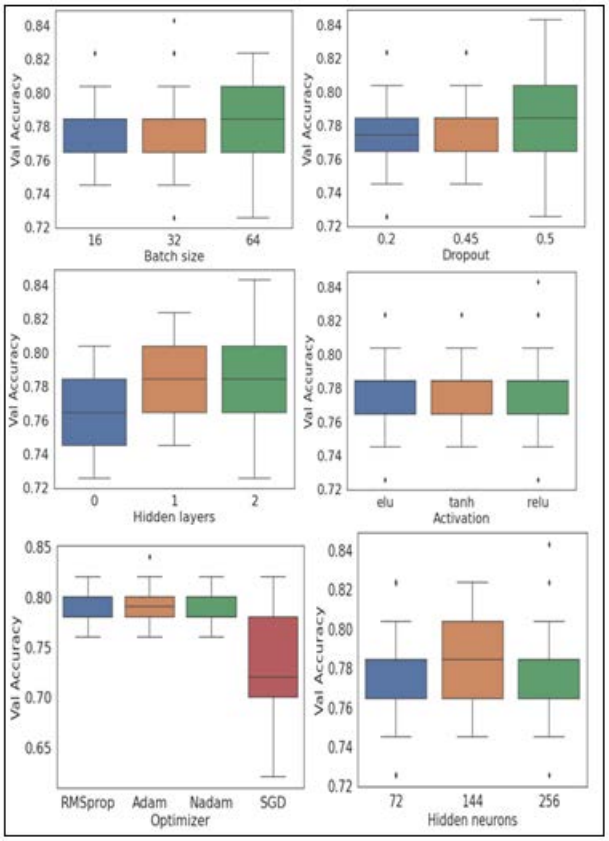


Figure 10. Comparison of hyperparameters.

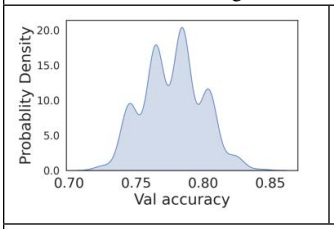
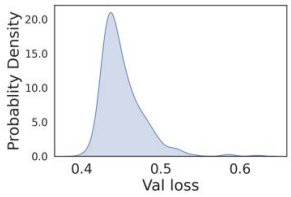
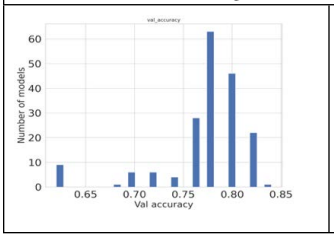
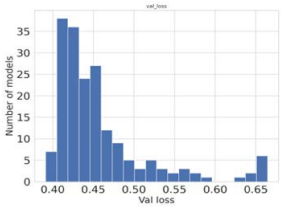
5.2 Validation Accuracy and Loss

The accuracy of most of the models formed by different combinations of hyperparameters lie in range of 0.75 – 0.85 and loss lies in the range of 0.4–0.5 for both the type of Models: 1 and 2 (Table 5). The loss function used for binary classification problems is the binary cross entropy/ Log loss (Godoy, 2018), it compares the predicted probabilities $(p(y_i))$ to actual output, given by:

$$Logloss = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log\left(p\left(y_i\right)\right) + \left(1-y_i\right)\cdot \log\left(1-p\left(y_i\right)\right)$$

When $y_i = 1 \Rightarrow \log\left(p\left(y_i\right)\right)$ is added to the loss and when $y_i = 0 \Rightarrow \log\left(1-p\left(y_i\right)\right)$ is added to the loss.

Table 5. Validation accuracy and loss of models

Validation accuracy	Validation Loss
Models formed during search for Model 1	
	
Models formed during search for Model 2	
	

5.3 Performance of Selected Models with Optimal Hyperparameters Over Test Data

5.3.1 Confusion Matrix

A confusion matrix (Table 6) depicts the model classification accuracy (Kalantar *et al.*, 2018) and performance. Accuracy of a model is given by:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

Precision is defined as "the probability that an object is relevant given that it is returned by the system", recall is "the probability that a relevant object is returned" and f1 score is the harmonic average of the above two (Goutte and Gaussier, 2005) and is given by:

$$Precision(p) = \frac{TP}{TP + FP}\ ,\ Recall = \frac{TP}{TP + FN}$$

$$F1\,Score = 2 \bullet \frac{Precision \bullet Recall}{Precision + Recall}$$

The confusion matrix of the models (Table 7) shows that the deep neural network (DNN) model 1 correctly identifies 28 cases of "fall" and misclassifies 5 cases as "no fall" whereas it correctly classifies 18 out of 20 cases of "No fall" category. The DNN model 2 correctly classifies 14 cases of "Local fall" and 16 cases of "Main fall" and misclassifies 1 case of each category. Thus, the classification ability of the models (Table 8) shows that Model 1 has an accuracy of 0.87 whereas Model 2 has accuracy of 0.94 over test data.

Table 6. Confusion matrix

Predicted Class / Fall No Fall	True Class No Fall Fall	
	True Positive (TP)	False Positive (FP)
	False Negative (FN)	True Negative (TN)

Table 7. Confusion matrix of DNN models

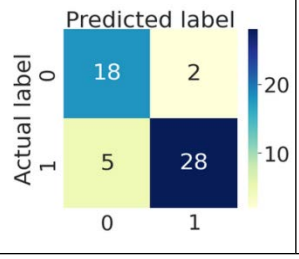
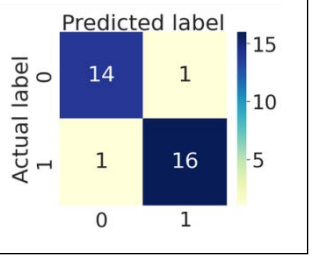
Model 1 (No fall(0)/Fall(1)	Model 2 (Local fall(0)/ Main fall(1)
	

Table 8. Classification Report of the Models

Model 1 (No fall(0)/ Fall(1))

	precision	recall	f1-score	support
0	0.78	0.90	0.84	20
1	0.93	0.85	0.89	33
accuracy			0.87	53
macro avg	0.86	0.87	0.86	53
weighted avg	0.88	0.87	0.87	53

Model 2 (Local fall(0)/ Main fall(1))

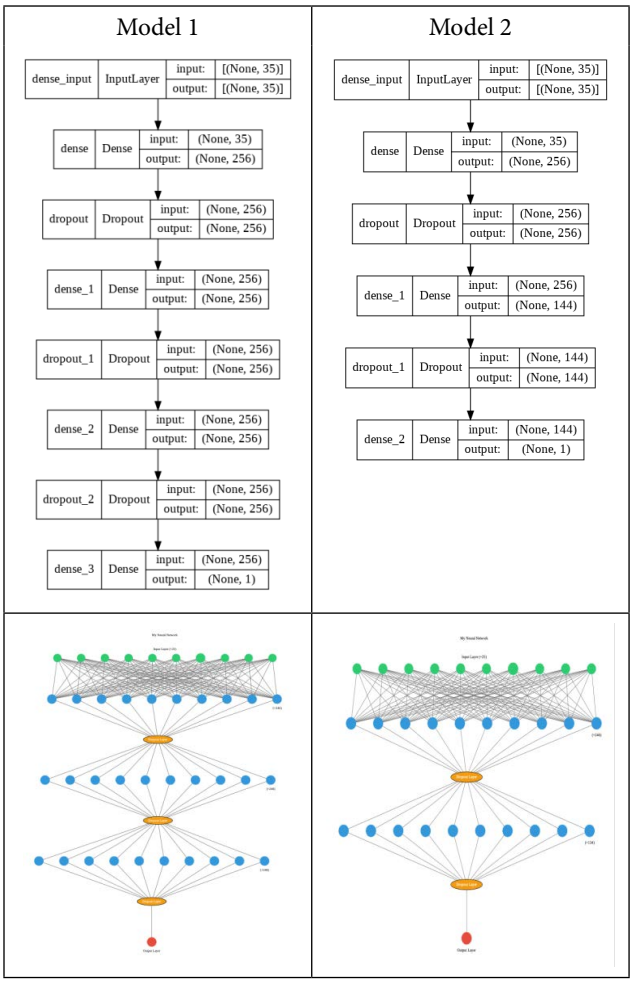
	precision	recall	f1-score	support
0	0.93	0.93	0.93	15
1	0.94	0.94	0.94	17
accuracy			0.94	32
macro avg	0.94	0.94	0.94	32
weighted avg	0.94	0.94	0.94	32

5.3.2 ROC Curves

In statistics, the receiver operating characteristics (ROC) curve, is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is changed. The curve is plotting the True positive rate (specificity) against false positive rate (1-Sensitivity). A

point in ROC space with coordinate (0, 1) represents the model with best prediction ability. At the default threshold, ROC curves and the area under the curve (AUC) of Model 1 is 0.874 (Figure 11) while that of Model 2 is 0.937 (Figure 12). The larger the area under the curve the better the classification ability of the model (Bradley, 1997).

Table 9. Model architecture



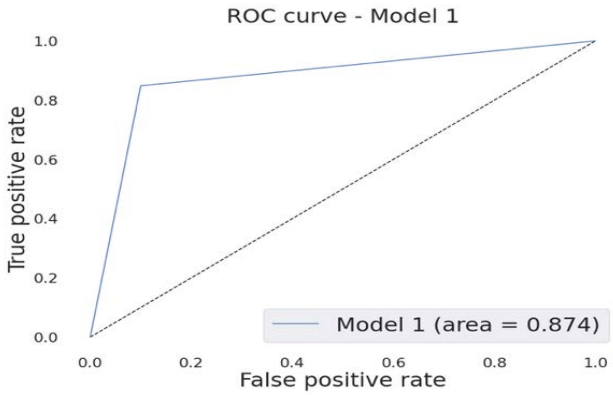


Figure 6. ROC and AUC of Model 1.

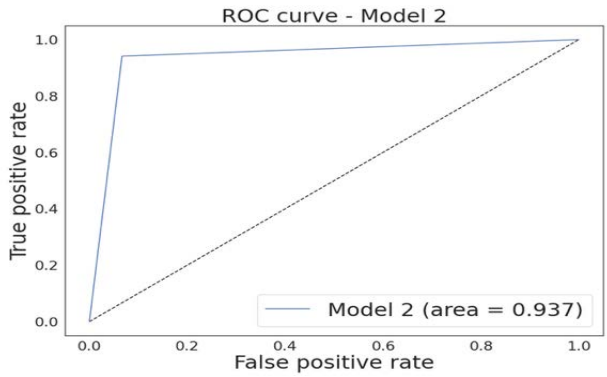


Figure 7. ROC and AUC of Model 2.

The structure of the models created can be visualized as (Table 9):

6. Conclusions

This study has been performed over two mining blocks of East Bokaro and South Karanpura coalfields of India. Analysis of each parameter from both the datasets, depicts that the occurrence of fall is highly correlated to the area of fall. The model forecasts the occurrence of fall and the sequence of local and main fall using extraction areas and other parameters. For field application, models are run with different anticipated extraction areas with other parameters to predict fall and to classify local or main fall. The set of such anticipated classified areas of local and main fall by the models are sequence of falls and used as an indicator for taking advance precautions against local and main fall in caving panel for safety of men and machineries. This study also tries to depict the importance of hyper parameter optimization while developing a neural network model, as a correct choice of hyper parameter results in a significant increase in accuracy over the same dataset (Figure 10). The model classifies local and main fall with substantial accuracy and precision, however model's ability to predict the occurrence of a fall can still be improved and can be worked upon.

7. Acknowledgement

Authors are thankful to the management of Central Coalfield Ltd., for extending their supports with the necessary information required for this paper. The views expressed in the paper are that of the authors and not necessarily of the organizations they belong to.

8. References

1. Simon, H.A. (1957). Models of Man: Social and Rational. John Wiley and Sons, Inc., 1957.
2. Malkowski, P., & Juszynski, D. (2021). Roof fall hazard assessment with the use of artificial neural network. International Journal of Rock Mechanics & Mining Sciences 143(2021): 104701. https://doi.org/10.1016/j.ijrmms.2021.104701
3. Isleyen, E., Duzgun, S., & Carter, R. (2021). Interpretable deep learning for roof fall hazard detection in underground mines. Journal of Rock Mechanics and Geotechnical Engineering, 13(6): 1246–1255. https://doi.org/10.1016/j.jrmge.2021.09.005
4. Razani, M.,Yazdani-chamzini, A., & Yakhchali, S. (2013). A novel fuzzy inference system for predicting roof fall rate in underground coal mines. Safety Science, 55: 26–33. https://doi.org/10.1016/j.ssci.2012.11.008
5. Deb, D., Kumar, A., & Rosha, R. (2006). Forecasting shield pressures at a longwall face using artificial neural networks. Geotechnical and Geological Engineering, 24: 1021–1037. https://doi.org/10.1007/s10706-005-4430-6
6. Monjezi, M., Hesami, S., & Khandelwal, M. (2009). Superiority of neural networks for pillar stress prediction in bord and pillar method. Arabian Journal of Geosciences, 4: 845–853. https://doi.org/10.1007/s12517-009-0101-x
7. Sheorey, P.R. (1984). Use of rock classification to estimate roof caving span in oblong workings. International Journal of Mining and Mineral Engineering, 2: 133–140. https://doi.org/10.1007/BF00880878
8. Singh, R., Singh, T.N., & Dhar, B.B. (1996). Coal pillar loading in shallow mining conditions. International Journal of Rock Mechanics and Mining Sciences & Geomechanics, 33, No.8: 757-768.
9. Sheorey P.R. (1994). A Theory for in situ stresses in isotropic and transversely isotropic. International Journal of Rock Mechanics and Mining Sciences & Geomechanics, 31(1): 23–34. https://doi.org/10.1016/0148-9062(94)92312-4
10. Jena, S., Prasad, K., Lokhande, R.D., & Pradhan, M. (2016). Analysis of strata control monitoring in underground coal mine for apprehension of strata movement. Recent Advances in Rock Engineering (RARE 2016), pp. 505–511. https://doi.org/10.2991/rare-16.2016.81
11. Obert, L., & Duvall, W. (1967). Rock mechanics and the design of structures in rock. New York: John Wiley and Sons, Inc.; 1967.
12. Majumdar, S. (1986). The support requirement at a longwall face — bending moment approach. In: Proceedings of 27th US Symposium on Rock Mechanics: Key to Energy Production (The University of Alabama, Tuscaloosa, Alabama); 1986: 325–332.
13. Pawlowicz, K. (1967). Classification of rock cavability of coal measure strata in upper Silesia coalfield. Prace GIG, Komunikat, No. 429, Katowice. (in Polish).
14. Peng, S.S., & Chiang, H.S. (1984). Longwall mining. In New York: John Wiley and Sons, Inc.
15. Ghose, A.K., & Dutta, D. (1987). A rock mass classification model for caving roofs. International Journal of Mining and Geological Engineering, 5: 257–271. https://doi.org/10.1007/BF01560777
16. Sarkar, S.K. (1998). Mechanized longwall mining — The Indian experiences. New Delhi: Oxford and IBH Publishing Company Private Limited.
17. Sarkar, S.K., & Dhar, B.B. (1993). Strata control failures at caved longwall faces in India — experience from Rana to Churcha (1964 to 1990). In Proceedings of the 4th Asian Mining (Organized by MGMI at Calcutta), 361–380.
18. Nimaje, D.S., & Sai, S. (2015). Development of software to evaluate roof fall risk in bord and pillar method — Depillaring Phase. GeoScience Engineering, LXI(2): 14–22. https://doi.org/10.1515/gse-2015-0014
19. McCulloch, W., & Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. The Bulletin of Math. Biophys., 5: 115–133. https://doi.org/10.1007/BF02478259
20. Fausett, L.V. (1993). Fundamentals of Neural Networks:Architectures, Algorithms and Applications. (1st ed). Pearson publication, India.
21. Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational capabilities. In Proceedings of National Academy of Sciences, (USA), 79: 2554–2558. https://doi.org/10.1073/pnas.79.8.2554. PMid:6953413. PMCid:PMC346238
22. Hopfield, J.J. (1984). Neurons with graded responses have collective computational properties like those of two-state neurons. In Proceedings of National Academy of Sciences (USA), 81: 3088–3092. https://doi.org/10.1073/pnas.81.10.3088. PMid:6587342. PMCid:PMC345226
23. Lee, S., Ryu, J., Lee, M., & Won, J. (2003). Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea. Environmental Geology, 44(7): 820–833. https://doi.org/10.1007/s00254-003-0825-y
24. Russell, S.J., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (Third ed). Pearson India Education Services Pvt. Ltd.
25. Haykin, S. (1999). Neural Networks:A Comprehensive Foundation. (2nd Ed.).
26. Nazzal, J., El-Emary, I., & Najim, S. (2008). Multilayer Perceptron Neural Network (MLPs) for analyzing the properties of Jordan Oil Shale. World Applied Sciences Journal, 5(5): 546–552.
27. Chollet, F., & Others. (2015). Keras. GitHub.

28. Van Rossum, G., & Drake, F.L. (2009). Python 3 Reference Manual.
29. Bisong, E. (2019). Google colaboratory. In: Building machine learning and deep learning models on Google cloud platform. (Apress Berkeley CA., pp. 59–64). https://doi.org/10.1007/978-1-4842-4470-8_7
30. Talos, A. (2019). Autonomio Talos [Computer Software] hyperparameter optimization for tensorflow, Keras and Pytorch.
31. Parashar, A., & Sonker, A. (2019). Application of hyperparameter optimized deep learning neural network for classification of air quality data. International Journal of Scientific & Technology Research, 8(11): 1435–1443.
32. Ghasemi, E., Ataei, M., Shahriar, K., Sereshki, F., & Esmaeil, S.(2012). Assessment of roof fall risk during retreat mining in room and pillar coal mines. International Journal of Rock Mechanics and Mining Science, 54: 80–89. https://doi.org/10.1016/j.ijrmms.2012.05.025
33. Kumar, A., Kumar, D., Singh, A.K., Ram, S., Kumar, R., Gautam, A., Singh, R., & Singh, A.K. (2019). Roof sagging limit in an early warning system for safe coal pillar extraction. International Journal of Rock Mechanics and Mining Sciences, 123: 104–131. https://doi.org/10.1016/j.ijrmms.2019.104131
34. Mark, C., & Michael, G. (2017). Preventing roof fall fatalities during pillar recovery: A ground control success story. International Journal of Mining Science and Technology, 27(1): 107–113. https://doi.org/10.1016/j.ijmst.2016.09.030
35. Mark, C., & Molinda, G. (2007). Development and application of the Coal Mine Roof Rating (CMRR). Proceedings of the International Workshop on Rock Mass Classification in Underground Mining, 95–110.
36. Palei, S.K., & Das, S.K. (2009). Logistic regression model for prediction of roof fall risks in bord and pillar workings in coal mines:An approach. Safety Science, 47: 88–96. https://doi.org/10.1016/j.ssci.2008.01.002
37. Torres, J. (2018). First contact with Deep Learning, Practical introduction with Keras.
38. Fan, C., Chen, M., Wang, X., Wang, J. & Huang, B. (2021). A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. Sustainable Energy Systems and Policies, Frontiers in Energy Research, 9: 18. https://doi.org/10.3389/fenrg.2021.652801
39. Brownlee, J. (2020). Data preparation for machine learning:Data cleaning, feature selection and data transforms in Python.
40. Godoy, D. (2018). Understanding binary cross-entropy/ log loss: A visual explanation.
41. Kalantar, B., Pradhan, B., Naghibi, S.A., Motevalli, A., & Mansor, S. (2018). Assessment of the effects of training data selection on the landslide susceptibility mapping: A comparison between Support Vector Machine (SVM), Logistic Regression (LR) and Artificial Neural Networks (ANN). Geomatics, Natural Hazards and Risk, 9(1): 49–69. https://doi.org/10.1080/19475705.2017.1407368
42. Goutte, C., & Gaussier, E.(2005). A probabilistic interpretation of precision, recall and F -score with implication for evaluation. Lecture Notes in Computer Science, 3408: 345–359. https://doi.org/10.1007/978-3-540-31865-1_25
43. Bradley, A.E. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition, 30(7): 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2