

METHODOLOGY

3.1 Introduction

In this study (case 1) the datasets which has been used to perform the prediction performance of the pillar stability previously compiled by (Lunder, 1997). The database includes 178 hard rock pillar examples from several mining locations, including the Elliot Lake uranium mines in Canada, the Mount Isa mines in Australia, the Selebi-Phikwe mines in Southern Africa, the Black Angel mine in Greenland, open stope mines in Canada, and the Zinkgruvan mine in Sweden. The database information described by Wattimena (2014) is included in Annexure 1. A summary of data has been indicated in Table 3.1. The w/h ratio and the proportionality of average induced pillar load over the intact rock's UCS (PL/UCS) are the two input parameters illustrated and also an output parameter, the pillar stability condition (PS) are used for the model prediction. Based on the definitions in Table 3.2, the pillar stability condition has been divided into three categories: stable cases, unstable cases, and failed cases. Figure 3.1 depicts the database, which comprises 60 stable cases, 50 unstable cases, and 68 failed cases. The data reduction process seeks to identify the reduced dataset, which is a subset of the original dataset, in order to maximize the performance criterion of the machine learning algorithm. 80:20 split has been taken as training and testing on the original and reduced datasets.

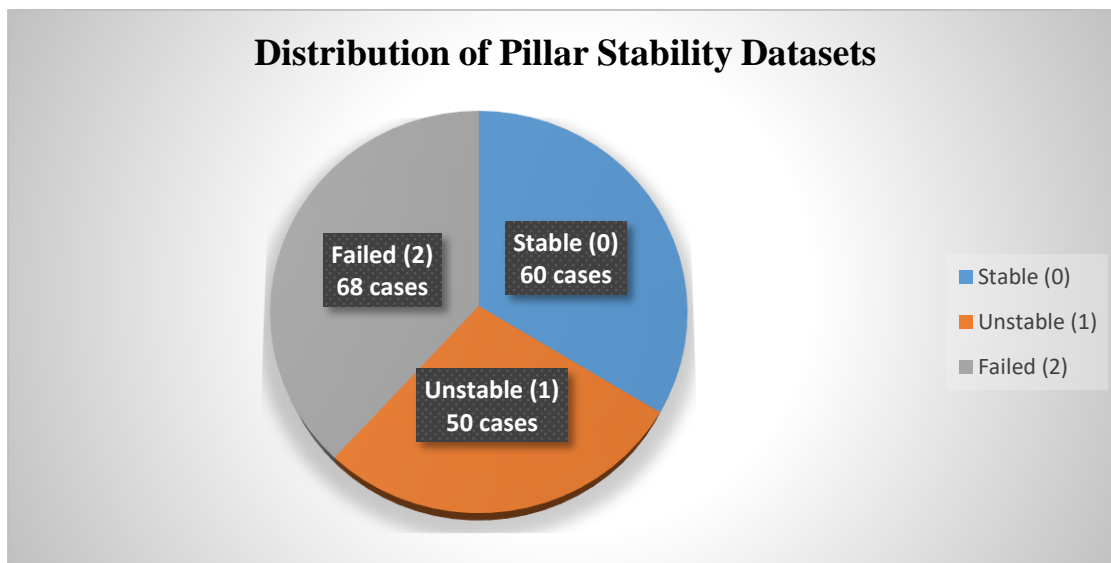


Figure 3.1 Distribution of Pillar Stability Datasets

Table 3.1: Descriptive characteristics of datasets.

Data Type	Parameters	Symbols used	Values in range
Input Parameters	Pillar width to Pillar height ratio	w/h	0.31-4.5
	Pillar load to Uniaxial compressive strength	PL/UCS	0.11-0.67
Output Parameters	Pillar stability condition	PS	Stable cases (0) Unstable cases (1) Failed cases (2)

Table 3.2. Representation of pillar stability condition (Wattimena, 2014).

Pillar stability conditions	Observed pillar condition
Stable cases	There is little spalling no sign of joint opening (Von Kimmelman et al., 1984).
Unstable cases	There is no evidence of stress-induced fracture (Lunder, 1994). Cracking and spalling in the rib pillar's development and rises; the measurable amount of noise in the pillar; distorted drill holes; excess sludge being obtained from stopes; pillar busting; serious lateral displacement within the pillar (Potvin et al. 1989).

	<p>spalling that is noticeable (Von Kimmelman et al. 1984).</p> <p>Fractures can also be found in the pillar's center (Krauland and Soder 1987).</p> <p>Only corner breaking up to fracturing in pillar walls with fracture apertures up to 10 mm is permitted (Lunder 1994).</p> <p>Failed in part (Hedley and Grant 1972).</p> <p>Slobbering (Hudyma 1988).</p>
Failed cases	<p>Excessive spalling, significant joint opening, and drill hole deformity (Von Kimmelman et al. 1984).</p> <p>The pillar disintegrates; stones fall out; cracks through the pillar with fracture holes more than 10 mm (Lunder 1994).</p>

The study of pillar stability prediction has been performed on the Waikato Environment for Knowledge Analysis (WEKA) tool, the obtained models are developed using the PART, JRIP, SMO, J48, and RF algorithms, which are all dependent on training data. WEKA was used to train the adopted models using the training dataset in ARFF (attribute relation file format) format. Following that, a testing dataset based on the specified training dataset is used to test the models. The flow chart of the research methodology and pre-processing of data are given in Figure 3.2 and 3.3 respectively.

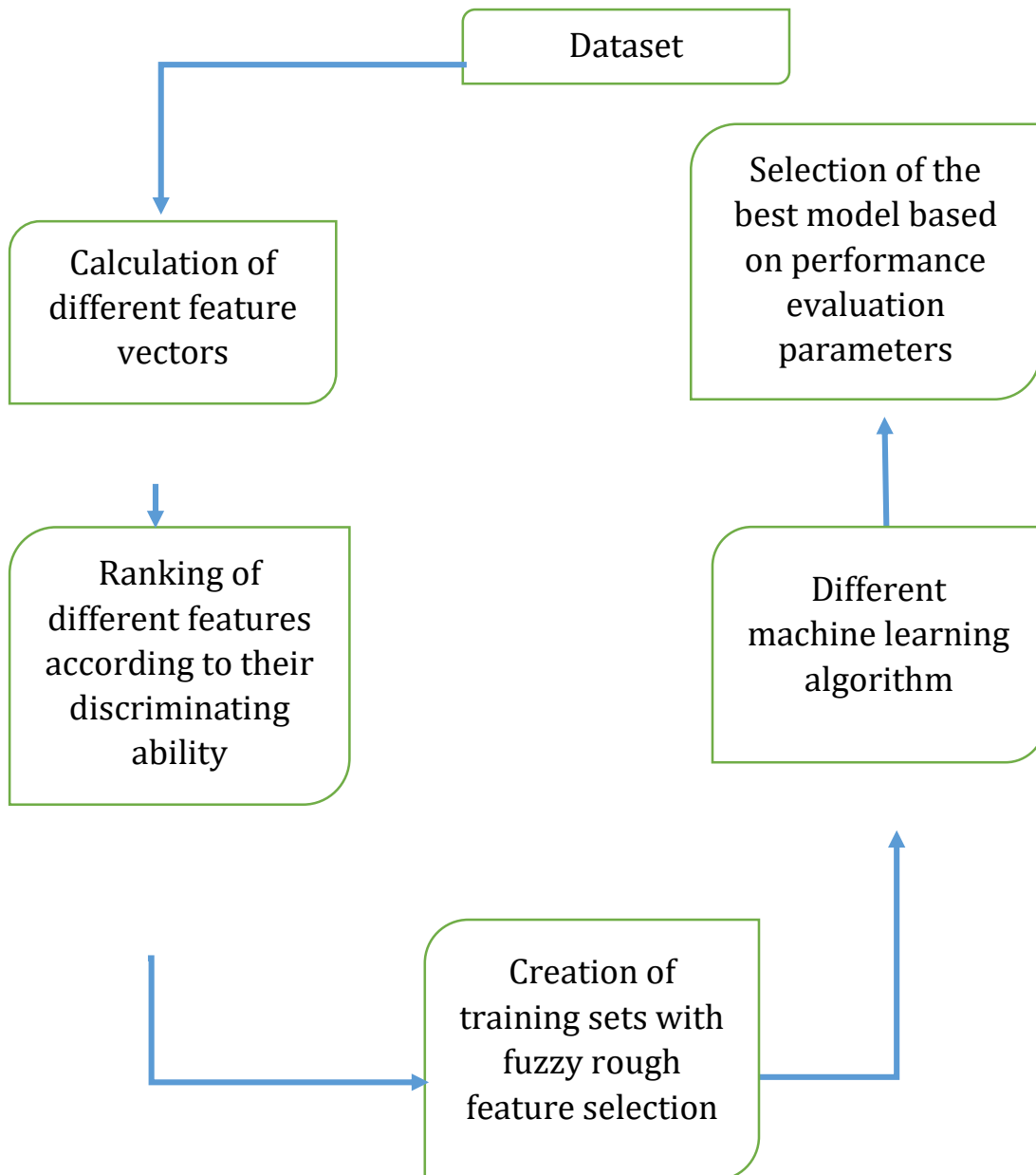


Figure 3.2: Flowchart of methodology for machine learning (Kumar et al.,2019).

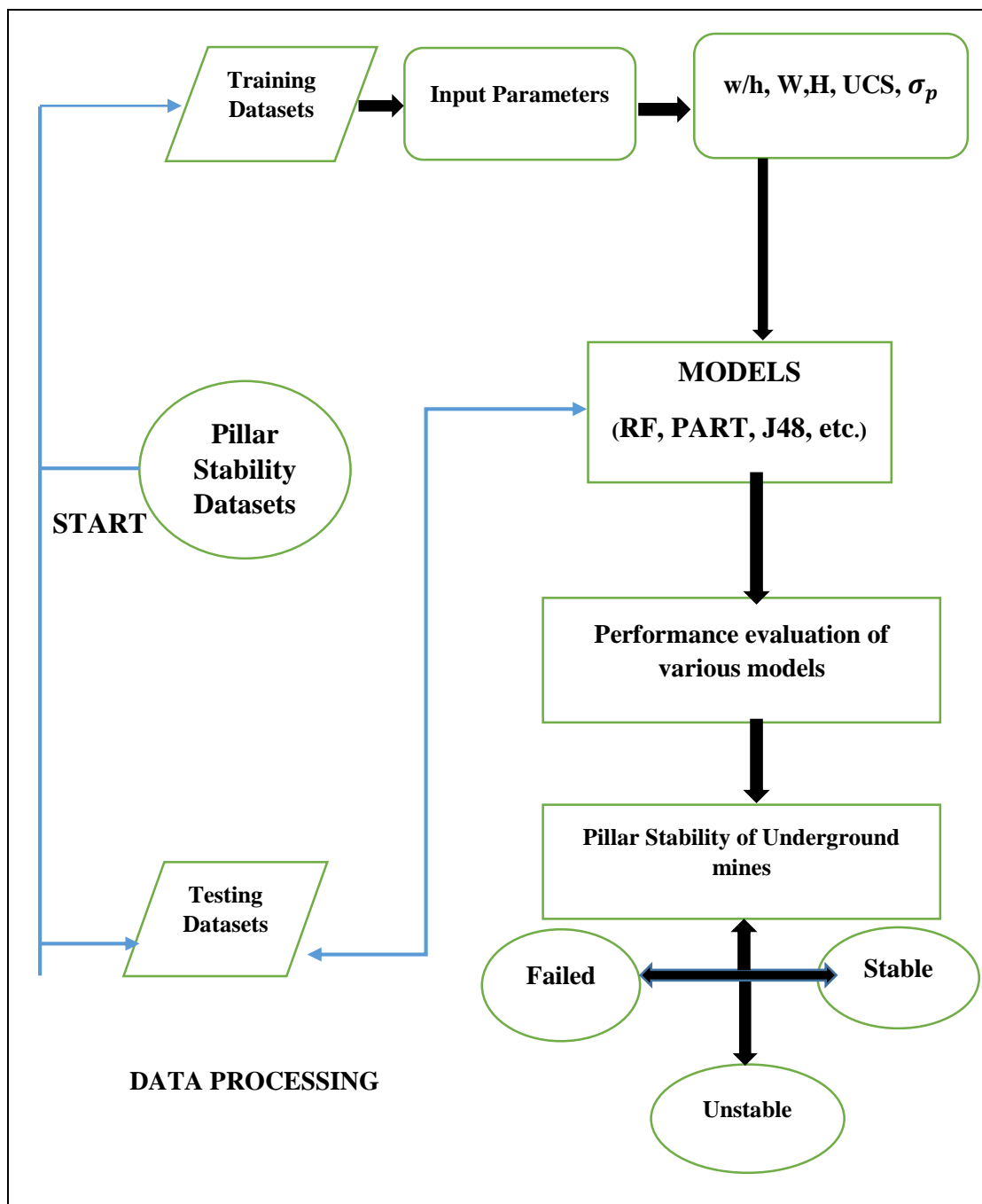


Figure 3.3. Data processing steps performed in WEKA tool.

Some pillar parameters should be employed to build classification models. Meanwhile, three parameters-selection principles must be followed. To begin, the discriminant indicators should be sensitive and stable metrics indicating pillar stability features. Second, the parameters should be physically separate from one another.

Methodology

Finally, the parameter data should be conveniently accessible or accessible. Many elements influence a pillar's stability, with the following being the most important:

- 1) Load on the pillar,
- 2) The dimensions of the pillar,
- 3) The geological arrangement and the ore body's own strength.

In general, pillar width, pillar height, the w/h ratio and shape of the pillars can all indicate the dimensions of the pillar. The w/h ratio, in particular, influences the possibility for catastrophic failure of pillars and is the key reference for estimating pillar strength in underground mines.

When compared to traditional safety factors and numerical simulation approaches, the Machine Learning methodology can uncover hidden correlations between variables and handle nonlinear situations effectively. As a result, it appears to be a viable strategy for determining pillar stability.

3.2 Feature Ranking

Due to the importance of features, the feature ranking algorithm: fuzzy rough attribute evaluator, was carried out to obtain the rank of the features that participated in the classification task. Then, experiments were conducted by using various classification algorithms by changing the number of features in the order of most significant to least significant. After performing the results of these experiments, it was possible to determine the redundant and irrelevant features and further remove them from the input feature set. This concept was highly efficient when input features were in large numbers.

3.2.1 Fuzzy-rough set-based feature selection technique

There are two key concepts in fuzzy rough sets: indiscernibility for rough sets and vagueness for fuzzy sets. Because of the uncertainty in various domains, the two concepts are established. Fuzzy set data always lacks precise boundaries. Rough sets can show uncertainty due to a lack of information by using approximations. In the current study, a fuzzy rough set-based feature selection approach was used to compute the reduct set, which may be expressed as follows (Jenson and Shen, 2008):

Expression of Algorithm: Fuzzy rough feature selection algorithm

Input: **1**, the set of all conditional attributes;

Output: **2**, the set of decision attributes;

$P \leftarrow \{ \}$

Do

$Q \leftarrow P$

foreach $x \in (1 - P)$

if $\gamma_{P \cup \{x\}}(2) > \gamma_Q(2)$

$T \leftarrow R \cup \{x\}$

$P \leftarrow Q$

until $\gamma_P(2) == \gamma_C(2)$

return P

3.3 Classification protocol

The experiments were conducted using different classifiers, namely Naïve Bayes, PART, Jrip, SMO, and Random Forest. A brief description of this algorithm is given as follows:

Naive Bayes is a technique for building classifiers from class labels of class labels that are obtained from a finite collection. There is not a single training algorithm for Naive Bayes classifiers. Instead, they are taught in accordance with a guiding

Methodology

principle. Given the class variable, all Naive Bayes classifiers assume that the value of a specific feature is self-governing of the importance of any other feature. Naive Bayes classifiers may be taught reasonably efficiently in a supervised learning environment for some probability models some types of probability models. Naive Bayes classifiers may be trained reasonably efficiently in a supervised learning environment. The maximum likelihood approach of maximum likelihood is used to estimate parameters for Naive Bayes models in many practical situations; in other words, the naive Bayes model may be employed without using Bayesian probability or any Bayesian methods. Despite their rudimentary assumptions and naive design, Naive Bayes classifiers have done effectively in a range of tough real-world circumstances. For some given probability models, Naive Bayes classifiers can be learned relatively efficiently in a SL situation.

PART algorithm (Classification Algorithm Based on Rules): PART is an abbreviation for Projective Adaptive Resonance Theory (Pinto and Mahado, 2014). PART is a refined method of generating rules (Chauhan et al. 2013). PART is a rule-following rule-breaker who divides and conquers. The algorithm produces "decision lists," which are pre-defined collections of rules. Each rule in the list is checked against new data, and the item is assigned the class of the first matching rule. The best tree is picked after rule creation, and its leaves are converted into rules. This algorithm supports all types of classes, including binary and nominal classes and all types of attributes.

Jrip values dataset instances in ascending order and generates a set of rules for each risk dataset. It uses sequential covering techniques to generate ordered rule lists and implements the Repeated Incremental Pruning to Produce Error Reduction

(RIPPER) propositional rule, learner. The Jrip (RIPPER) algorithm analyses each dataset in a given database and creates a set of rules including all of the class's properties. The next class is assessed in the same manner as the preceding class. This process is repeated until all of the classes have been covered (Bhargav et al. 2016).

SMO (sequential minimum optimization) is a solution approach for the quadratic programming (QP) problem that emerges during support vector machine training (SVM). SMO is a technique that is often employed in the training of SVM. SMO is a straightforward method for dealing with the SVM QP problem quickly and without the need for any extra matrix storage or numerical QP optimization phases. SMO is used to breakdown the overall QP problem into QP sub-problems. The SMO algorithm's release in 1998 sparked a lot of enthusiasm in the SVM community, as earlier approaches for SVM training were far more difficult and time-consuming (Rifkin, 2002).

Random Forest (RF): RF is a supervised learning algorithm. Random forests are a set of tree predictors in which each tree is dependent on the values of a single predictor. For all trees in the forest, a random vector was sampled separately and with the same distribution. This algorithm was first created by Brieman (2001). Vital improvements in portrayal accuracy has been achieved on account of growing an outfit of trees and allowing them to cast a ballot for the most notorious class. In order to build up these collections, often-unpredictable vectors are created that administer the improvement of each tree in the gathering.

Each tree is constructed according to the following procedure: (a) Accept the number of cases N in the preparation dataset randomly. (b) At each center point, m number of features are picked randomly out of everything being equivalent, and the

Methodology

advanced split subject to these ‘m’ features is used to part the center point. (c) Each tree should be created as tremendous as possible without trimming. By building up all the tree Resulting of building up all trees, another article would then have the option to be classified as the class name with the most votes, where each vote is picked by each tree in the forest and the majority votes will have the prediction class. Depending on the size and type of the training set, a few hundred to several thousand trees are typically employed. Formation of prediction class has been presented in a given below in figure 3.4.

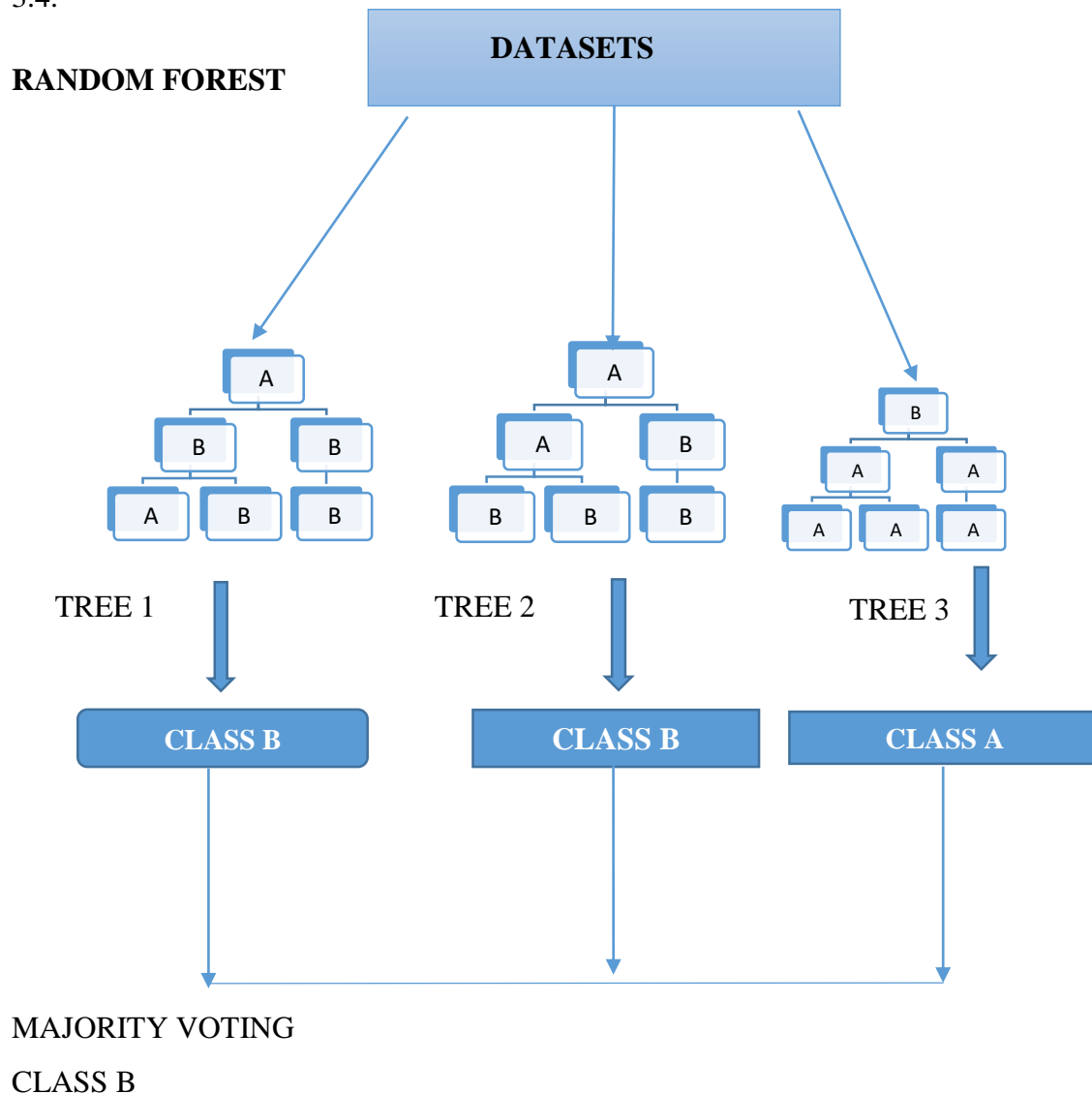


Figure 3.4 Representation of Random forest picking up majority class.

The RF algorithm is expressed as:

```
for i ← 1 to B do  
    Using the training data, create a bootstrap sample of size N;  
    while node size != minimum node size do  
        Choose a subset of m predictor variables at random from the total p.  
        for j ← 1 to m do  
            if jth The predictor optimizes the splitting condition then  
                splitting of the internal node into two child nodes;  
            break;  
        End  
    End  
End  
end  
return the ensemble tree comprising all B subtrees created in the for loop's  
outer loop;
```

Expression for RF Algorithm.

3.4 Performance Evaluation Metrics:

The five machine learning methods' relative prediction accuracy is assessed using threshold-dependent and threshold-independent metrics. True Positives (TP), which is the number of correctly predicted pillar stability, False Negatives (FN), which is the number of incorrectly predicted pillar stability, True Negatives (TN), which is the number of correctly predicted pillar un-stability with failure, and False Positives (FP), which is the number of incorrectly predicted pillar un-stability with failure.

a) Accuracy:

The percentage calculation of the correctly predicted pillar's stability and un-stability with failure is presented as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100.$$

b) Area Under Curve (AUC):

It denotes the area under the Receiver Operating Characteristic curve (ROC), and the closer it is to one, the better the predictor. It is one of the assessment factors that are resistant to the skewed nature of the datasets.

c) MCC: Mathew's correlation coefficient was estimated by using the following equation:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

It is a prominent performance metric for binary classifications. A MCC value of one is seen to be the strongest predictor. The experiments were carried out using the open-source Java-based machine-learning platform WEKA 3.8.

d) Receiver Operating Curve (ROC)

It's a technique that is useful for forecasting the likelihood of a binary result. It is a plot of the false positive rate (x-axis) vs the true positive rate (y-axis) for a variety of threshold values ranging from 0.0 to 1.0.

ROC curve was used to perform a visual representation of the classifiers. It is one of the best ways to estimate the overall performance of different classifiers at different decision thresholds. The development of the ROC model has been presented in figure 3.5.

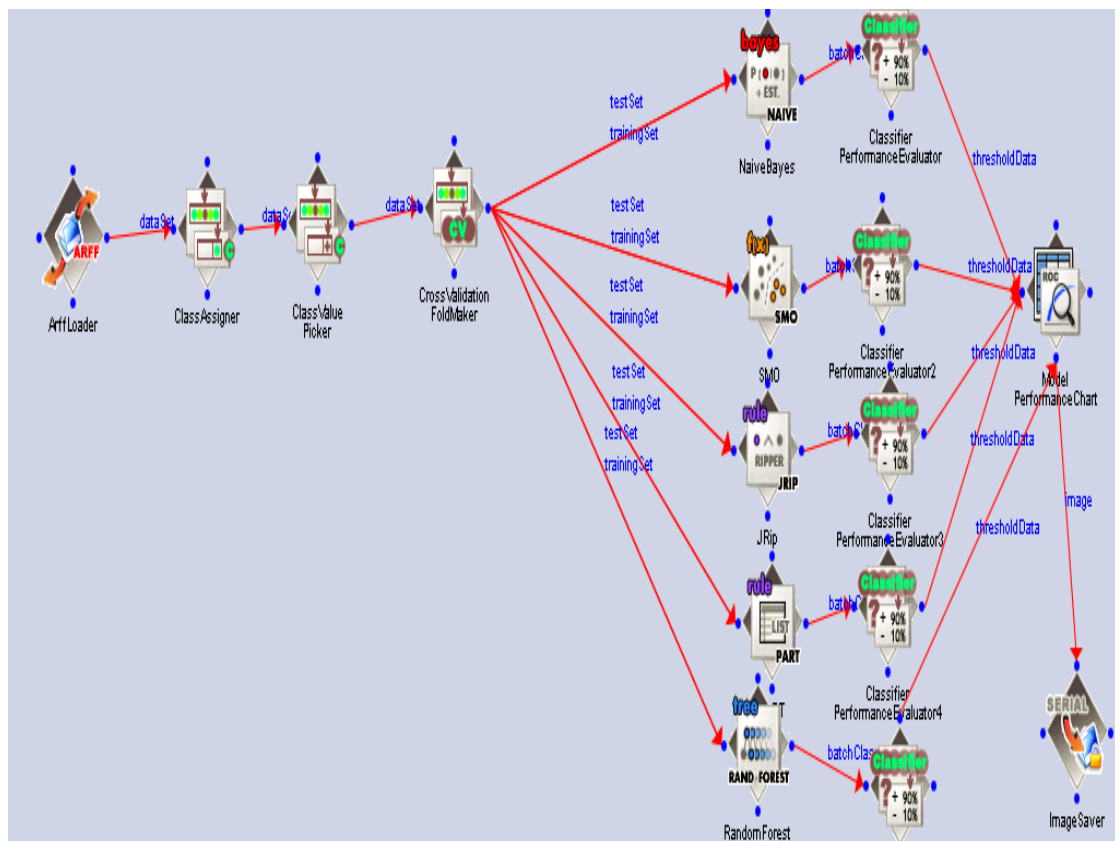


Figure 3.5: Development of ROC model

3.5 Theoretical estimation of Pillar strength

The pillar's strength reflects the consistency of the materials and, as a result, the distribution of stresses within the pillar; consequently, it is critical to consider the variable that influences the pillar's stress distribution. The vertical and horizontal components of the stress vector at the point where the pillar meets the surrounding strata are likely to be the most relevant (Martin and Maybee, 2000). The following is a general equation for determining pillar strength (PS): Equation 3.1

$$PS = k \cdot W^\alpha \cdot h^\beta \quad (3.1)$$

Where W is Pillar width, h is the pillar height, α , β are constants, and k are constant represents the strength of coal material.

The strength of pillars in Table 1 was estimated by the equation proposed by (Sheorey, 1992) as Equation 3.2:

$$PS = 0.27\sigma_c h^{-0.36} + \left(\frac{D}{250} + 1\right)\left(\frac{W}{H} - 1\right) \quad (3.2)$$

Where

PS	Pillar strength in (MPa)
h	Pillar height (m)
σ_c	Uniaxial compressive strength in (MPa)
W	Pillar width in (m)
D	Depth of cover in (m)

3.6 Theoretical estimation of Pillar stress

Estimating a pillar's stress is one of the most difficult tasks in pillar design. When necessary, pillar layouts are made up of pillars that are evenly placed and of roughly the same size. To determine the pillar stress, the tributary area method will be used. The whole load of the overburden is evenly distributed across all pillars in this approach. Despite the fact that it denies the fact that the amount of load that may be transferred to barrier pillars or solid abutment is dependent on the stiffness of the pillars and the surrounding overburden, it is a valid and well-developed approach for

determining proper stable pillar design. (Madden,1991). The overburden of coal-bearing strata is assumed to be 0.025MN/m³, and the pillar load is approximated by Equation 3.3.

$$PL = 0.025D(W + B)^2W^2 \quad (3.3)$$

Where

PL Pillar stress (pillar load) in (MPa)

B Gallery width in (m).

W Pillar width in (m)

D Depth of cover in (m).

3.7 Principal Component Analysis (PCA)

The principal component analysis approach is used to reduce the dimensionality of such datasets, boosting interpretability while minimizing information loss. It accomplishes this by generating new uncorrelated variables that gradually optimize variance. Examine the size and direction of the coefficients for the original variables to interpret each primary component. The PCA statistical approach was utilized to identify the significant parameters influencing Pillar Stability (PS). The PCA has selected the pillar design parameters in the form of principal components. The principal components were selected on the basis of the eigenvalue (when Eigen value>1). Subsequently, from the principal components group, only those pillar design parameters were selected which offered high order of correlations with Pillar Stability.

Methodology

Following parameter selection, it was discovered that several of the parameters had multi-collinearity. A phrase known as the Variance inflation factor (VIF) was used to reduce the multi-collinearity. The Variance Inflation Factor (VIF) was used to assess the strength of the predictor variables' connection (Polhemus,2005). A VIF larger than 10 suggests that the independent variables are multi-collinear (Wang et al., 2016). Therefore, to remove the multi-collinearity, we further selected only those independent variables (pillar design parameters) which revealed VIF values of less than 10. The flowchart of the methodology has been illustrated in Figure 3.6.

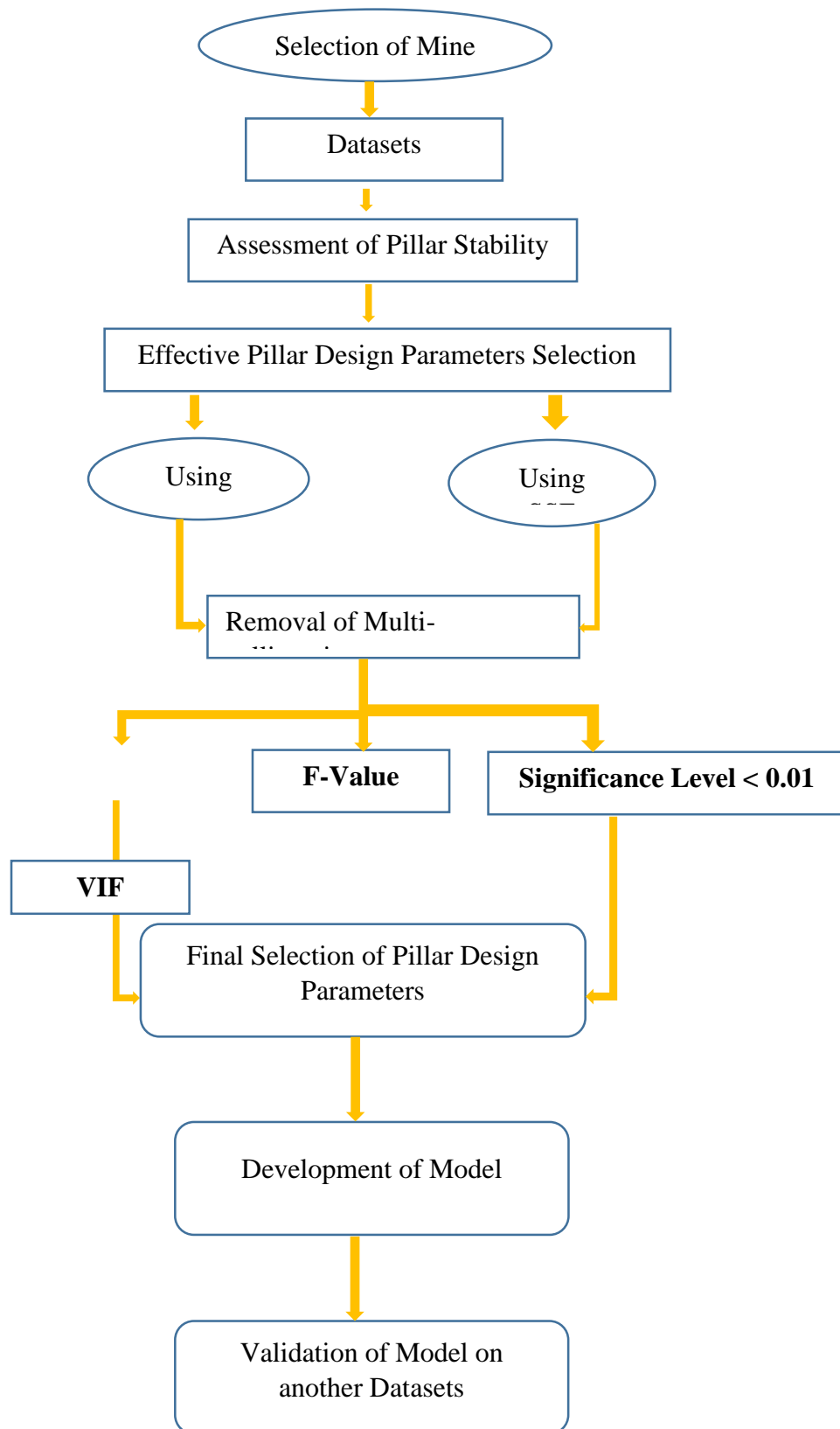


Figure 3.6: Flowchart of methodology

Methodology

The screenshots of the PCA technique, as explained in the preceding passages is sequentially revealed in series of screenshots in (Figure 3.7).

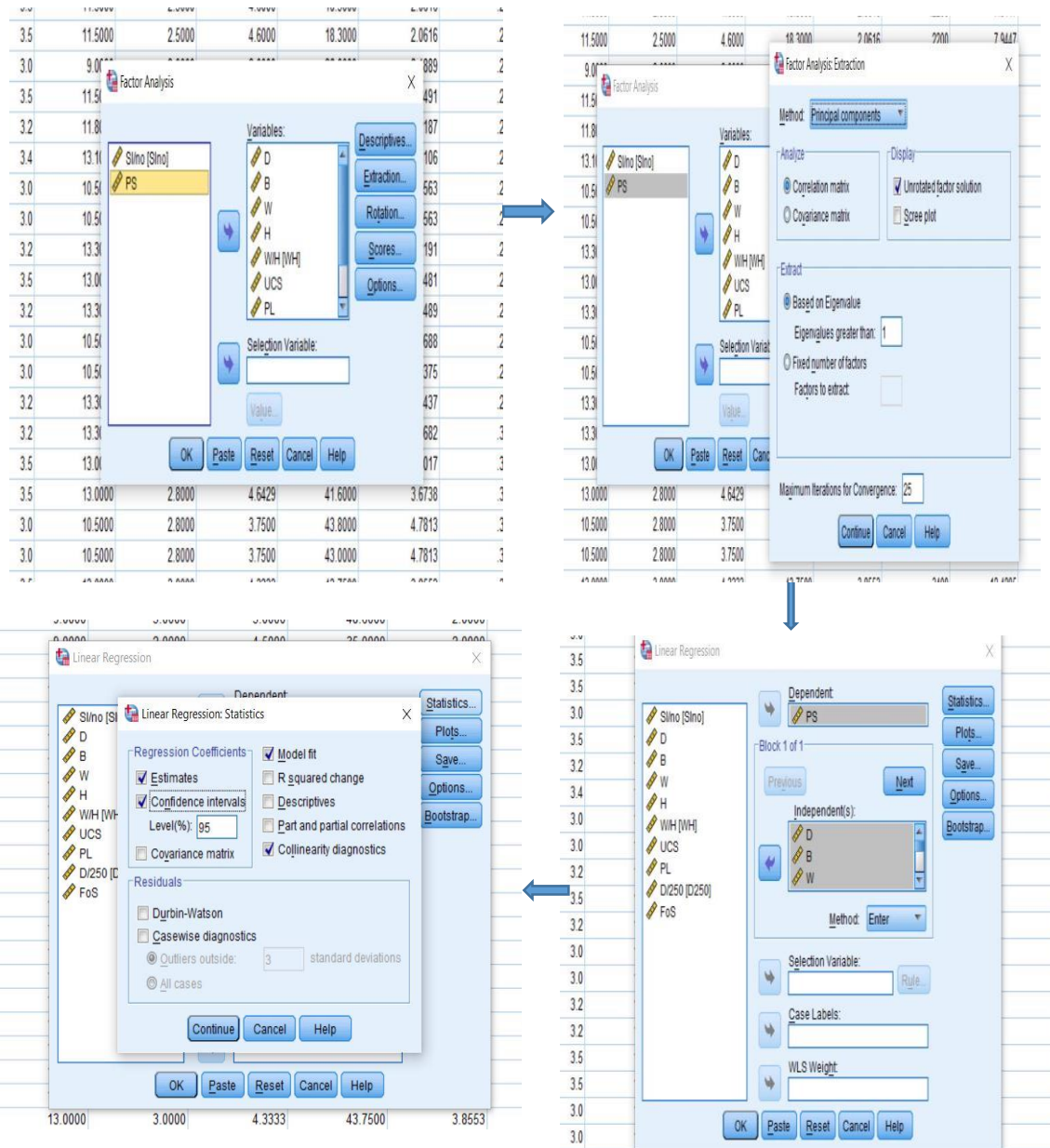


Figure 3.7: Sequential screenshots of the PCA method

3.8 Step-wise Selection and Elimination (SSE) Technique

SSE is a well-known approach for selecting explanatory variables (pillar design parameters in this example) to be included in a multi-variate regression model. The

parameters impacting Pillar Stability are chosen and eliminated based on Pearson's correlation coefficient and the significance value of the input parameters. On the basis of Pearson's correlation coefficient, the parameters with significance less than 0.05 were chosen, whereas those with significance, more than 0.05 were deemed inconsequential and so eliminated. After the selection of pillar design parameters, it was found that some of the parameters were having multi-collinearity. To remove the multi-collinearity, a term called Variance inflation factor (VIF) was deployed. To remove the multi-collinearity, we further selected only those independent variables (pillar design parameters) which revealed VIF values of less than 10.

The screenshots of the SSE technique, as explained in the preceding passages are sequentially revealed in a series of screenshots in Figure 3.8.

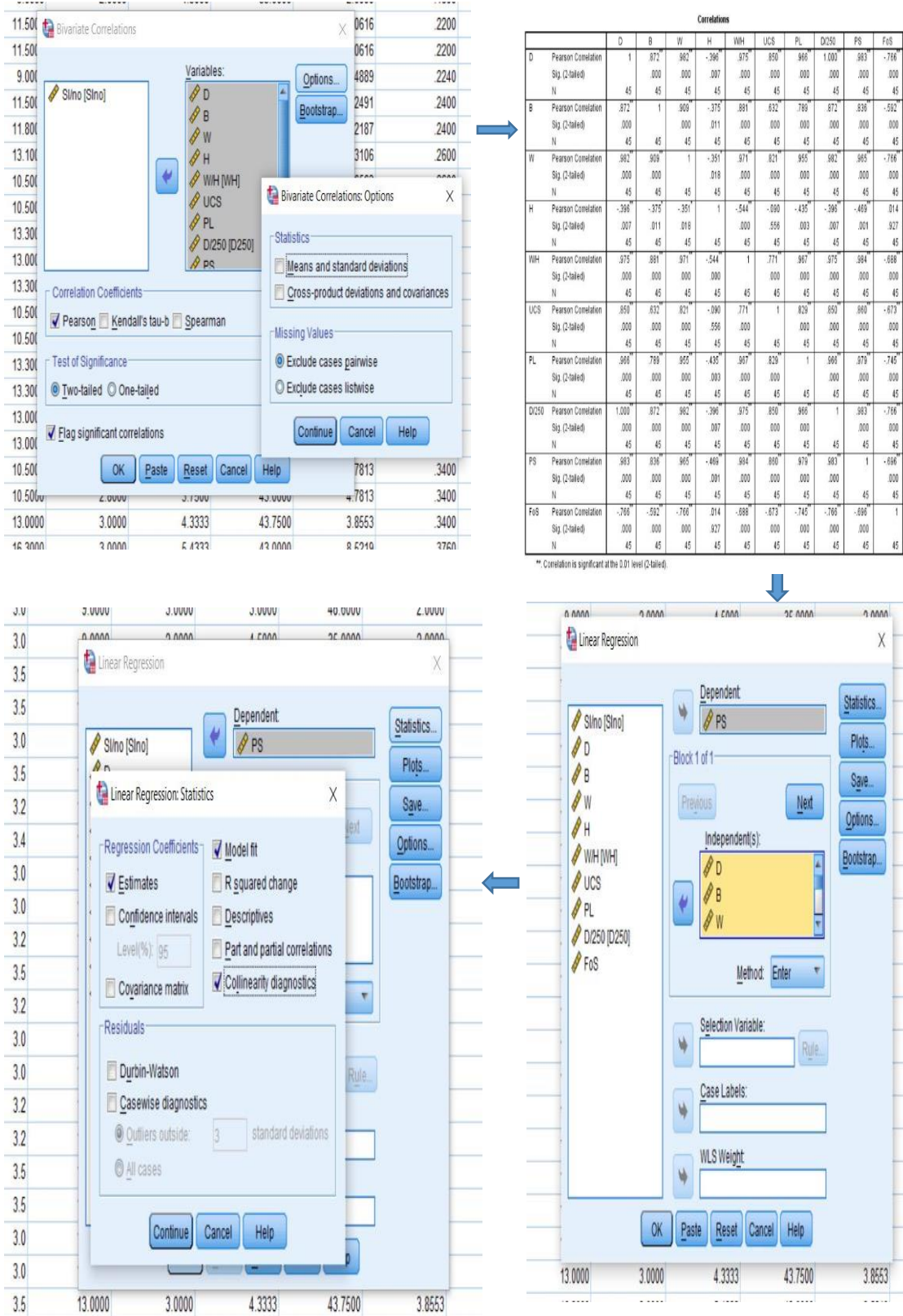


Figure 3.8: Sequential Screenshots of SSE method

3.9 Multi-variate Linear Regression (MLR) Technique

When using MLR, the multi-collinearity among input variables must be discovered and eliminated. The Variance inflation factor (VIF) is used to reduce multi-collinearity. Another metric, the (F-ratio), reflects the dependability of the given results. As the F-ratio raises, the results become more reliable. As a result, the methodology is built on picking the model with the highest F-ratio value.

Following the final parameter selection, the MLR approach was utilized to create the model for both PCA and SSE. As shown in Equation (3.4), the MLR method was utilized to create the relationship between input and output parameters for creating the model:

$$\bar{Y} = a + b_1X_1 + b_2X_2 + \dots + b_n X_n \tag{3.4}$$

where,

- \bar{Y} Predicted value of Y
- a Intercept
- b Partial regression coefficient

Based on the above equation, the model has been developed for pillar strength as an output parameter for both PCA and SSE techniques. Further, a factor of safety has been estimated using the observed and theoretical value of pillar strength. Also, the model has been validated on another dataset to examine the accuracy of the developed model in the prediction of pillar strength.

