

Exploring Machine Learning Techniques for Predicting Open Stope Stability in Underground Mining: Evaluating Accuracy and Applicability

by

Alicja Szmigiel

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mining Engineering

Department of Civil and Environmental Engineering

University of Alberta

©Alicja Szmigiel, 2024

ABSTRACT

Underground mining operations are inherently dangerous due to a variety of factors present in a mining environment. Firstly, the confined spaces and limited ventilation, the use of heavy machinery, explosives, and drilling equipment poses significant risks to the safety of workers. Moreover, underground mines are susceptible to geological hazards such as rockfalls, collapses, and seismic events.

Collapsed and caving openings in underground mining are particularly hazardous due to the potential for catastrophic events. When openings collapse or cave in, they can trap workers underground, leading to injuries, fatalities, and the disruption of rescue operations. Furthermore, collapses can destabilize the surrounding rock mass, leading to further collapses.

Observing and assessing the stability of underground openings is important for several reasons. Firstly, it ensures the safety of workers by identifying potential hazards before accidents occur. By monitoring the stability of openings, mining companies can implement preventative measures such as reinforcement and support systems. Additionally, assessing the stability of underground openings allows for informed decision-making regarding mining operations, ensuring the sustainability and efficiency of production while minimizing risks to personnel and equipment.

Machine learning methods offer promising solutions to the stability assessment problem in underground mining. Through various techniques such as classification and feature importance analysis, machine learning algorithms can effectively predict and evaluate the stability of underground openings. Classification models can classify openings as stable, unstable, or caved based on input features such as geological characteristics and historical stability data. Feature

importance analysis helps identify critical factors influencing stability, enabling targeted interventions.

This research study presents a comprehensive investigation of various machine learning models applications, aimed to predict the stability of underground mining openings, particularly stopes. Open stopes are integral to underground mining operations, where they serve as excavated voids created during the extraction of mineral resources from underground deposits.

Chapter 1 of this thesis presents the groundwork by providing a comprehensive overview of the research topic, outlining its primary objectives, the methodology employed, and the structure of the thesis.

Chapter 2 of this study provides a comprehensive engineering background and overview of open stopes mining operations. The chapter begins with an explanation of the terminology associated with mining methods. Moreover, the chapter elaborates on the most popular methods used for rock mass classification.

In Chapter 3, an extensive literature review of contemporary methods for assessing the stability of open stopes is presented. This review presents a diverse range of approaches, including empirical methods, statistical analyses, and applications of machine learning techniques, which have been proposed by various researchers to address the challenge of evaluating stope stability.

Chapters 4, 5, and 6 present the results of various machine learning models, that were developed to predict the stability of open stopes. Chapter 4 utilizes a Potvin database, where stability number N' and shape factor HR of each historical case were used, and each case had a stability assessment assigned. Random Forest (RF) and Logistic Regression models were employed, evaluated, and compared to achieve the most accurate predictions. In Chapter 5 an

extensive analysis of the Potvin database was performed and used to develop the most effective Artificial Neural Network (ANN) model. In this study all the parameters that combine into the stability number N' were employed and treated as a separate input features for the model. Various ANN model configurations were utilized and evaluated to find the most effective network configuration. The feature importance analysis was then performed to find the parameters that have the highest influence on the stability of an open stope. A final chapter 6, presents an analysis of a larger database obtained from literature, followed by a comparison of several machine learning models. The models' results were then analyzed and the most important features for each model were determined.

In essence, this thesis systematically integrates machine learning techniques to predict the stability of open stopes in underground mining. Through this approach, feature importance analysis was conducted to determine the parameters exerting the greatest influence on stope stability. By evaluating the dataset and identifying key influencing factors, the study enables better control over stope stability, it allows for a more precise understanding of the conditions that lead to failure, enhancing safety and operational efficiency in underground mining environment.

PREFACE

This thesis is an original work by Alicja Szmigiel. Some parts are based on published papers of the author. Chapter 3 is based on the paper Szmigiel, Alicja and Apel, Derek B. (2022) "Predicting the stability of open stopes using Machine Learning, "Journal of Sustainable Mining: Vol. 21 : Iss. 3 , Article 7. Available at: <https://doi.org/10.46873/2300-3960.1369>

This study was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Collaborative Research and Development (CRD)

This research was led by Professor Derek Apel at the University of Alberta.

DEDICATION

This thesis is dedicated to:

*I dedicate this thesis to my wonderful parents, Joanna and Stanislaw Szmigiel,
for their unwavering support, love, and encouragement.*

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Derek Apel. I am profoundly thankful for the opportunity to pursue my PhD under his guidance at the University of Alberta in Canada. His consistent support, immense knowledge, and endless patience were invaluable throughout this journey. Dr. Derek, your encouragement and insights have inspired me to achieve my best and have significantly contributed to the success of this research.

I owe a deep sense of gratitude to my parents, Joanna and Stanislaw Szmigiel. Your support, encouragement, and belief in me have been the foundation of my journey. From the very beginning, you initiated the idea of pursuing higher education and have stood by me every step of the way. Your confidence in my abilities and constant motivation kept me going even during the toughest times. Thank you for always believing in me and for helping me fulfill this dream.

I would like to express my deepest gratitude to my wonderful boyfriend, Peter Kasprzak. Your unwavering support and dedication have been incredible. Moving from Poland to Canada, leaving behind your job and the security of your established life, just to be by my side, is a testament to your love and commitment. Even during the most challenging times, you never stopped believing in me. You gave me the space I needed to focus on my studies and stood by me with support. Without your encouragement, completing this journey would have been much more challenging. I will forever cherish and remember the sacrifices you made for me. Thank you for being my rock and for making this accomplishment possible.

I would like to express my gratitude and appreciation to the thesis examining committee members: Dr. Hooman Askari-Nasab, Dr. Hassan Dehghanpour Hossein Abadi, Dr. Gordon Wilson, Dr. Robert Hall and Dr. Rudrajit Mitra, for their precious time and constructive criticisms and comments.

Thank you for all your encouragement

TABLE OF CONTENTS

ABSTRACT.....	ii
PREFACE	v
DEDICATION.....	vi
ACKNOWLEDGEMENTS.....	vii
TABLE OF CONTENTS.....	viii
LIST OF FIGURES	xiii
LIST OF TABLES	xvii
LIST OF ABBREVIATIONS.....	xviii
CHAPTER 1: INTRODUCTION.....	1
1.1. General background of this research.....	2
1.2. Research objectives and methodology.....	2
1.3. Organization of this thesis	4
CHAPTER 2: ENGINEERING BACKGROUND ON OPEN STOPES MINING OPERATIONS AND ROCK MASS CLASSIFICATION SYSTEMS	7
2.1 Open stope mining operations	8
2.2. Classification of open stope mining methods	11
2.3. Dillution and Hangingwall overbreak.....	16
2.4. Rock mass characteristics	17
2.5. Rock mass classification systems	19

2.5.1. Rock Quality Designation (RQD).....	20
2.5.2. Rock mass characteristics (Q value).....	22
2.5.3. Rock Mass Rating (RMR)	28
2.6. Summary	29
References.....	30
CHAPTER 3: LITERATURE REVIEW ON STOPE STABILITY ASSESSMENT METHODS	
.....	36
3.1. Analytical Methods.....	37
3.1.1 Kinematic, Beam Failure, and Plate Buckling Analysis.....	37
3.1.2. Voussoir Beam Analysis.....	38
3.1.3. Mohr - Coulomb Criterion	40
3.1.4. Hoek and Brown failure criterion	41
3.2. Empirical Methods for open stope design	42
3.2.1. Design tools for open stope.....	42
3.2.1.1. Rock Stress Factor A	43
3.2.1.2. Joint orientation adjustment Factor B	46
3.2.1.3. Surface orientation factor.....	48
3.2.1.4. Shape Factor (Hydraulic Radius HR)	51
3.2.2. Stability graphs	52
3.3. Statistical analysis and numerical modeling.....	58
3.4. Machine learning methods.....	59
3.5 Summary	60
3.6. References.....	61

CHAPTER 4: PREDICTING THE STABILITY OF OPEN STOPES USING MACHINE LEARNING	64
4.1 Introduction.....	65
4.2. Open stope mining method	67
4.2.1. Matthew stability graph method review	67
4.2.2 Database.....	69
4.3. Model development, evaluation, and results.....	76
4.3.1. K-fold Cross-Validation.....	76
4.3.2. Confusion Matrix and ROC AUC score	77
4.3.3 Logistic Regression.....	79
4.3.3.1 Development of the model.....	79
4.3.3.2 Results.....	79
4.3.4. Random forest.....	81
4.3.4.1 Development of the model.....	82
4.3.4.2 Hyper-parameters tuning	83
4.3.4.3 Results.....	84
4.4 Summary and conclusions	85
4.5 References.....	87
CHAPTER 5: ENHANCING UNDERGROUND EXCAVATIONS STABILITY PREDICTION IN MINING ENGINEERING: OPTIMAL CONFIGURATION OF AN ARTIFICIAL NEURAL NETWORK MODEL	89
5.1 Introduction.....	90
5.2 Engineering background and literature review	91

5.2.1 Stability Parameters	92
5.2.2 Literature review	94
5.3 Artificial Neural Network Model Overview	96
5.3.1 Preprocessing the data.....	99
5.3.2 ANN model structure.....	102
5.3.3 Model evaluation	105
5.4 Model development and results	107
5.4.1 Model results.....	112
5.4.2 Exploring feature importance with SHAP	119
5.5. Summary	123
5.6 References.....	125
CHAPTER 6: EXPLORING MACHINE LEARNING TECHNIQUES FOR OPEN STOPE STABILITY PREDICTION: A COMPARATIVE STUDY AND FEATURE IMPORTANCE ANALYSIS.....	130
6.1 Introduction.....	131
6.2 Background of stope stability assessment.....	132
6.3 Database analysis and pre-processing.....	137
6.3.1 Database analysis	145
6.3.2 Database pre-processing	150
6.4 Comparative study overview	152
6.4.1 Investigated Machine Learning models	153
6.4.1.1 Random Forest	154
6.4.1.2 Support Vector Machine	154

6.4.1.3 Adaptive Boosting - AdaBoost	155
6.4.1.4 Extreme Gradient Boosting - XGBoost	156
6.4.1.5 LightGBM.....	156
6.4.1.6 Artificial Neural Network	157
6.5 Evaluation and results	158
6.5.1 Training and evaluation	158
6.5.2 Results.....	160
6.5.3 Feature importance analysis with SHAP	164
6.6 Summary	169
6.7 References	172
CHAPTER 7: SUMMARY, CONCLUSION AND RECOMMENDATIONS	177
7.1 Summary of the research	178
7.2 Conclusions of this research	179
7.3 The prospects of future research	183
BIBLIOGRAPHY	185

LIST OF FIGURES

Figure 2.1. A basic open stope geometry (Modified from Heidarzadeh et al. 2019)	9
Figure 2.2. A typical cross-section of a single lift open stoping mine (Modified from Potvin and Hudyma 2000)	11
Figure 2.3. Longitudinal open stope mining, or sublevel retreat (Modified from Potvin and Hudyma 2000)	13
Figure 2.4. Transverse blasthole open stoping (Modified from Potvin and Hudyma 2000)	13
Figure 2.5. Long - hole blasting stoping (Modified from Harraz 2010).....	15
Figure 2.6. Shrinkage stoping (Modified from Hamrin 1997).....	16
Figure 2.7. Planned and unplanned dilution (Modified from Scoble and Moss 1994).....	17
Figure 2.8. RQD procedure and measurement (Modified from Potvin and Nedin 2003)	21
Figure 3.1. Compression arch in a deflecting beam (From Diederichs and Kaiser 1999).....	39
Figure 3.2. Mohr Coulomb failure criterion (From Goodman 1989)	41
Figure 3.3. Graph of the stresses induced on the major surface of a stope vs. the ratio of opening dimensions. (Modified from Potvin, 1988).....	44
Figure 3.4. Graph of the stresses induced on the minor surface of a stope vs. the ratio of opening dimensions. (Modified from Potvin, 1988).....	45
Figure 3.5. Graph for the estimation of Rock Stress Factor A. (Modified from Potvin, 1988)....	46
Figure 3.6. Sketch for the determination of B factor. (Modified from Mathews et al, 1980)	47
Figure 3.7. Graph for the determination of Rock Stress Factor B. (Modified from Hutchinson and Diederichs 1996).....	48
Figure 3.8. Determination of Stability Factor C Gravity Fall & Slabbing. (Modified from Hutchinson and Diederichs 1996).....	49

Figure 3.9. Determination of Stability Factor C Sliding. (Modified from Hutchinson and Diederichs 1996).....	50
Figure 3.10. Determination of Hydraulic Radius.....	51
Figure 3.11. Mathews Stability Graph (After Mathews et al., 1981, From Capes 2009).....	53
Figure 3.12. Modified Stability Graph (After Potvin, 1988, From Capes 2009).....	54
Figure 3.13. Modified Stability Graph (After Nickson, 1992, From Capes 2009).....	55
Figure 3.14. Modified Stability Graph (After Stewart and Forsyth, 1993, From Capes 2009)....	56
Figure 3.15. Modified Stability Graph (After Hadjigeorgiou et al. 1995, From Capes 2009)	57
Figure 4.1. Distribution of the Potvin database.	70
Figure 4.2. The process of k-fold Cross-Validation.	77
Figure 4.3. Decision boundaries for Logistic Regression.....	80
Figure 4.4. Confusion Matrix for Logistic Regression.	80
Figure 4.5. The architecture of Random Forest algorithm (Modified from Qi et al., 2018)	82
Figure 4.6. The AUC score for several numbers of decision trees	83
Figure 4.7. The AUC score for different values of decision trees depth.	84
Figure 4.8. Confusion Matrix for Random Forest multiclass classification.	85
Figure 5.1 Open slope basic geometry.....	94
Figure 5.2 Pairs plot visualization of Potvin database.....	98
Figure 5.3 The steps of dividing the database into subsets.....	101
Figure 5.4 Diagram of one single processing element.....	102
Figure 5.5 Bar plot of stability assessment count	108
Figure 5.6 ReLU and Swish activation functions	109
Figure 5.7 Final ANN model structure	111

Figures 5.8 and 5.9 Accuracy and loss score vs epochs, for a data scaled with MinMaxScaler ReLu activation function	112
Figures 5.10 and 5.11 Accuracy and loss score vs epochs, for a data scaled with MinMaxScaler Swish activation function.....	113
Figures 5.12 and 5.13 Accuracy and loss score vs epochs, for a data scaled with StandardScaler Swish and ReLu activation functions	113
Figures 5.14 and 5.15 Accuracy and loss score vs epochs, for a data scaled with StandardScaler Swish activation function.....	114
Figure 5.16 Confusion Matrix of the prediction performance on the testing set.....	117
Figure 5.17 N vs HR plot for actual classes.....	118
Figure 5.18 N vs HR plot for predicted classes	118
Figure 5.19 SHAP Summary Plot: Feature Importance Analysis.....	121
Figures 5.20 and 5.21 Distribution of HR and Q values across stability classes.....	122
Figure 6.1. Calculation of Hydraulic radius.....	133
Figure 6.2 Distribution of cases in database. Green – stable cases, blue – unstable and red – caved.	139
Figure 6.3 Histograms of the parameters	146
Figure 6.4. Box – plots of the parameters	147
Figure 6.5. Correlation heatmap of all parameters	148
Figure 6.6 Pre-processing workflow for the database preparation	152
Figure 6.7 Accuracy score for all the Machine Learning models.....	161
Figure 6.8 Classification report for all the Machine Learning models.....	161
Figure 6.9 Precision scores for all Machine Learning models.....	163
Figure 6.10 Recall scores for all Machine Learning models	163

Figure 6.11 Random Forest - SHAP value summary plot	165
Figure 6.12. SVM - SHAP value summary plot	166
Figure 6.13 AdaBoost- SHAP value summary plot.....	166
Figure 6.14 XGBoost- SHAP value summary plot.....	167
Figure 6.15 LightGBM- SHAP value summary plot	167
Figure 6.16 ANN model - SVM - SHAP value summary plot	168

LIST OF TABLES

Table 2.1. Summary of rock mass characteristics (From Hoek et al. 2000).....	18
Table 2.2. Classification of all parameters used in the NGI Q classification system (From Hoek and Brown, 1996).....	23
Table 2.3. Classifications of rock mass quality based on Q (From Barton et al. 1974a).....	27
Table 2.4. RMR inputs (From Bienawski 1976).....	28
Table 4.1. Complete Database from Potvin, 1988.....	70
Table 4.2. The Confusion Matrix.....	78
Table 5.1 Potvin's data set investigated – first 15 examples (complete database can be found in Chapter 4).....	98
Table 5.2 Classification report for testing set	116
Table 6.1 Complete Adoko et al. (2022) Database.....	140

LIST OF ABBREVIATIONS

AB	Adaptive Boosting
AI	Artificial Intelligence
ANN	Artificial Neural Network
AUC	Area Under the Curve
BAC	Bootstrap Aggregating Classifier
BP	Backpropagation
CMS	Cavity Monitoring System
CNN	Convolutional Neural Network
CV	Cross Validation
DTR	Decision Tree
ELOS	Equivalent Linear Overbreak/Slough
FFNN	Feed Forward Neural Network
FN	False Negative
FP	False Positive
GB	Gradient Boosting
GBM	Gradient Boosting Machine
HR	Hydraulic Radius
HW	Hanging Wall
KNN	K-Nearest Neighbors
LR	Logistic Regression

NB	Naive Bayes
NGI	Norwegian Geotechnical Institute
ML	Machine Learning
MLPNN	Multilayer Perceptron Neural Networks
RBF	Radial Basis Function
ReLU	Rectified Linear Units
RF	Random Forest
RMR	Rock Mass Rating
ROC	Receiver Operator Characteristic
RQD	Rock Quality Designation
SHAP	Shapley Additive exPlanations
SRF	Stress Reduction Factor
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
UCS	Uniaxial Compressive Stress

CHAPTER 1: INTRODUCTION

This chapter is an overview of this thesis, which provides the research background, research objectives, and methodologies. The organization of this thesis is also outlined at the end of this chapter.

1.1. General background of this research

Underground mining operations present inherent risks to worker safety due to various environmental factors existing in its surroundings. Confined spaces and limited ventilation amplify the potential for accidents and exposure to many hazards. Moreover, the use of heavy machinery, explosives, and drilling equipment further heightens the safety risks faced by workers. Additionally, underground mines are vulnerable to geological threats such as rockfalls, collapses, and seismic events, which can occur suddenly and without warning.

Among the most dangerous scenarios in underground mining are collapsed and caving openings, which create significant risks of catastrophic events. Such incidents can lead to workers being trapped underground, resulting in injuries, fatalities, and disruptions to rescue efforts. Furthermore, collapses have the potential to destabilize the surrounding rock mass, leading to subsequent collapses and ongoing risks to both workers and mine infrastructure.

To prevent these catastrophes, assessment of the stability of underground openings is critical, for ensuring worker safety. By identifying and addressing potential hazards proactively, mining companies can minimize the risk of accidents before they occur. Implementing reinforcement and support systems based on stability assessments helps mitigate the risk of collapses. Additionally, evaluating the stability of underground openings allows for informed decision-making in mining operations, balancing production efficiency with safety considerations.

1.2. Research objectives and methodology

The primary objective of this research study is to assess the stability of open stopes in underground mining using various machine learning models and feature importance evaluations. By applying machine learning techniques, including Random Forest, Logistic Regression, and Artificial Neural Networks, I aim to develop predictive models capable of accurately determining the stability of open stopes based on input features such as stability number and shape factor. Additionally, I seek to identify the features that have the most significant impact on the final stability predictions, providing insights into the key factors influencing stope stability. Through

this analysis, I aim to contribute to the advancement of predictive capabilities in underground mining stability assessment, ultimately enhancing safety and efficiency in mining operations.

While previous attempts by researchers to utilize machine learning in the prediction of stope stability were limited and lacked a systematic approach, this thesis takes a comprehensive and structured approach to address this challenge. Unlike isolated and unsystematic trials in the past, this research incorporates various facets of machine learning, including data preprocessing, supervised learning, and model performance evaluation, and feature importance analysis. Moreover, state-of-the-art machine learning platforms and toolkits such as TensorFlow, Keras, and Scikit-Learn were employed to ensure successful modeling. Consequently, this thesis can be considered a systematic exploration of data-driven strategies for stope stability prediction, demonstrating an accomplished effort to integrate machine learning techniques into traditional engineering domains such as mining and mineral engineering. Through this comprehensive approach, the thesis aims to bridge the gap between conventional engineering practices and emerging data-driven methodologies, introducing the way for more accurate and reliable stability predictions in underground mining operations.

To achieve the objectives of this research, the following tasks have been fulfilled.

1. An extensive literature review was implemented for this research, which mainly included aspects such as:
 - empirical methodologies used for stope stability assessment,
 - numerical and statistical methodologies used for stope stability assessment,
 - previous achievements and applications of machine learning methods for stope stability assessment.
2. A thorough data collection from public publications with stope features and stability conditions was conducted. First database utilized consisted of 176 cases, followed by a second more detailed one that consisted of 225 cases.
3. Different machine learning models were developed in this study to predict the stability of open stopes. For each model, the mathematic mechanism was adhered to exhibit the model's practicability. The model's prediction results were evaluated with various metrics to ensure

their predictive accuracy and avoid undesirable behavior of the model, such as underfitting and overfitting.

4. In the conclusive phase, a comprehensive examination of feature importance was conducted to identify the key parameters exhibiting the most significant influence on both the model predictions and the stability of the underground openings. This detailed analysis aimed to provide insights into the critical factors driving the predictive capabilities of the models and their implications.

1.3.Organization of this thesis

This thesis is comprised of seven chapters in total. All of the chapters are titled as follows: Chapter 1 (Introduction); Chapter 2 (Engineering background on open stope mining operations and rock mass classification systems); Chapter 3 (Literature review on stope stability assessment methods); Chapter 4 (Predicting the stability of open stopes using Machine Learning); Chapter 5 (Enhancing underground excavations stability in mining engineering: optimal configuration of an artificial neural network model); Chapter 6 (Exploring machine learning techniques for open stope stability prediction: a comparative study and feature importance analysis); Chapter 7 (Summary, conclusion, and prospects).

Chapter 1 serves as an introductory foundation for the thesis, presenting the broader context and risks associated with unstable openings in mining environments. It explains the rationale behind the research, highlighting the essential need to address the hazards posed by collapsing and unstable openings. Furthermore, Chapter 1 provides the objectives and methodologies adopted for conducting the study.

Chapter 2 of this thesis presents an extensive exploration of the engineering fundamentals and operational methods where open stopes are utilized. It delves into the traditional and widely recognized approaches developed for classification of the rock masses. The chapter begins by explaining the terminology essential to these mining techniques, ensuring a coherent understanding of the fundamental principles important for evaluating stope stability. Furthermore, it provides in-depth insights into the diverse methods and parameters developed to classify and assess the quality of the rock mass. This comprehensive analysis in Chapter 2 serves as a

fundamental framework for subsequent examinations and discussions referring to methodologies for assessing slope stability.

Chapter 3 presents a thorough examination of recent approaches in assessing the stability of open slopes presented in literature. This review demonstrates various methods, including empirical techniques, statistical analyses, and the utilization of machine learning algorithms, proposed by researchers to address the challenge of evaluating slope stability. Through a detailed exploration of each method and outcomes, the chapter offers insights into their strengths, limitations, and potential applications within the field of mining engineering. By integrating the latest research findings and advancements in slope stability assessment methods, Chapter 3 serves as a valuable reference for understanding the current state-of-the-art methods and identifying paths for further research and development in this area of study.

Chapter 4 of the thesis focuses on implementing machine learning models for predicting the stability of open slopes. In this study, two widely used machine learning algorithms, Logistic Regression and Random Forest, were employed. The analysis investigated the Potvin database, consisting of 176 case studies, where two key variables, stability number (N) and shape factor (HR), were considered alongside their respective stability conditions: stable, unstable, or caved. Extensive performance evaluations were conducted, and hyperparameters were fine-tuned to optimize model performance and address issues like overfitting. By utilizing these machine learning techniques and adjusting the models, the study aimed to enhance the accuracy and reliability of slope stability predictions.

Chapter 5 presents a comprehensive analysis of the Potvin database to construct an optimized Artificial Neural Network (ANN) model. The study incorporates five parameters contributing to the stability number N for a thorough examination. Various configurations of ANN models were explored and assessed to determine the optimal network structure adjusted to the Potvin database. Furthermore, feature importance analysis was conducted to identify the key parameters exerting the greatest influence on the stability of open slopes. By employing ANN models and analyzing feature importance, the chapter aims to refine slope stability predictions and enhance overall model performance.

In the chapter 5, an exhaustive analysis is conducted on a substantial database acquired from the literature, offering a robust foundation for the subsequent evaluation. This analysis encompasses a detailed comparison of various machine learning models, including Random Forest (RF), Support Vector Machine (SVM), AdaBoost, LightGBM, XGBoost, and Artificial Neural Network (ANN). These models are specifically developed to predict the stability conditions of open stopes, leveraging a comprehensive set of seven features combining into shape factor and stability number N . Following the implementation of these models, their outcomes are thoroughly analyzed to determine the most influential features driving their predictive accuracy. This detailed examination explains the complex relationships between input variables and stability outcomes, ultimately informing strategies for optimizing predictive performance and enhancing the applicability of stope stability assessment models.

In the final Chapter 6, I summarize the insights gathered from the previous chapters, explain the essence of our research, while presenting the course for future exploration. With a careful examination of the datasets and a systematic evaluation of various machine learning models, I have aimed to discover the predictive potential inherent in stope stability assessments. This final chapter presents future prospects of this field of study, highlighting the escalating significance of machine learning models in predicting stope stability. As technological advancements continue to evolve, the integration of machine learning algorithms could revolutionize the way we approach, and address risks associated with underground mining operations.

CHAPTER 2: ENGINEERING BACKGROUND ON OPEN STOPES MINING OPERATIONS AND ROCK MASS CLASSIFICATION SYSTEMS

The chapter provides an engineering background on open stope mining operations, presenting various methods employed in this mining technique. It outlines the challenges in open stope mining and discusses the basic systems and applications used to overcome these challenges. Additionally, the chapter presents the most popular and widely employed rock mass classification systems in open stope mining, emphasizing their role in assessing the stability of underground openings. The chapter presents insights into the complexities of open stope mining operations by covering the operational methods and geomechanical considerations.

2.1 Open stope mining operations

Open stope mining involves extracting a large ore block through drilling and blasting techniques from pre-existing access drives within an underground ore deposit. This method is a widely adopted approach for bulk underground mining operations. Open stoping is categorized as a non-entry method, meaning that once production has begun, miners are not required to physically enter the stope. This characteristic contributes to the method's reputation for safety, as workers conduct their activities exclusively along the periphery of the stope, away from the potentially hazardous production face. In contrast, entry mining methods such as cut and fill, longwall, room and pillar, and shrinkage stoping involve miners entering the excavation area for various tasks. The stopes are also designed to be stable, which means that there shouldn't be any large releases of energy caused by a sudden change in the opening geometry caused by caving (Potvin and Hudyma 1989).

The use of heavy machinery to remove the blasted material results in an open void or stope, which can be the size of a large house, surrounded by rock walls, the basic geometry of a stope is shown in Figure 2.1. These voids are then backfilled with various materials, to allow the extraction of the adjacent ore with secondary stopes. The orientation of the walls around the stope varies based on local geological conditions and limitations imposed by mining equipment. An inclined wall that hangs over the stope is commonly known as the hangingwall (HW). Due to its orientation and geometry relative to the open void, the hangingwall of a stope tends to be less stable compared to other stope walls (Pakalnis 1986).

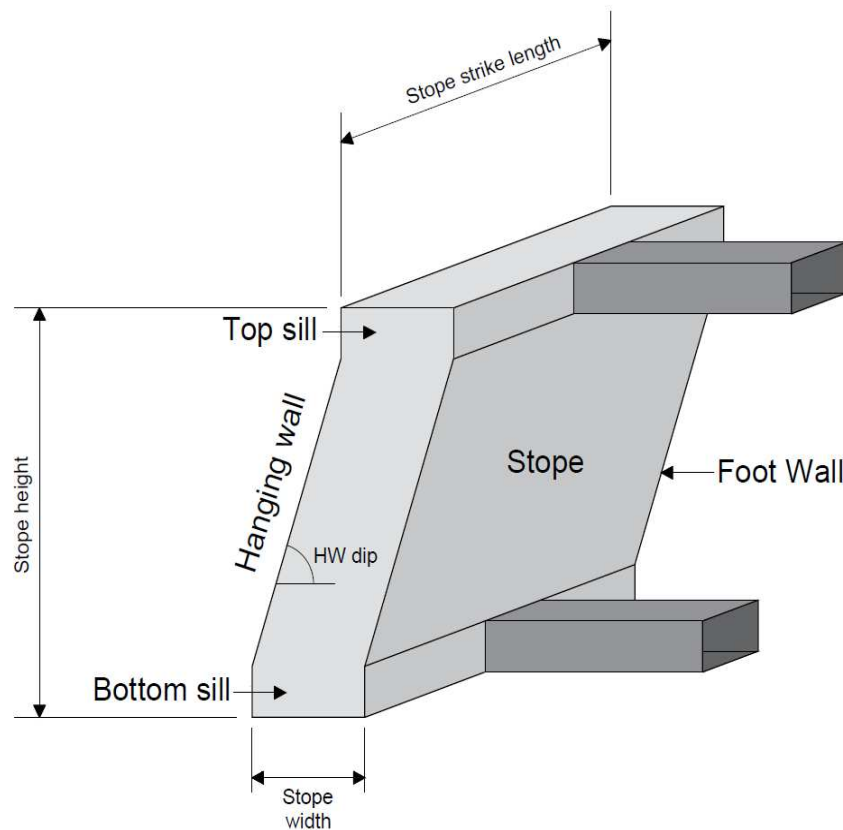


Figure 2.1. A basic open stope geometry (Modified from Heidarzadeh et al. 2019)

Open stope mining demonstrates efficiency under specific orebody and geological conditions. This method relies on the natural gravity flow of ore material toward the stope bottom. Therefore, the inclination of the stope should exceed the angle of repose of the fragmented ore material, typically greater than 50° to 55° . While open stope mining is more compatible with steeply dipping orebodies, it can still be successfully used in shallow dipping orebody (typically less than 30°). In such cases, the stopes must be oriented in a sub-vertical manner, and the orebody should have a minimum thickness of 15 to 20 meters. The orebody contour should have a relatively consistent shape since open stoping is not highly selective. Furthermore, a minimum width of approximately 3 meters is typically required to avoid the risk of excessive dilution coming from damage to the walls caused by blasting vibrations or deviations in drillhole. Open stoping typically employs large opening dimensions. As no significant support system like backfill or pillars is

present inside the stope during mining operations, it is necessary for a rock mass to have a strength ranging from "fair" to "good" for the stope back and walls to provide self-support. Enhanced strength in the rock mass enables the creation of larger stopes which improves stoping efficiency (Potvin and Hudyma 1989)

The approach to open stope mining can vary significantly based on factors such as the direction of access to the orebody, the relationship between orebody geometry and the stability of the hangingwall, as well as economic and safety considerations within the mine. Different methods may be employed to accommodate these varying conditions and requirements. Typically, access to the orebody is established through stope crosscuts from a footwall drive offset from the deposit. The stopes are excavated at vertical intervals of either 30 or 60 meters, with careful consideration given to the quality of the local rock mass and its relationship with stope geometry. The excavation of stopes involves the use of drill and blast techniques, and they are categorized as either primary or secondary stopes based on their position within the mining sequence. Primary stopes are excavated as the first phase of mining operations and subsequently filled with a variety of filling materials. Following this, secondary stopes are typically excavated once the neighboring primary stopes have been mined to a level that is one sublevel higher than the elevation of the secondary stope. Additionally, there is a waiting period of 28 days after filling the primary stopes to allow the filling material to gain adequate strength, which enables it to function effectively as a sidewall during the extraction process. The typical horizontal dimensions - strike spans of the stopes generally range from 10 to 20 meters, with variations based on factors such as local rock mass characteristics, stope geometry, constraints related to filling material, and various other considerations that influence the specific design of each stope and the overall stability of the area. Overbreak, which refers to the unintended displacement of rock beyond the intended stope design, can potentially occur from any of the stope walls. However, the hangingwall typically plays a significant role in contributing to overbreak due to its particular orientation and shape relative to the void created by mining activities. Figure 2.2 (redrawn from Potvin and Hudyma 2000) illustrates a typical cross-section of a single lift stoping operation, with the resultant flat-bottom stopes.

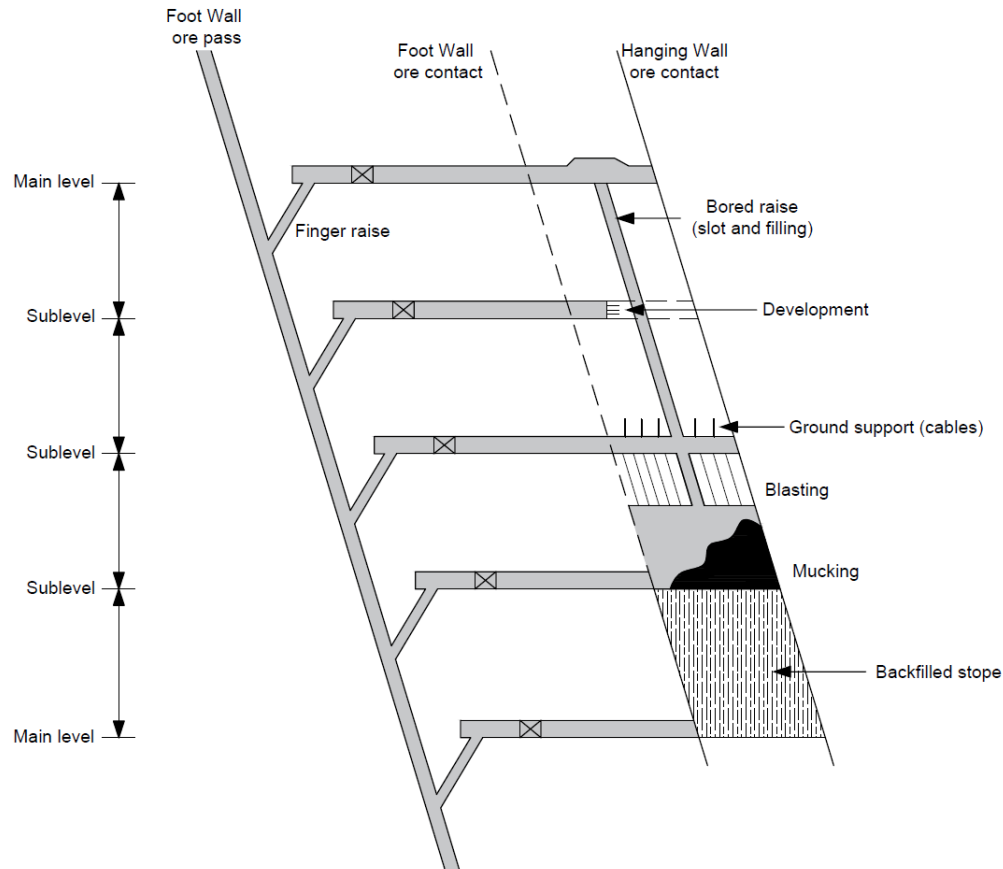


Figure 2.2. A typical cross-section of a single lift open stoping mine (Modified from Potvin and Hudyma 2000)

2.2. Classification of open stope mining methods

Open stope mining methods involve a variety of elements, resulting in unique applications developed for specific conditions. Consequently, no two open stoping scenarios are identical. These methods can be classified based on three main specifications:

- Mining direction (longitudinal or transverse),
- Use of pillars and backfill
- Drillhole diameter (longhole or blasthole)

These classifications determine the level of development and stope preparation needed, the retreat method, as well as the mining sequence.

In longitudinal mining (Figure 2.3), the extraction process progresses along the direction of the orebody's strike, meaning that exploitation advances parallel to the orientation of the mineral deposit. This method is often favored when the orebody is relatively narrow and elongated, as it requires less development work to establish the mining operation. Engineers can access the orebody more quickly and begin extraction sooner, leading to faster production and reduced costs associated with preparatory work. However, a key consideration in longitudinal mining is the stability of the stope backs. If the orebody width exceeds a certain threshold, the stope backs may struggle to maintain stability, posing safety risks and complicating the mining process.

On the other hand, transverse mining (Figure 2.4) involves extracting ore perpendicular to the strike of the orebody. This method is typically employed when the orebody is wider and more expansive, requiring a different approach to extraction. Transverse mining needs greater development work upfront to establish access and infrastructure within the wider orebody. However, once operational, transverse mining can efficiently extract ore from larger areas, leading to higher production rates. A minimum orebody width of around 15 meters is often considered necessary for transverse mining to be economically viable and operationally efficient. This method allows for effective utilization of the orebody's width, optimizing the extraction process while ensuring the stability of the stope backs (Potvin and Hudyma 1989).

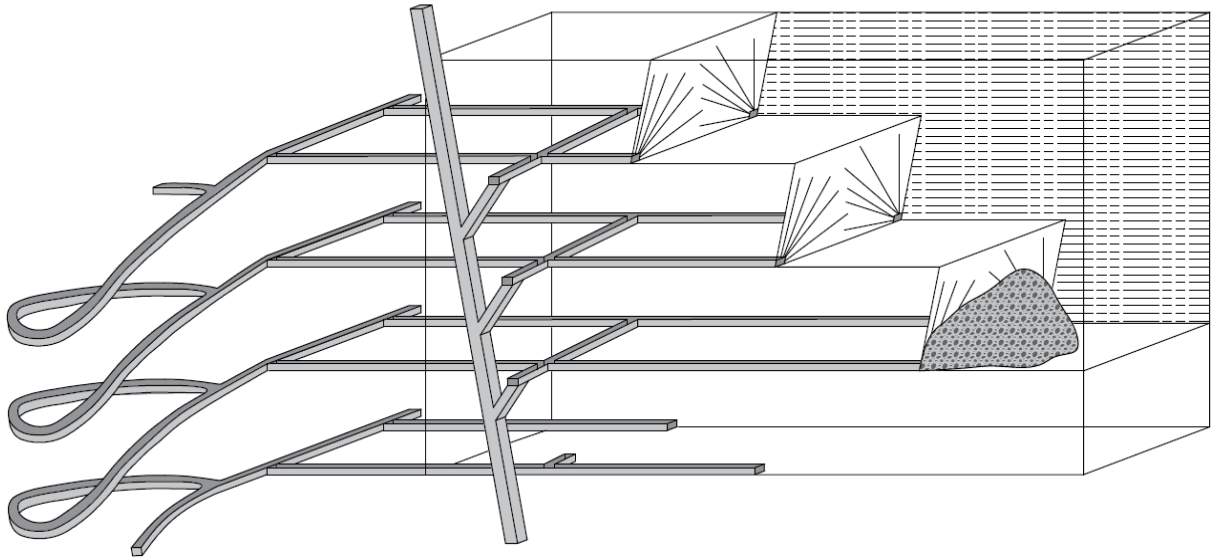


Figure 2.3. Longitudinal open stope mining, or sublevel retreat (Modified from Potvin and Hudyma 2000)

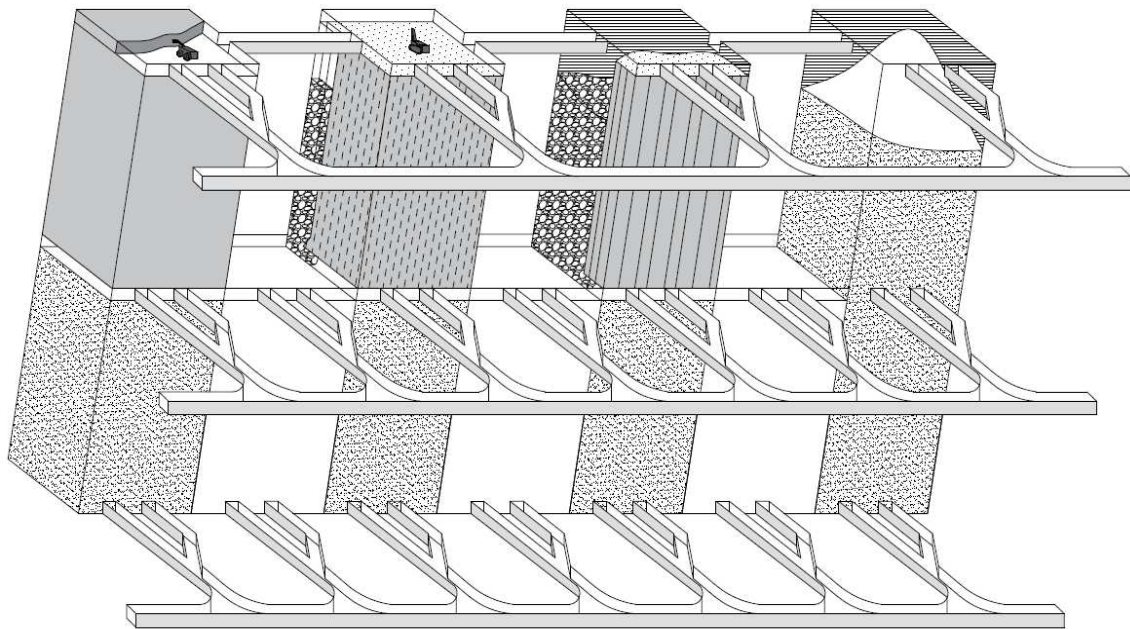


Figure 2.4. Transverse blasthole open stoping (Modified from Potvin and Hudyma 2000)

The most simple and cost-effective method of open stope mining is full lens extraction, characterized by the absence of backfill and pillars. This method is particularly suitable for small orebodies or isolated lenses, provided that the quality of the rock mass allows for self-supporting stope surfaces, including walls and backs. However, in cases where the value of the ore doesn't warrant backfill usage but the orebody is too extensive to be mined in a single stope, permanent pillars are left in place. These pillars are minimized to maximize orebody recovery but must remain stable to ensure overall mine stability and safe access to the stope. When backfill is incorporated, the sequencing of stope extraction becomes integral to an overall strategy aimed at optimizing the recovery of secondary and tertiary stopes. Primary stopes are typically mined against rock walls, while secondary stopes are often mined against one or more cemented backfill walls. Tertiary stopes, surrounded by backfilled stopes, typically utilize uncemented fill. Consequently, primary and secondary stopes often require cemented backfill, whereas tertiary stopes can be filled with uncemented material, highlighting the importance of strategic planning and sequencing in open stope mining operations.

Longhole stoping, shown in Figure 2.5, represents a variation of sublevel open stoping that employs longer blast holes with larger diameters ranging from 140 to 165 mm, typically drilled using the in-the-hole (ITH) technique, with depths of up to 100 m. This method requires miners to initially create a vertical slot at one end of the stope and subsequently working in sublevels to drill a radial pattern of drill holes. Following the loading of these holes, blocks of ore body are blasted to open the stope. In longitudinal longhole stoping, stopes may be retreated in the direction of cross-cuts using either a top-down or bottom-up sequence. Transverse longhole stoping is adopted when the rock mass quality of the hanging wall constrains the length of the open mining span, offering greater flexibility in sequencing and scheduling due to independent access for each stope. Stope sequencing in transverse longhole stoping can effectively mitigate the effects of mining-induced stress by creating an active stress shadow, thereby sheltering existing and future excavations. Many open stope mines employ transverse longhole stoping and sequence their stopes based on the high-stress conditions encountered in underground mining (Kumar et al. 2014).

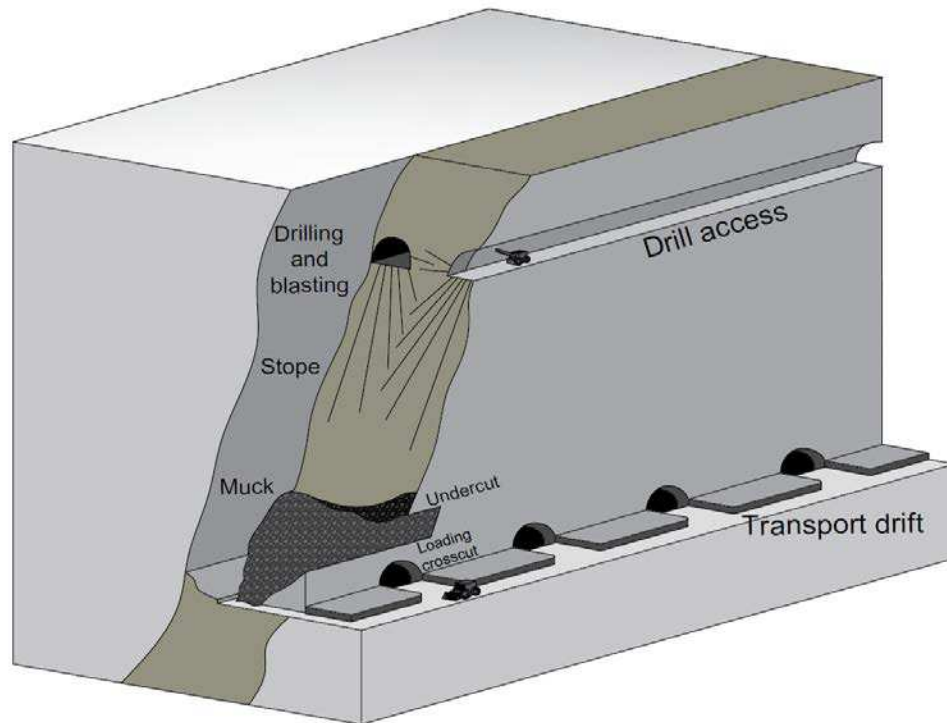


Figure 2.5. Long - hole blasting stopeing (Modified from Harraz 2010)

Shrinkage stoping, illustrated in Figure 2.6, is employed in steeply dipping ores due to its reliance on gravity to channel ore into chutes resembling inverted truncated cones. The orebody should have a regular shape, aside from its dip, which is crucial to prevent the lodging of loose ore. Typically, the orebody is divided into blocks and mined in slices to facilitate progressive mining between levels, thereby enhancing safety by ensuring miners always remain below the stope's top. These overhand stopes utilize broken ore both as a working platform and to support the stope walls. Given the increased volume of broken rocks compared to solid ones, partial muck removal is necessary as the stope advances, hence the term 'shrinkage.' Mining in shrinkage stoping progresses from bottom to top, similarly, to cut and fill mining, with broken ore left in place. Depending on rock conditions, the stope may be backfilled or left empty. The process continues until reaching the next level or a predetermined elevation, leaving horizontal crown pillars at the stope's top.

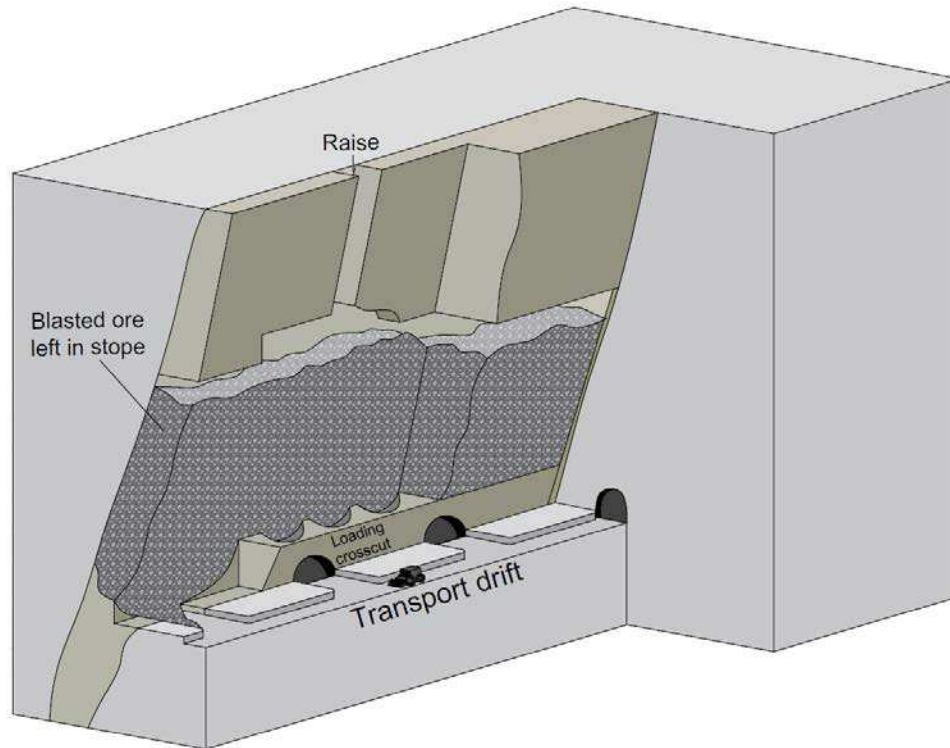


Figure 2.6. Shrinkage stoping (Modified from Hamrin 1997)

2.3. Dillution and Hangingwall overbreak

Dilution (Figure 2.7) refers to the excess material extracted from the stope by heavy machinery, which then enters the ore processing stream. Overbreak refers to the quantity of unstable rock that collapses into the stope from beyond the intended design boundary of a specific wall. This phenomenon occurs when rock masses outside the planned excavation shape become detached and intrude into the stope area. Overbreak can pose significant challenges in mining operations, leading to safety concerns, increased operational costs, and potential delays in production schedules. Hangingwall overbreak and dilution are often closely linked, particularly when the majority of overbreak from stope walls originates from non-economic material in the hangingwall. Unplanned HW dilution poses a significant expense for many open stope mining operations. Additionally, due to orebody boundary, waste material outside the orebody may be mined as planned dilution due to constraints in drilling and blasting operations. The impact of unplanned dilution on the mining cycle encompasses both direct and indirect costs. Direct costs

pertain to the physical handling of materials, while indirect costs are associated with downstream effects of instability (Pakalnis et al. 1996).

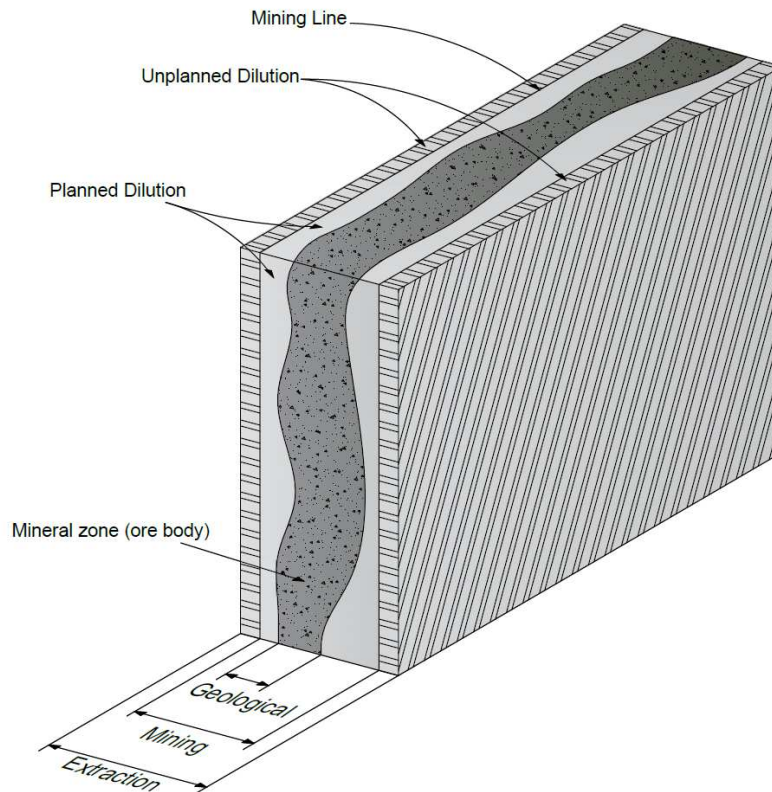


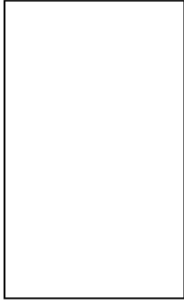
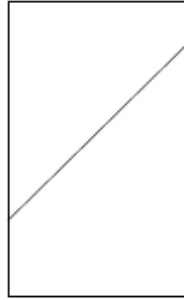
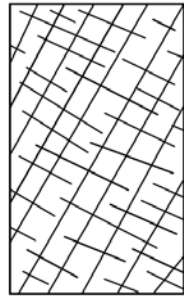
Figure 2.7. Planned and unplanned dilution (Modified from Scoble and Moss 1994)

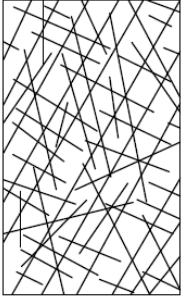
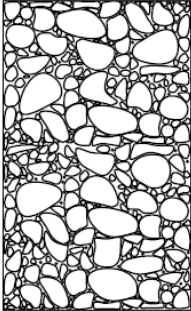
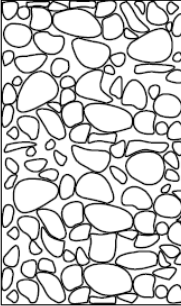
2.4. Rock mass characteristics

Information regarding the rock mass is typically obtained from core samples extracted from drill holes and observations made during underground and surface mapping. This data commonly includes details about the rock type, mineral composition, and characteristics of discontinuities or joints present in the rock mass. These discontinuities play a significant role in determining the overall strength of the rock mass. While the rock type and mineralogy can provide insights into the strength of individual components within the mass, understanding the properties of these discontinuities is crucial for assessing the mass's overall strength. Although there are laboratory tests available to estimate various strength parameters such as intact tensile and compressive

strength, as well as the shear strength of individual joints, it is often challenging or impossible to conduct tests on representative sized samples of heavily jointed rock masses commonly encountered in underground mines. This limitation contributes to a lack of understanding regarding how certain rock masses respond to loading conditions (Goodman 1989). A comparison of rock mass characteristics, testing methods, and theoretical understanding is presented in Table 2.1.

Table 2.1. Summary of rock mass characteristics (From Hoek et al. 2000)

	Description	Strength characteristics	Strength testing	Theoretical considerations
	Intact rock	Brittle, elastic and generally isotropic behavior	Triaxial testing of core specimens relatively simple and inexpensive and results are usually reliable	Behavior of elastic isotropic rock is adequately understood for most practical applications
	Intact rock with a single inclined discontinuity	Highly anisotropic, depending on shear strength and inclination of discontinuity	Triaxial tests difficult and expensive. Direct shear tests preferred. Careful interpretation of results required	Behavior of discontinuities adequately understood for most practical applications
	Massive rock with a few sets of discontinuities	Anisotropic, depending on number, orientation and shear strength of discontinuities	Laboratory testing very difficult because of sample disturbance and equipment size limitations	Behavior of complex block interaction in sparsely jointed rock masses poorly understood

	Heavily jointed rock masses	Reasonably isotropic, highly dilatant at low stress levels with particle breakage at high stress levels	Triaxial testing of representative samples extremely difficult because of sample disturbance	Behavior of interlocking angular pieces poorly understood
	Compacted rockfill or weakly cemented conglomerates	Reasonably isotropic, less dilatant and lower strength than in situ rock due to destruction of fabric	Triaxial testing simple but expensive due to large equipment required to accommodate samples	Behavior reasonably well understood from soil mechanics studies on granular materials
	Loose waste rock or gravel	Poor compaction and grading allow particle movement resulting in mobility and low strength	Triaxial or direct shear testing simple but expensive due to large size of equipment	Behavior of loosely compacted waste rock and gravel adequately understood for most applications

2.5. Rock mass classification systems

During the early stages of project feasibility and preliminary design, when detailed information about the rock mass and its stress and hydrological characteristics is limited, employing a rock mass classification system is highly beneficial. This may involve using the classification system as a checklist to ensure all relevant information has been considered or utilizing one or more classification systems to develop an understanding of the rock mass composition and properties. This helps in making initial estimations of support requirements, as well as the strength and deformation properties of the rock mass. It is important to note that while rock mass classification schemes are valuable tools, they do not replace more elaborate design

procedures that require detailed information on in situ stresses, rock mass properties, and planned excavation sequences. However, since such detailed information may not be available in the early project stages, rock mass classification schemes should be regularly updated and used alongside site-specific analyses as more information becomes accessible.

Various classification systems have been developed over more than 100 years, since Ritter (1879) attempted to formalize an empirical approach to tunnel design for determining support requirements. These systems were developed to characterize rock masses, serving as input data for assessing the stability of underground openings. These systems offer a systematic, quantitative, and repeatable method to evaluate the conditions of the rock mass for design purposes. Each classification system assigns weights to different parameters and may consider factors such as block size, joint strength, stress, and groundwater. These parameters are quantified as an index, which can be empirically correlated with engineering design needs like excavation span and ground support requirements.

Among the various classification systems developed, the "Rock Mass Rating" (RMR) introduced by Bieniawski (1978) and the "Tunnelling Quality Index" (Q) developed by Barton et al. (1974a), or their derivatives, have become most widely adopted in recent years. Both systems rely on the Rock Quality Designation (RQD), developed by Deere (1963) as a key parameter. A Q system and its variations were integrated into the Mathews stability graph method (Mathews et al. 1980) and Modified Stability Graph Method (Potvin 1988a), a commonly used approach for open stope design. Additionally, the RMR is applied for assessing back stability using the Span Graph Method (Lang et al. 1991) and for estimating hanging wall dilation in stopes (Pakalnis 1986).

2.5.1. Rock Quality Designation (RQD)

RQD, or Rock Quality Designation, serves as a straightforward tool for classifying rock masses and was originally developed by Deere (1963) to evaluate the quality of 54 mm diamond drill core samples. It plays an important role in various classification systems used in the field. RQD is computed by summing the length of intact core pieces exceeding 10 cm in length and dividing it by the total length of the drilled core within a specified interval (Figure 2.8). The resulting value, expressed as a percentage ranging from 0 to 100%, offers insights into the competence of the rock mass. Higher RQD values typically indicate better rock quality and greater

resistance to fracturing and fragmentation, whereas lower values suggest poorer quality and increased likelihood of instability.

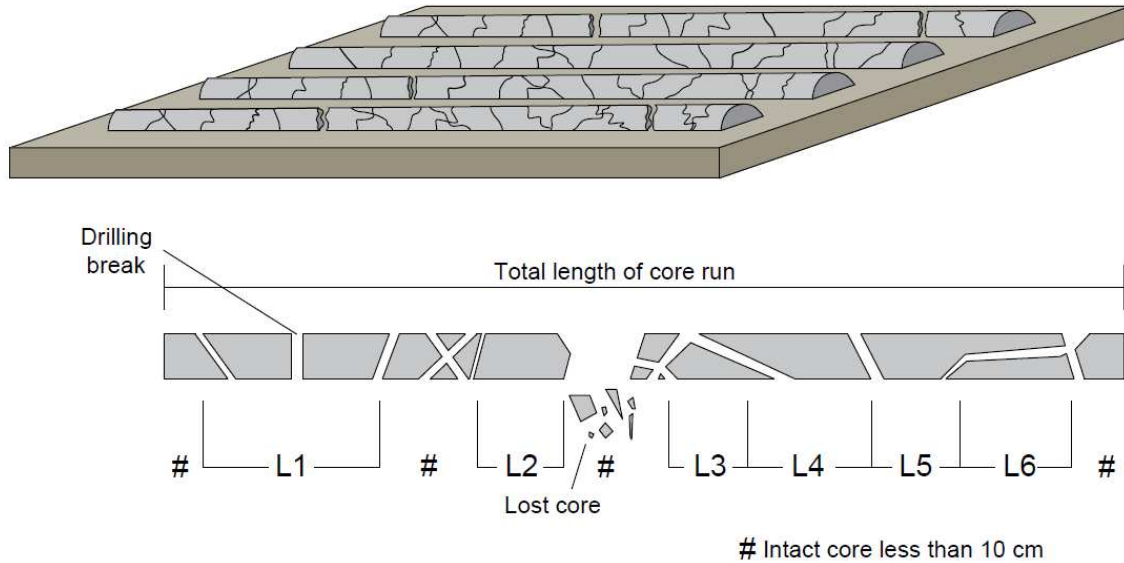


Figure 2.8. RQD procedure and measurement (Modified from Potvin and Nedin 2003)

The RQD value is calculated as:

$$RQD (\%) = 100 \cdot \frac{\text{Length of pieces in core} > 100 \text{ mm}}{\text{Length of a core}} \quad (2.1)$$

$$RQD (\%) = 100 \cdot \frac{L1+L2+L3+L4+L5+L6}{\text{Length of a core}} = (\%) \quad (2.2)$$

2.5.2. Rock mass characteristics (Q value)

Q value was first presented Barton et al. (1974a) of the Norwegian Geotechnical Institute (NGI) to evaluate the characteristics of rock mass and ground conditions. The purpose of calculating the Q value was to determine the requirement of support in mining excavations, tunnels and rock caverns. The NGI rock mass classification system is a function of six independent variables defining a constant, with each quotient in the equation representing a specific physical characteristic of the rock mass.

$$Q = \frac{RQD}{J_n} \cdot \frac{J_r}{J_a} \cdot \frac{J_w}{SRF} \quad (2.3)$$

Where:

RQD – Rock Quality Designation

J_n - Joint set number

J_r - Joint roughness number

J_a - Joint alteration number

J_w - Joint water reduction number

SRF – Stress Reduction Factor

A description of all the Q parameters is shown in Table 2.2. Table 2.3 is illustrating the range of values along with subjective definitions used for assessment purposes. The system has been modified for input into slope stability assessment methods and is referred to as modified Q', which is similar to original Q value, but it assumes that the joint water reduction factor (J_w) and the stress reduction factor (SRF) are equal to one. The formula is as follows:

$$Q' = \frac{RQD}{J_n} \cdot \frac{J_r}{J_a} \quad (2.4)$$

The first quotient ($\frac{RQD}{J_n}$) represents the structure of the rock mass, and the latter one ($\frac{J_r}{J_a}$) is representing roughness and frictional properties of the joint wall or filling material.

Table 2.2. Classification of all parameters used in the NGI Q classification system (From Hoek and Brown, 1996)

Description	Value	Note
1. ROCK QUALITY DESIGNATION (RQD)		
A. Very poor	0 – 25	1. Where RQD is reported or measured as ≤ 10 (including 0), a nominal value of 10 is used to evaluate Q.
B. Poor	25 – 50	
C. Fair	50 – 75	
D. Good	75 – 90	2. RQD intervals of 5, i.e. 100, 95, 90 etc are sufficiently accurate
E. Excellent	90 – 100	
2. JOINT SET NUMBER (Jn)		
A. Massive, no or few joints	0.5 – 1.0	
B. One joint set	2	
C. One joint set plus random	3	
D. Two joint sets	4	1. For intersections use $(3.0 \times J_n)$
E. Two joint sets plus random	6	
F. Three joint sets	9	2. For portals use $(2.0 \times J_n)$
G. Three joint set plus random	12	
H. Four or more joint sets, random, heavily jointed ‘sugar cube’, etc	15	
J. Crushed rock, earthlike	20	
3. JOINT ROUGHNESS NUMBER (Jr)		
<i>a. Rock wall contact and</i>		
<i>b. Rock wall contact before 10 cms shear</i>		
A. Discontinuous joints	4	
B. Rough and irregular, undulating	3	Add 1.0 if the mean spacing of the relevant joint set is greater than 3m.
C. Smooth, undulating	2	
D. Slickensided, undulating	1.5	
E. Rough or irregular, planar	1.5	
F. Smooth, planar	1.0	Jr = 0.5 can be used for planar, slickensided joints having lineations, provided the lineations are orientated for minimum strength.
G. Slickensided, planar	0.5	
<i>c. No rock contact when sheared</i>		

H. Zone containing clay minerals thick enough to prevent rock wall contact	1.0		
J. Sandy gravelly or crushed zone thick enough to prevent rock wall contact	1.0		
4. JOINT ALTERATION NUMBER (Ja)		ϕ_r (Approx.)	
<i>a. Rock wall contact</i>			
A. Tightly healed, hard, non-softening, impermeable	0.75	-	1. Values of ϕ_r , the residual friction angle, are intended as an approximate guide to the mineralogical properties of the alteration products, if present.
B. Unaltered joint walls, surface staining only	1.0	(25° - 35°)	
C. Slightly altered joint walls non-softening mineral coatings, sandy particles, clay-free disintegrated rock, etc	2.0	(25° - 30°)	
D. Silty-, or sandy-clay coatings, small clay fraction (non- softening)	3.0	(20° - 25°)	
E. Softening or low friction clay mineral coatings, i.e. kaolinite, mica. Also chlorite, talc, gypsum and graphite etc., and small quantities of swelling clays. (Discontinuous coatings, 1-2mm or less in thickness)	4.0	(8° - 16°)	
<i>b. Rock wall contact before cms shear.</i>			
F. Sandy particles, clay-free disintegrated rock etc.	4.0	(25° - 30°)	
G. Strongly over-consolidated, non-softening clay mineral fillings (continuous, < 5mm thick)	6.0	(16° - 24°)	

H. Medium or low over consolidation, softening, clay mineral fillings, (continuous, < 5mm thick)	8.0	(12° - 16°)	
J. Swelling clay fillings, i.e. Montmorillonite (continuous, <5mm thick). Values of J_w depend on percent of swelling clay-size particles, and access to water	8.0 – 12.0	(6° - 12°)	
<i>c. No rock wall contact when sheared.</i>			
K. Zones or bands of disintegrated	6.0		
L. or crushed rock and clay (see	8.0		
M. G,H and J for clay conditions)	8.0 – 12.0	(6° - 24°)	
N. Zones or bands of silty or sandy clay, small clay fraction, (non-softening)	5.0		
Q. Thick, continuous zones or	10.0 – 13.0		
P. bands of clay (see G, H and	13.0 – 20.0	(6° - 24°)	
R. J for clay conditions)			
5. JOINT WATER REDUCTION FACTOR (J_w)		Approx. water pressure (kgf/cm ²)	
A. Dry excavations or minor inflow, i.e. < 5 lit/min, locally	1.0	<1.0	1. Factors C to F are crude estimates. Increase J_w if drainage measures are installed.
B. Medium inflow or pressure, occasional outwash of joint fillings	0.66	1.0 - 2.5	2. Special problems caused by ice formation are not considered.
C. Large inflow or high pressure in competent rock with unfilled joints	0.5	2.5 – 10.0	

D. Large inflow or high pressure, considerable outwash of fillings	0.33	2.5 – 10.0	
E. Exceptionally high inflow or pressure at blasting, decaying with time	0.2 – 0.1	>10	
F. Exceptionally high inflow or pressure continuing without decay	0.1 – 0.05	>10	
6. STRESS REDUCTION FACTOR (SRF)			
<i>a. Weakness zones intersecting excavation, which may cause loosening of rock mass when tunnel is excavated.</i>			
A. Multiple occurrences of weakness zones containing clay or chemically disintegrated rock, very loose surrounding rock (any depth)			10.0
B. Single weakness zones containing clay, or chemically disintegrated rock (excavation depth < 50m)			5.0
C. Single weakness zones containing clay, or chemically disintegrated rock (excavation depth > 50m)			2.5
D. Multiple shear zones in competent rock (clay free), loose surrounding rock (any depth)			7.5
E. Single shear zones in competent rock (clay free), (depth of excavation < 50m)			5.0
F. Single shear zones in competent rock (clay free), (depth of excavation > 50m)			2.5
G. Loose open joints, heavily jointed or 'sugar cube' (any depth)			5.0
<i>b. Competent rock, rock stress problems</i>			
	σ_c/σ_1	σ_t/σ_1	
H. Low stress, near surface	>200	>13	2.5
J. Medium stress	200 - 10	13-0.66	1.0

K. High stress, very tight structure (usually favorable to stability, may be unfavorable for wall stability)	10-5	0.66-0.33	0.5-2
L. Mild rock burst (massive rock)	5-2.5	0.33-0.16	5-10
M. Heavy rock burst (massive rock)	<2.5	<0.16	10-20
<i>c. Squeezing rock, plastic flow of incompetent rock under the influence of high rock pressure</i>			
N. Mild squeezing rock pressure			5-10
O. Heavy squeezing rock pressure			10-20
<i>d. Swelling rock, chemical swelling activity depending upon presence of water</i>			
P. Mild swelling rock pressure			5-10
R. Heavy swelling rock pressure			10-20

Table 2.3. Classifications of rock mass quality based on Q (From Barton et al. 1974a)

Quality Index Q	Rock Mass Description
0.001 – 0.01	Exceptionally Poor
0.01 – 0.1	Extremely Poor
0.1 – 1	Very Poor
1 – 4	Poor
4 – 10	Fair
10 – 40	Good
40 – 100	Very Good
100 – 400	Extremely Good
400 – 1000	Exceptionally Good

2.5.3. Rock Mass Rating (RMR)

The Geomechanics Classification, also known as the Rock Mass Rating (RMR) system, was developed by Bieniawski (1976) primarily to assess the support requirements in civil engineering tunneling projects. The RMR index is determined by summing up five parameters within a range of 0 to 100, with an additional adjustment factor to accommodate joint orientation considerations. These parameters encompass the strength of intact materials, Rock Quality Designation (RQD), spacing and condition of joints, and groundwater. Bieniawski later refined the system (Bienawski 1989), assigning greater weight to joint condition values while reducing the significance of joint spacing values compared to the original model. Table 2.4 provides a breakdown of values associated with each variable from the 1976 classification system.

Table 2.4. RMR inputs (From Bienawski 1976)

Rock Strength (R1)			Groundwater (R5)		
<i>Description</i>	<i>MPA</i>	<i>Rating</i>	<i>Description</i>	<i>Inflow</i>	<i>Rating</i>
Very low	1-25	0 - 2	Dry	0	10
Low	25-50	4	Moist	>25L/min	7
Medium	50-100	7	under moderate water pressure	25- 125L/min	4
High	100-200	12	sever water problems	>125L/min	0
Very high	>200	15			
Joint Density (R2)			Joint Spacing (R3)		
<i>RQD</i>	<i>Joints/m3</i>	<i>Rating</i>	<i>Description</i>	<i>Distance</i>	<i>Rating</i>
90-100%	0-8	20	very wide	>3m	30
75-90%	8-12	17	wide	1-3m	25
50-75%	12-20	13	moderately close	.3-1m	20
25-50%	20-27	8	close	.05-.3m	10
<25%	>27	3	very close	<.05m	5

Joint Condition (R4) <i>Description</i>	<i>Rating</i>
Very Rough, not continuous, no separation, hard joint wall rock	25
Slightly rough surfaces, separation <1mm, hard joint wall rock	20
Slightly rough surfaces, separation < 1mm, soft joint wall rock	12
Slickensided or gouge, <5mm thick or Joints open 1 – 5mm	6
Soft gouge > 5mm thick or Joints open > 5mm	0

2.6. Summary

Dividing rock masses into zones with similar characteristics and utilizing rock mass classification systems is a valuable technique for characterizing the behavior of rock mass which can be challenging to describe. Initial steps involve gathering data on intact rock properties and joint conditions to establish geomechanical domains. This foundational information is used to describe rock mass domains based on factors like rock type, mineralogy, and joint characteristics. Rock mass classification systems build upon this data define critical parameters such as block size and shear strength, which play crucial roles in assessing rock mass stability. When integrated with considerations of opening geometry and stress conditions, these classification systems provide essential inputs for evaluating the stability of underground excavations.

References

- Barton N, Lien R, Lunde J (1974b) Engineering classification of rock masses for the design of tunnel support. *Rock Mech Felsmech Mec Roches* 6(4):189–236. <https://doi.org/10.1007/BF01239496>
- Beer G, Meck JL (1982) Design curves for roofs and hanging-walls in bedded rock based on “voussoir” beam and plate solutions. *Trans - Inst Min Met Sect U K* 91:Pages: A18-A22
- Bienawski ZT (1976) Rock Mass Classification in Rock Engineering. In *Proceedings of the Symposium on Exploration for Rock Engineering*
- Bienawski ZT (1989) *Engineering rock mass classification.*, John Wiley&Sons. New York
- Bieniawski ZT (1978) Determining rock mass deformability: experience from case histories. *Int J Rock Mech Min Sci Geomech Abstr* 15(5):237–247. [https://doi.org/10.1016/0148-9062\(78\)90956-7](https://doi.org/10.1016/0148-9062(78)90956-7)
- Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit* 30(7):1145–1159. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2)
- Brady BHG, Brown ET (2006) *Rock mechanics for underground mining*, 3. ed., repr. with corr. Springer, Dordrecht
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
- Capes GW (2009a) Open stope hangingwall design based on general and detailed data collection in unfavourable hangingwall conditions, NR62618 Ph.D. The University of Saskatchewan (Canada)
- Carbonell JG, Michalski RS, Mitchell TM (1983) Machine Learning: A Historical and Methodological Analysis. *AI Mag* 4(3):69–69. <https://doi.org/10.1609/aimag.v4i3.406>
- Clark LM (1998) Minimizing dilution in open stope mining with a focus on stope design and narrow vein longhole blasting. <https://doi.org/10.14288/1.0081111>

Deere DU (1963) Technical description of rock cores for engineering purposes. *Rock Mech Eng Geol* :1:16-22

Diederichs MS, Kaiser PK (1999) Tensile strength and abutment relaxation as failure control mechanisms in underground excavations. *Int J Rock Mech Min Sci* 36(1):69–96. [https://doi.org/10.1016/S0148-9062\(98\)00179-X](https://doi.org/10.1016/S0148-9062(98)00179-X)

Erdogan EG, Bozkurt KS, Yavuz M (2021) Grid Search Optimised Artificial Neural Network for Open Stope Stability Prediction. *Int J Min Reclam Environ* 35(8):600–617. <https://doi.org/10.1080/17480930.2021.1899404>

Evans WH (1941) The strength of underground strata., 50th edn. Trans. Instn. Min. Metall.

Goodman RE (1989) Introduction to rock mechanics, 2nd ed. Wiley, New York

Hackston A, Rutter E (2016) The Mohr–Coulomb criterion for intact rock strength and friction – a re-evaluation and consideration of failure under polyaxial stresses. *Solid Earth* 7(2):493–508. <https://doi.org/10.5194/se-7-493-2016>

Hadjigeorgiou J, Leclair J, Potvin Y (1995) An update of the stability graph method for open stope design. *CIM Rock Mech Strata Control Sess Halifax N S* :14–18

Hamrin H (1997) Guide to Underground Mining: Methods and Applications, Orebo Sweden: Atlas Copco

Harraz HZ (2010) Underground mining Methods. <https://doi.org/10.13140/RG.2.1.2881.1124>

Heidarzadeh S, Saeidi A, Rouleau A (2019) Evaluation of the effect of geometrical parameters on stope probability of failure in the open stoping method using numerical modeling. *Int J Min Sci Technol* 29(3):399–408. <https://doi.org/10.1016/j.ijmst.2018.05.011>

Henning JG, Mitri HS (2007) Numerical modelling of ore dilution in blasthole stoping. *Int J Rock Mech Min Sci* 44(5):692–703. <https://doi.org/10.1016/j.ijrmms.2006.11.002>

Hoek E, Brown ET (1980b) Underground excavations in rock, Rev. Institution of Mining and Metallurgy, London

Hoek E, Brown ET (1996) Underground excavations in rock, Rev. 1. ed., repr. The Institution of Mining and Metallurgy, London

Hoek E, Kaiser PK, Bawden WF (2000) Support of Underground Excavations in Hard Rock, 0 edn. CRC Press

Hosmer DW, Lemeshow S, Sturdivant RX (2013) Applied logistic regression, Third edition. Wiley, Hoboken, New Jersey

Hu H, Cao Y (2009) Numerical Simulation Modeling and Calculation Analysis on Stope Roof Stability under the Complex Geological Conditions in Deep Mining. In: 2009 International Conference on Engineering Computation. IEEE, Hong Kong, China, pp 175–177

Hutchinson DJ, Diederichs MS (1996) Cablebolting in Underground Mines. Bi Tech Publishers, Richmond.

Janitza S, Strobl C, Boulesteix A-L (2013) An AUC-based permutation variable importance measure for random forests. BMC Bioinformatics 14(1):119. <https://doi.org/10.1186/1471-2105-14-119>

Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp 1137–1143

Kumar HV, Kushwaha A, Sinha A (2014) Design of stoping parameters and support system for longhole stoping method by numerical modelling. Conf 5th Asian Min Congr Exhib

Lang B, Pakalnis R, Vongpaisal S (1991) Span design in wide cut and fill stopes at Detour Lake Mine. 93rd AGM-CIMM Calg Can

Marcot BG, Hanea AM (2021) What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? Comput Stat 36(3):2009–2031. <https://doi.org/10.1007/s00180-020-00999-9>

Mathews KE, Hoek E, Wyllie DC, Stewart, S (1980) Prediction Of Stable Excavation Spans for Mining at Depths below 1000 Metres in Hard Rock. Ottawa, ON

McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5(4):115–133. <https://doi.org/10.1007/BF02478259>

Nickson SD (1992) Cable support guidelines for underground hard rock mine operations. <https://doi.org/10.14288/1.0081080>

Obert L, Duval W (1967) *Rock mechanics and the design of structure in rock*. John Wiley & Sons, Inc., New York

Pakalnis R, Poulin R, Hadjigeorgiou J (1996) Quantifying the cost of dilution in underground mines. *Int J Rock Mech Min Sci Geomech Abstr* 33(5):A233. [https://doi.org/10.1016/0148-9062\(96\)80170-7](https://doi.org/10.1016/0148-9062(96)80170-7)

Pakalnis RT (1986) Empirical stope design at the Ruttan Mine, Sherritt Gordon Mines Ltd. <https://doi.org/10.14288/1.0081095>

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12(85):2825–2830

Potvin Y (1988a) Empirical open stope design in Canada. <https://doi.org/10.14288/1.0081130>

Potvin Y, Hudyma M (2000) Open Stope Mining in Canada. In *Metallurgy*, editor, *Proceedings of the MassMin2000*. Brisbane, Australia ed. Vol. n/a. Carlton, Victoria: Australasian Institute of Mining and Metallurgy. 2000. p. 661-674

Potvin Y, Hudyma M (1989) Open Stope Mining Practices in Canada. In: *Proceedings of the 1989 CIM annual general meeting*. Quebec City, Canada.

Potvin Y, Nedin P (2003) *Management of rockfall risks in underground metalliferous mines : a reference manual*. Minerals Council of Australia

Pu Y, Apel DB, Szmigiel A, Chen J (2019) Image Recognition of Coal and Coal Gangue Using a Convolutional Neural Network and Transfer Learning. *Energies* 12(9):1735. <https://doi.org/10.3390/en12091735>

Pu Y, Apel DB, Wang C, Wilson B (2018) Evaluation of burst liability in kimberlite using support vector machine. *Acta Geophys* 66(5):973–982. <https://doi.org/10.1007/s11600-018-0178-2>

- Pu Y, Szmigiel A, Apel DB (2020) Purities prediction in a manufacturing froth flotation plant: the deep learning techniques. *Neural Comput Appl* 32(17):13639–13649. <https://doi.org/10.1007/s00521-020-04773-2>
- Purwanto, Shimada H, Sasaoka T, Wattimena RK, Matsui K (2013) Influence of Stope Design on Stability of Hanging Wall Decline in Cibaliung Underground Gold Mine. *Int J Geosci* 04(10):1–8. <https://doi.org/10.4236/ijg.2013.410A001>
- Qi C, Fourie A, Du X, Tang X (2018) Prediction of open stope hangingwall stability using random forests. *Nat Hazards* 92(2):1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>
- Qi C, Fourie A, Ma G, Tang X, Du X (2018c) Comparative Study of Hybrid Artificial Intelligence Approaches for Predicting Hangingwall Stability. *J Comput Civ Eng* 32(2):04017086. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000737](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000737)
- Ritter W (1879) *Die statik der tunnelgewölbe*. Germany
- Rosenblatt F (1957) *The Perceptron, a Perceiving and Recognizing Automaton Project Para*. Cornell Aeronautical Laboratory
- Saadaari FS, Mireku-Gyimah D., Olaleye BM (2020) Development of a Stope Stability Prediction Model Using Ensemble Learning Techniques - A Case Study. *Ghana Min J* 20(2):18–26. <https://doi.org/10.4314/gm.v20i2.3>
- Santos AEM, Amaral TKM, Mendonça GA, Silva D de FS da (2020a) Open stope stability assessment through artificial intelligence. *REM - Int Eng J* 73(3):395–401. <https://doi.org/10.1590/0370-44672020730012>
- Scoble MJ, Moss A (1994) Dilution in underground bulk mining: implications for production management. *Geol Soc Lond Spec Publ* 79(1):95–108. <https://doi.org/10.1144/GSL.SP.1994.079.01.10>
- Sofianos AI (1996) Analysis and design of an underground hard rock voussoir beam roof. *Int J Rock Mech Min Sci Geomech Abstr* 33(2):153–166. [https://doi.org/10.1016/0148-9062\(95\)00052-6](https://doi.org/10.1016/0148-9062(95)00052-6)

Suorinen FT, Kaiser PK, Tannant DD (2001) Likelihood statistic for interpretation of the stability graph for open stope design. *Int J Rock Mech Min Sci* 38(5):735–744. [https://doi.org/10.1016/S1365-1609\(01\)00033-8](https://doi.org/10.1016/S1365-1609(01)00033-8)

Tishkov M (2018) Evaluation of caving as a mining method for the Udachnaya underground diamond mine project. In: *Proceedings of the Fourth International Symposium on Block and Sublevel Caving*. Australian Centre for Geomechanics, Perth, pp 835–846

Villaescusa E (1996) Excavation design for bench stoping at Mount Isa. *Trans Inst Min Metall Sect Min Ind*

Visa S, Ramsay B, Ralescu A, Knaap E (2011). Confusion Matrix-based Feature Selection. *Proceedings of The 22nd Midwest Artificial Intelligence and Cognitive Science Conference 2011*, Cincinnati, Ohio, USA.

CHAPTER 3: LITERATURE REVIEW ON STOPE STABILITY ASSESSMENT METHODS

The evaluation of open stope stability requires a comprehensive approach that combines data from various sources, including rock mass classification, numerical modeling, on-site observations, and historical case studies. Given the challenge of explicitly defining rock masses, underground observations of rock mass behavior are correlated with parameters believed to influence stability. These parameters are essential to both analytical and empirical assessment techniques. Analytical methods rely on theoretical frameworks to identify parameters affecting opening stability, while empirical methods, such as rock mass classification, are based on parameters known to define the explicit factors influencing stability. Through the compilation of extensive case histories, confidence is established in the correlation between empirical parameters and opening stability. In this chapter, the various popular design methodologies available for HW design in open stopes are presented.

3.1. Analytical Methods

Analytical design methodologies involve the evaluation of stress-induced (non-dynamic) failure, gravity-induced failure, or a combination of both. A thorough understanding of the relationship between stress conditions and rock properties create a significant aspect of the design process. When analyzing gravity-wedge failures, important considerations include the shear strength along joints in relation to the normal stress clamping the wedge, as well as the orientations of the joints. Whereas, when assessing the behavior of intact rock subjected to varying load conditions, understanding the magnitude and orientation of principal stresses acting on the rock, alongside its strength, becomes necessary. However, it is worth noting that certain assumptions associated with analytical methods may limit their applicability in practice.

During underground excavation, stress is relieved from the exposed rock face, leading to the relaxation of the surrounding rock near the opening. Depending on factors such as the local stress environment, rock mass classification, and orientation of bedding and structures relative to the opening, various gravity-induced failures of the rock may occur. Analyses of gravity-driven failures include Kinematic Wedge Failure, Beam and Plate Buckling Analysis, and Voussoir Beam Analysis. Moreover, excavation leads to stress concentration in certain areas around the opening. Stress-driven failure analyses employ criteria such as Mohr-Coulomb and Hoek-Brown to assess the relationship between rock mass strength and induced stress conditions.

3.1.1 Kinematic, Beam Failure, and Plate Buckling Analysis

The analysis of kinematic failures relies on structural mapping to identify probable wedge geometries that could develop within an excavation. Such failures typically occur where the intersection of joints creates conditions favorable for a block of material to detach from a stope wall, primarily when the force of gravity exceeds the resisting forces, including joint cohesion and frictional strength. Major faults often cause a larger-scale failures within stope walls, whereas joint intersections may create smaller failures. When data is accessible for a stope wall, wedge failure analysis is employed in open stope design.

Beam failure and plate buckling analysis, explained by Obert and Duval (1967), have historically been utilized to assess the stability of tunnel roofs in laminated rock formations. However, the method's reliance on assumptions of continuous, homogeneous, isotropic, and linear elastic conditions makes it challenging to apply in underground mining, where discontinuities significantly influence stability of the opening. Despite equations available to estimate plate stability based on parameters like gravity loading, thickness, length, width, and elastic deformation coefficients, the assumption that the plate can mobilize tensile stresses makes this approach unsuitable for jointed rock masses. In scenarios where joints are perpendicular to the bedding, the rock mass cannot develop tensile strength, thus limiting the applicability of plate analysis in open stope excavation designs.

3.1.2. Voussoir Beam Analysis

The Voussoir beam model represents an advancement from plate and beam analyses, particularly in situations where joints are nearly perpendicular to laminations. Evans (1941) introduced a model to address the stability of a jointed beam, which considers the transfer of vertical gravity loading to a nearly horizontal thrust across individual blocks and onto the opening abutment. The method was later modified by several authors (Beer and Meck 1982; Sofianos 1996; Diederichs and Kaiser 1999; Brady et al. 2006).

The Voussoir beam theory is a technique for evaluating the stability of a layered roof by leveraging the arching mechanism, as shown in Figure 3.1. This theory has found extensive application in both the design and stability assessment of underground excavations. It operates on the principle of distributing loads across successive blocks or "voussoirs," mimicking the structural behavior of an arch. By considering the interaction between the individual layers and their ability to support the overlying weight, the Voussoir beam theory provides insights into the overall stability of the roof structure in underground opening.

Design graphs have been derived from the theory to analyze the stability of plates and beams. However, the method has some limitations mentioned by Diederichs and Kaiser (1999). When considering open stope design for poor quality rock masses, certain constraints need to be acknowledged. For instance, it may not be suitable for rock masses with more than three joint sets

and less than 50% Rock Quality Designation (RQD). Additionally, the validity of design charts may be compromised in the presence of low to mid-angle jointing. Moreover, the method only accounts for the first lamination, which may not fully represent the complexities of the rock mass.

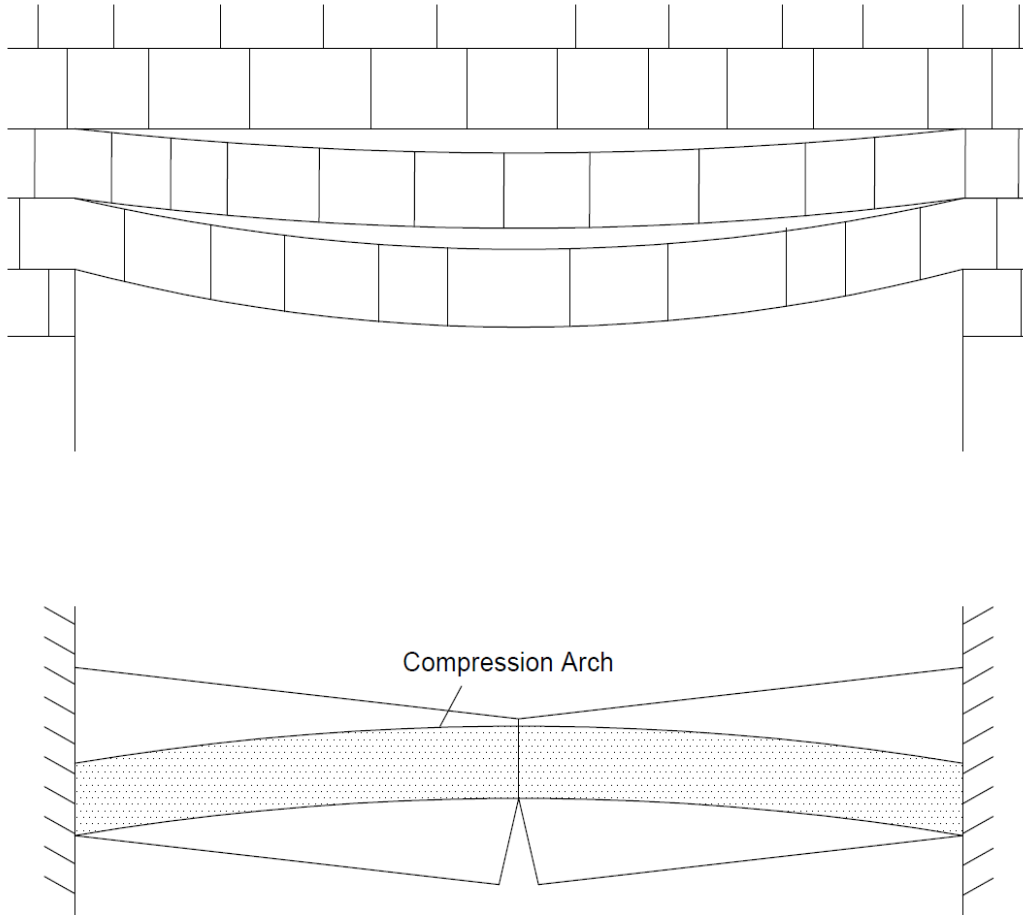


Figure 3.1. Compression arch in a deflecting beam (From Diederichs and Kaiser 1999)

3.1.3. Mohr - Coulomb Criterion

The Mohr-Coulomb failure criterion is a well-known, straightforward, and widely employed criteria for rock failure analysis. It considers how the strength of rock increases with increasing confinement, and that failure happens at specific combinations of the maximum and minimum principal stresses, disregarding the influence of the intermediate principal stress on failure. Additionally, it establishes failure criteria solely based on the stress state, neglecting any insights into strain or deformation mechanisms leading to failure. This criterion is widely adaptable to various geomechanical modeling challenges and frequently employed for problems involving polyaxial loading, particularly when information regarding the behavior of specific rocks under such stress conditions is not established (Hackston and Rutter 2016). Illustrated in Figure 3.2 with a cutoff for tensile stress, this criterion outlines failure in terms of the maximum principal stress, σ_1 , which can be expressed as:

$$\sigma_1 = \sigma_{UCS} + \sigma_3 \tan\left(45 + \frac{\Phi}{2}\right) \quad (3.1)$$

Where:

σ_{UCS} - Unconfined compressive strength

Φ - Angle of internal friction

At exceedingly high confinements, the maximum stress at failure becomes unrealistic due to the formation of a linear failure surface. The Hoek and Brown (1980a) failure criterion aims to address this issue by flattening the failure surface and considering the jointed characteristics of the rock mass.

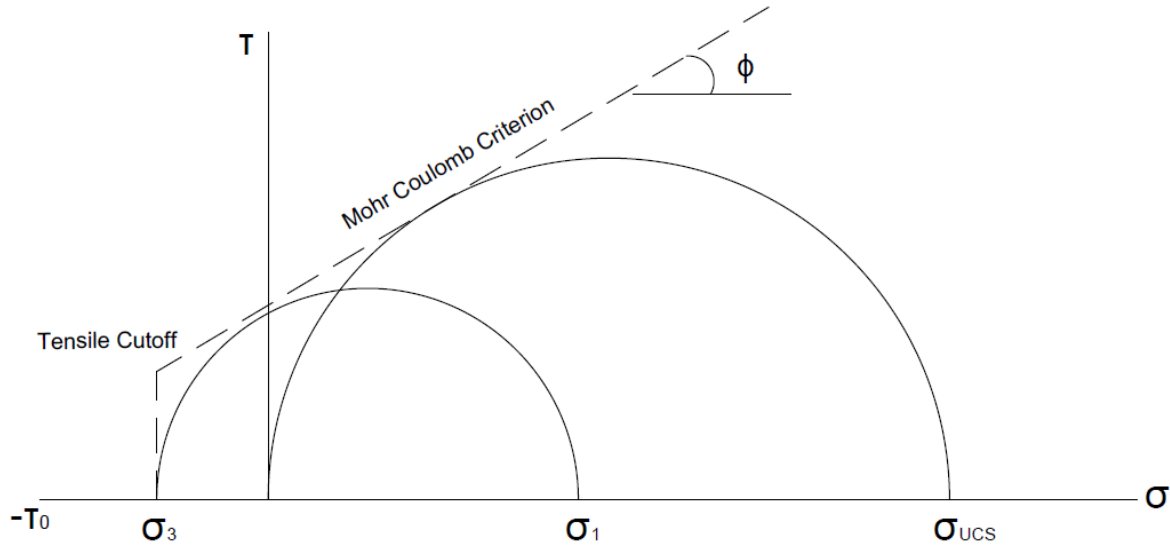


Figure 3.2. Mohr Coulomb failure criterion (From Goodman 1989)

3.1.4. Hoek and Brown failure criterion

Hoek and Brown established an empirical correlation for the maximum failure stress across various levels of confinement. This criterion incorporates two parameters, denoted as m and s , and is also based on the unconfined compressive strength (UCS) of the rock to characterize its strength properties. The values of m and s are determined based on the properties or classification of the rock mass, as well as its type. The equation formulated to calculate the maximum load at failure is expressed as:

$$\sigma_1 = \sigma_3 + (m\sigma_3\sigma_{UCS} + s\sigma_{UCS}^2)^{\frac{1}{2}} \quad (3.2)$$

Where the parameters m and s represent constants that vary according to the type of rock and the classification of the rock mass. An advantage of this failure criterion lies in its correlation with rock classification values such as RMR and Q , which consider the jointed characteristics of the rock mass.

3.2. Empirical Methods for open stope design

Considerable effort has been dedicated to the advancement and utilization of empirical design techniques for open stope design (Mathews et al. 1980; Pakalnis 1986; Potvin 1988; Nickson 1992; Clark 1998). These methods utilize key features identified through observations as essential for design, incorporating data from rock mass characterization and opening geometry to analyze and compare previous outcomes. Essential inputs for these methods include gravity, stress, rock strength, rock mass classification, and joint orientation. Empirical methods play a crucial role in rock mechanics due to the complexity of rock masses, which lack predetermined properties for standard theoretical analysis. While individual blocks and joints can be evaluated in laboratories, assessing their interaction on a wide mining scale is highly challenging. Empirical methods in open stope design can be categorized into stability methods, focusing on qualitative assessments, and dilution methods, which quantify the extent of dilution.

3.2.1. Design tools for open stope

Over the course of the last decades, a range of empirical design graphs has emerged as practical tools for design applications in mining industry. These design graphs establish a correlation between the geometry of the opening, the quality of the rock mass, and the prevailing loading conditions to assess the stability of stope walls.

The Mathews stability graph design method was specifically established for open stope surfaces in deep underground mining excavations. It was developed and presented for the first time in 1980 (Mathews et al. 1980). This widely used method, relates two calculated factors: shape factor (SF) or hydraulic radius (HR) and stability number (N), determining zones of stability. The major principal theory that stands behind the Mathews stability graph, is that the dimensions of an excavation surface can be associated with the rock mass conditions and indicate either instability or stability of the opening.

This approach gained increased acceptance following Potvin's (1988) collection of a more extensive database and the refinement of input factors, leading to the development of a modified

stability number (N'). This modification effectively reduced the substantial transition zone between stable and unstable behavior, further enhancing the method's applicability.

Stability Number N' was developed specifically for designing span dimensions and support, and it yields the physical conditions of the stopes. To calculate Stability Number certain rating systems are being applied. N' is defined as follows:

$$N' = Q' \cdot A \cdot B \cdot C \quad (3.3)$$

Where:

Q' – Modified Q value

A – Rock stress factor

B – Joint orientation adjustment factor

C – Surface orientation factor

3.2.1.1. Rock Stress Factor A

The Stress Reduction Factor (SRF) that was proposed in the Norwegian Geotechnical Institute (NGI) classification is being replaced by Factor A , which quantifies more accurately the effect of stresses that are acting on the open stopes exposed surface. This factor is a function represented by a ratio of the intact rock strength, determined by Uniaxial Compressive Strength (USC) test, and the induced stress, which is a maximum tangential stress that is acting parallel to the exposed surface at the boundary of a stope.

To determine Uniaxial Compressive Strength (UCS) laboratory tests need to be conducted while the induced stress should be estimated by numerical modeling to obtain best results. Mathews has developed two graphs (Figures 3.3. and 3.4.) that are describing the approximate stresses inducing in the walls and roof of mining openings. Then factor A is being determined based on the third graph (figure 3.5).

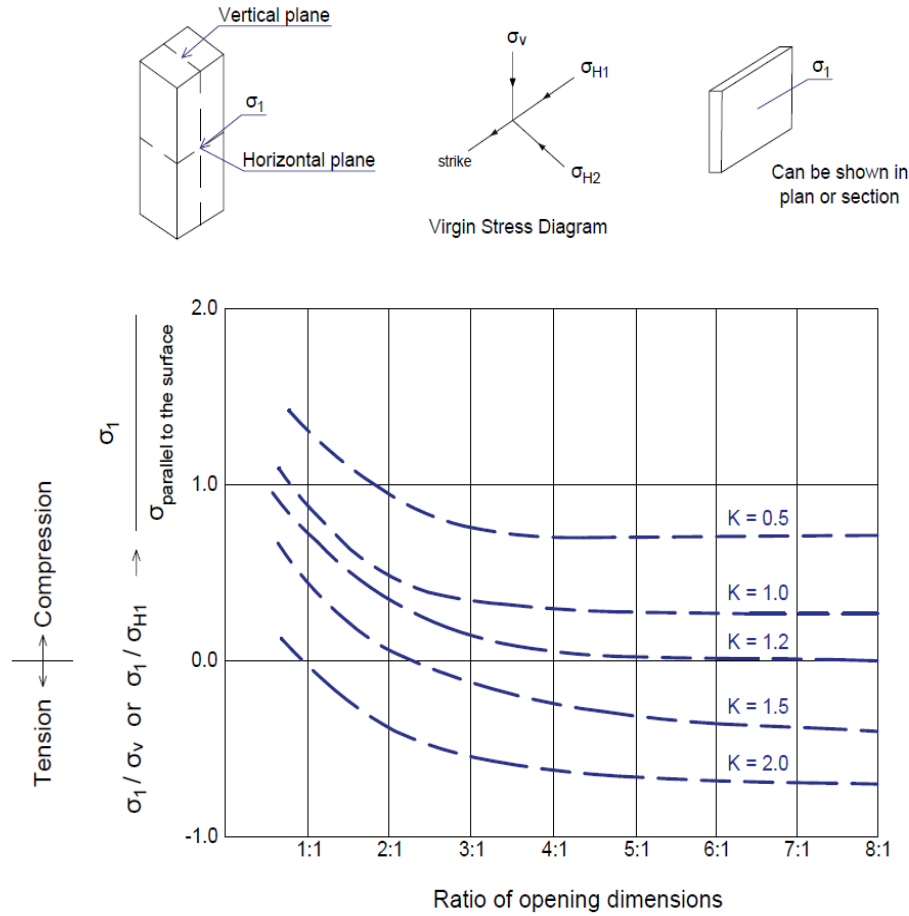


Figure 3.3. Graph of the stresses induced on the major surface of a stope vs. the ratio of opening dimensions. (Modified from Potvin, 1988)

Where:

σ_1 – Induced stress

σ_v – Vertical virgin stress

σ_{H1} – Horizontal virgin stress on strike

σ_{H2} – Horizontal virgin stress normal to strike

Horizontal plane: $K = \sigma_{H2} / \sigma_{H1}$

Vertical plane: $K = \sigma_{H2} / \sigma_v$

} $\frac{\sigma_{\text{normal to the surface}}}{\sigma_{\text{parallel to the surface}}}$

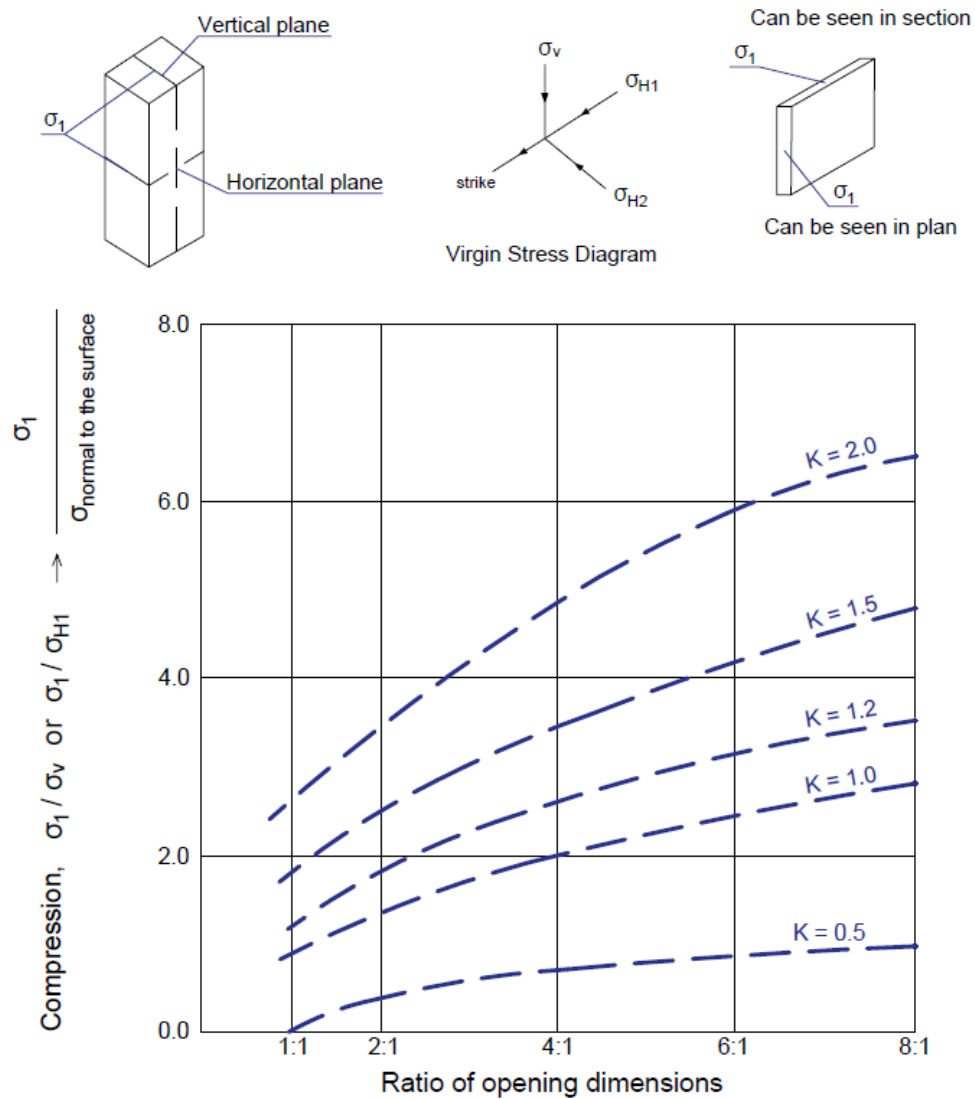


Figure 3.4. Graph of the stresses induced on the minor surface of a stope vs. the ratio of opening dimensions. (Modified from Potvin, 1988)

Where:

σ_1 – Induced stress

σ_v – Vertical virgin stress

σ_{H1} – Horizontal virgin stress on strike

σ_{H2} – Horizontal virgin stress normal to strike

$$\left. \begin{array}{l} \text{Horizontal plane: } K = \sigma_{H2} / \sigma_{H1} \\ \text{Vertical plane: } K = \sigma_{H2} / \sigma_V \end{array} \right\} \frac{\sigma_{\text{parallel to the surface}}}{\sigma_{\text{normal to the surface}}}$$

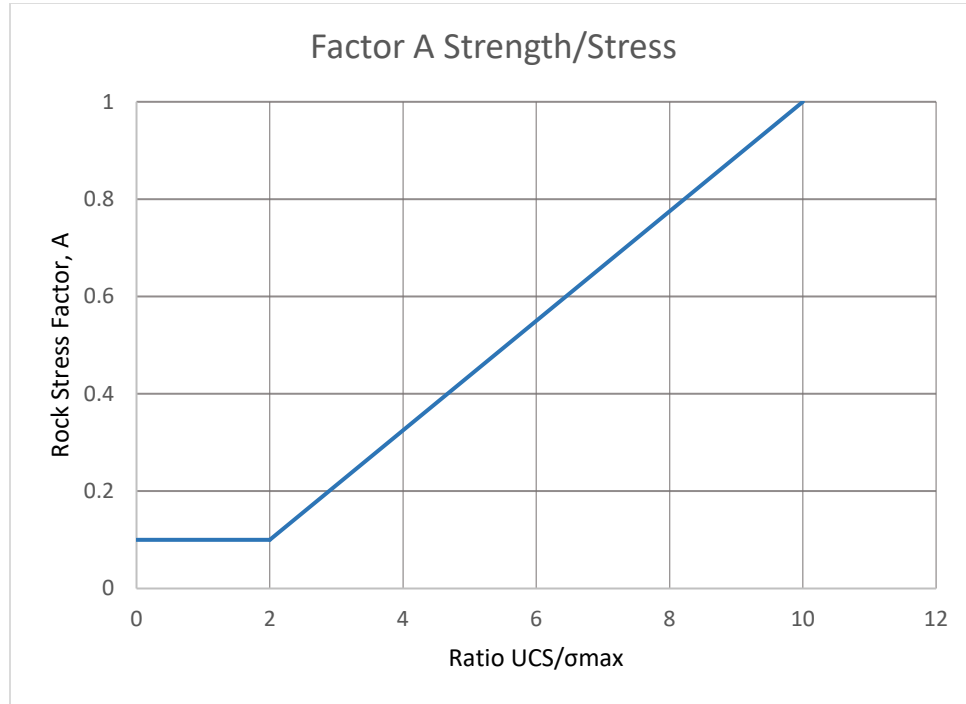


Figure 3.5. Graph for the estimation of Rock Stress Factor A. (Modified from Potvin, 1988)

3.2.1.2. Joint orientation adjustment Factor B

Factor B accounts for the orientation of the geological structures (joint sets) with respect to the plane that is investigated. It is determined by the angle of intersection between the surface that was exposed and the structure that is mostly predominant, shown on Figure 3.6. Factor B can be also determined using the graph on Figure 3.7.

In many cases a rock mass has multiple joint sets, in that situation Factor B should be determined for each joint set and then the lowest value should be used to calculate the stability number N.



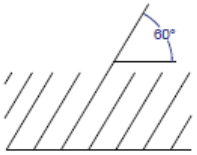
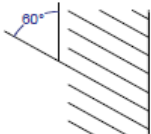
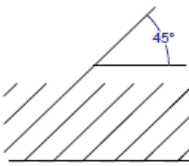
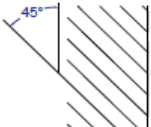


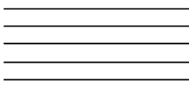
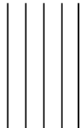
Orientation of Roof	Factor B	Orientation of Wall
	1.0	
	0.8	
	0.4	
	0.3	
	0.5	

Figure 3.6. Sketch for the determination of B factor. (Modified from Mathews et al, 1980)

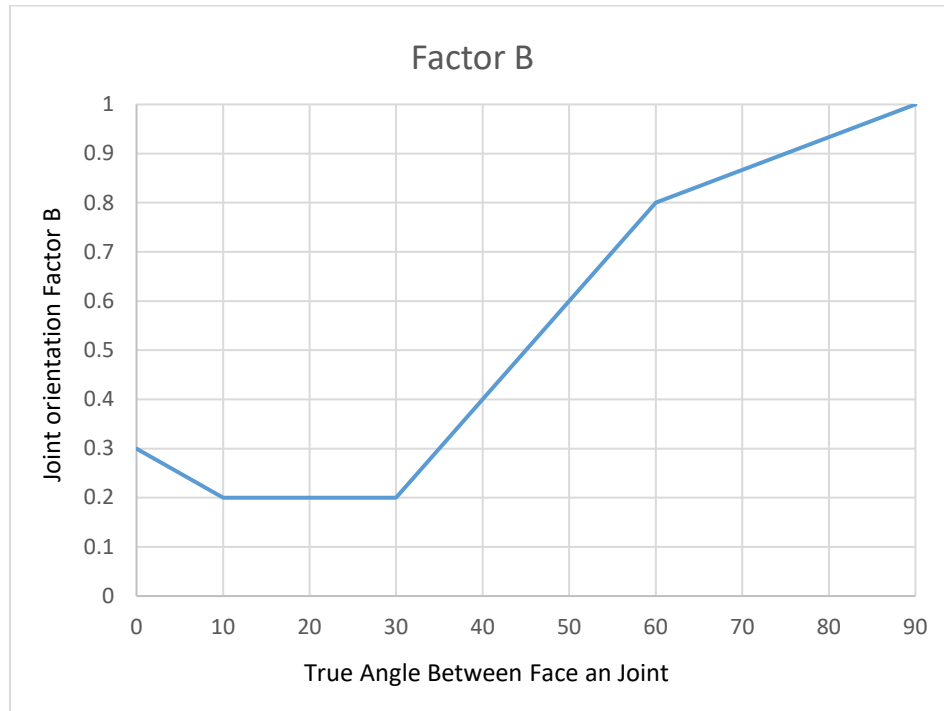


Figure 3.7. Graph for the determination of Rock Stress Factor B. (Modified from Hutchinson and Diederichs 1996)

3.2.1.3. Surface orientation factor

Factor C is the factor describing the influence of surface inclination. Slope backs create greater stability challenges compared to walls caused by the effect of gravity. Barton (1974) suggested that the rock quality within a tunnel wall could theoretically be enhanced up to five times when compared to a horizontally oriented roof. Mathews et al. (1980) proposed that a vertical open slope wall is eight times the stability of a horizontal roof, considering the tolerable level of minor instability in non-entry mining scenarios. Figures 3.8 and 3.9 are used in determining factor C, which quantifies the escalating risk of instability with a surface becoming more horizontal (Potvin 1988).

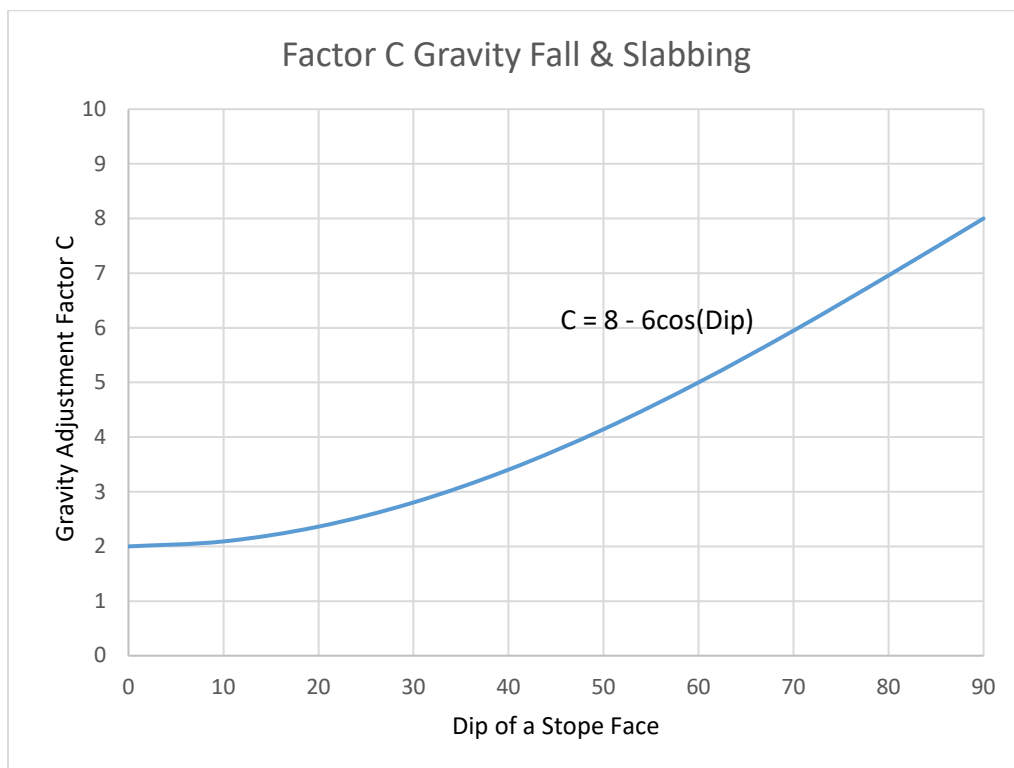
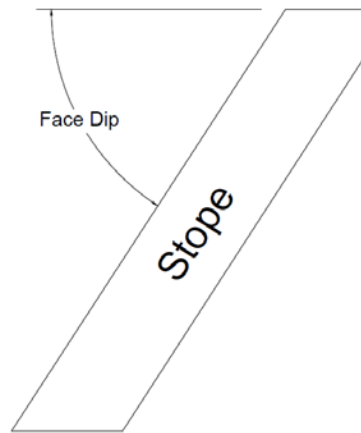


Figure 3.8. Determination of Stability Factor C Gravity Fall & Slabbing. (Modified from Hutchinson and Diederichs 1996)

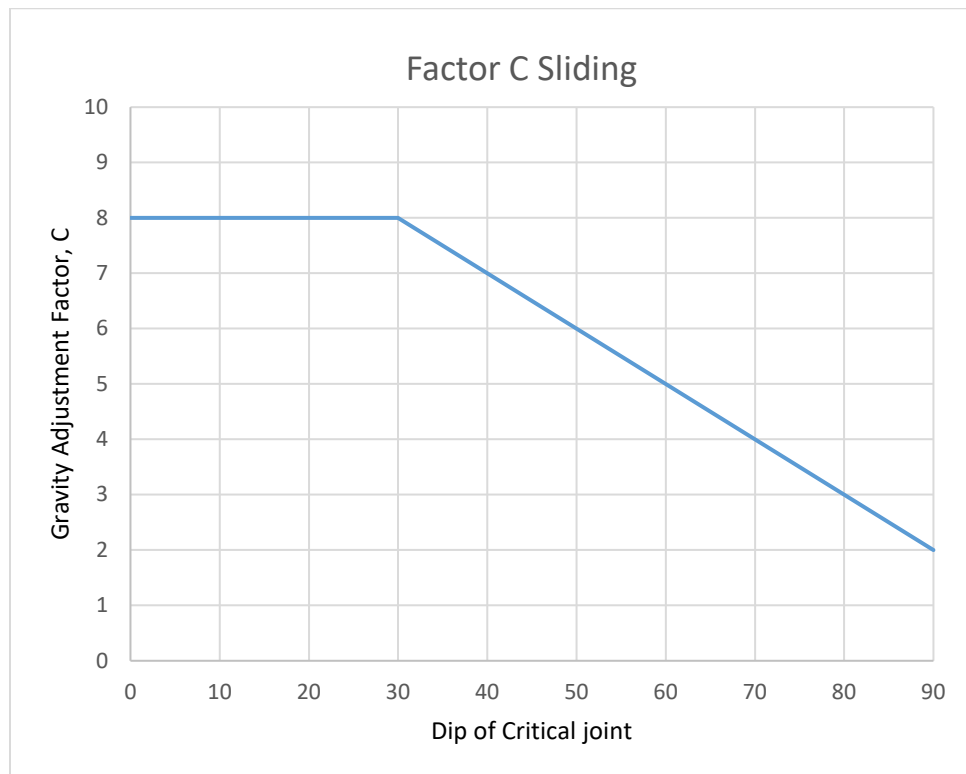
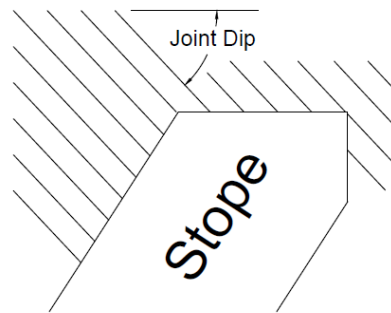


Figure 3.9. Determination of Stability Factor C Sliding. (Modified from Hutchinson and Diederichs 1996)

3.2.1.4. Shape Factor (Hydraulic Radius HR)

The crucial factor for the successful assessment of opening stability in underground mining is the shape factor, which relates the dimensions of the opening and is commonly referred to as the hydraulic radius (HR). The hydraulic radius is a critical geometric parameter that characterizes the shape of the opening and is essential in determining its stability.

The term hydraulic radius is commonly understood as the ratio of the area of exposure of the hanging wall to its perimeter. In the context of inclined stopes, where the stope is not in a vertical position, the most critical aspect for calculating the HR is the exposure of the hanging wall (Figure 3.10). The calculation of HR takes into account the spans of the stopes along the dip (h) and along the strike (w) (Tishkov 2018).

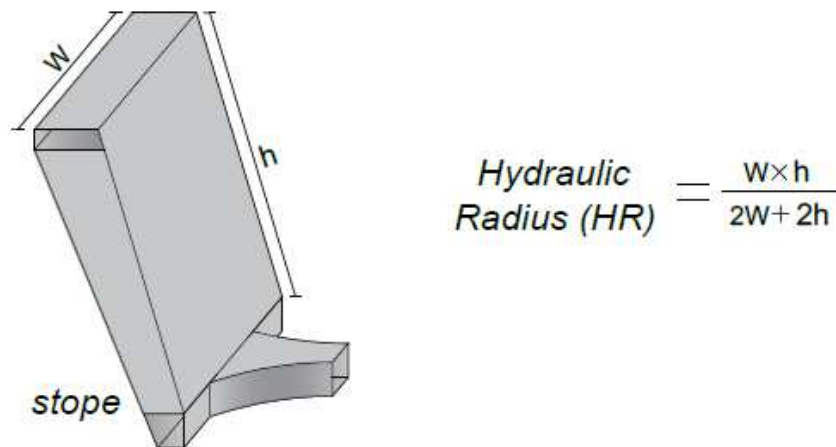


Figure 3.10. Determination of Hydraulic Radius

3.2.2. Stability graphs

Several authors have approached the stability graph method presented by Matthews (Figure 3.11) and Potvin (3.12), proposing various adjustments and refinements to enhance its applicability and accuracy.

Nickson (1992) introduced an additional zone in the stability graph method, delineating stable spans achievable with the implementation of pattern cable bolting (Figure 3.13). Nickson was first in statistically determining the boundaries' positions. He employed discriminant analysis on the three-dimensional multivariate stability database and utilized Mahalanobis' distance to partition the data into two distinct groups. Nickson established a linear separation between stable and caved unsupported histories through a logarithmic transformation.

Stewart and Forsyth (1993) deliberated on various rock mass classification systems applicable in mining, noting the limitations of each method. They critiqued Potvin's approach for its perceived precision, particularly its division into only two stability zones separated by a transition zone. Drawing from their practical experience, they proposed a qualitative redefinition of the stability graph, with four stability zones, correlating them with percentage dilution estimates, as illustrated in Figure 3.14. Hadjigeorgiou et al. (1995) made adjustments to the graph to enhance its accuracy in representing the design considerations for large hanging walls in high-quality rock conditions.

Hutchinson and Diederichs (1996) identified several limitations of the Modified Stability Graph as a design approach. Firstly, the method assumes that the wall under consideration is fully bounded, making it unsuitable for situations where adjacent stopes are not tightly filled or for intersections. Secondly, discrete wedges, shears, and delamination zones are not considered, and these are better assessed using analytical methods. Lastly, corners or bulges in the stope wall are not factored into the hydraulic radius calculation and can significantly impact stope stability. Villaescusa (1996) developed a bench stability method tailored for the strongly laminated stope hanging walls. This method shares a similar design principle with the Modified Stability Graph, where a stability parameter is compared to a shape factor. However, the input data for this rating was customized to suit particular conditions. The input primarily relies on factors such as the

number of bedding plane breaks per meter, the presence and continuity of joint sets, with minor considerations given to stress normal to the orebody and blasting practices.

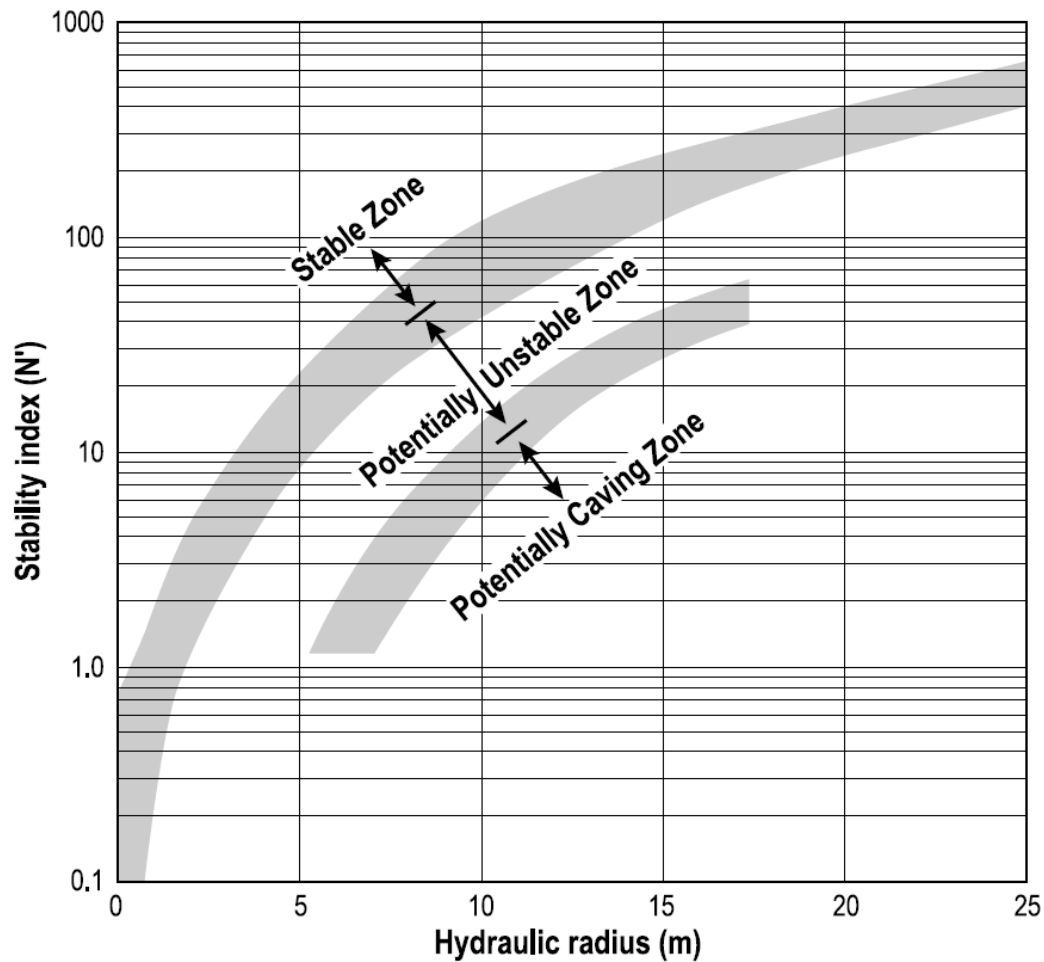


Figure 3.11. Mathews Stability Graph (After Mathews et al., 1981, From Capes 2009)

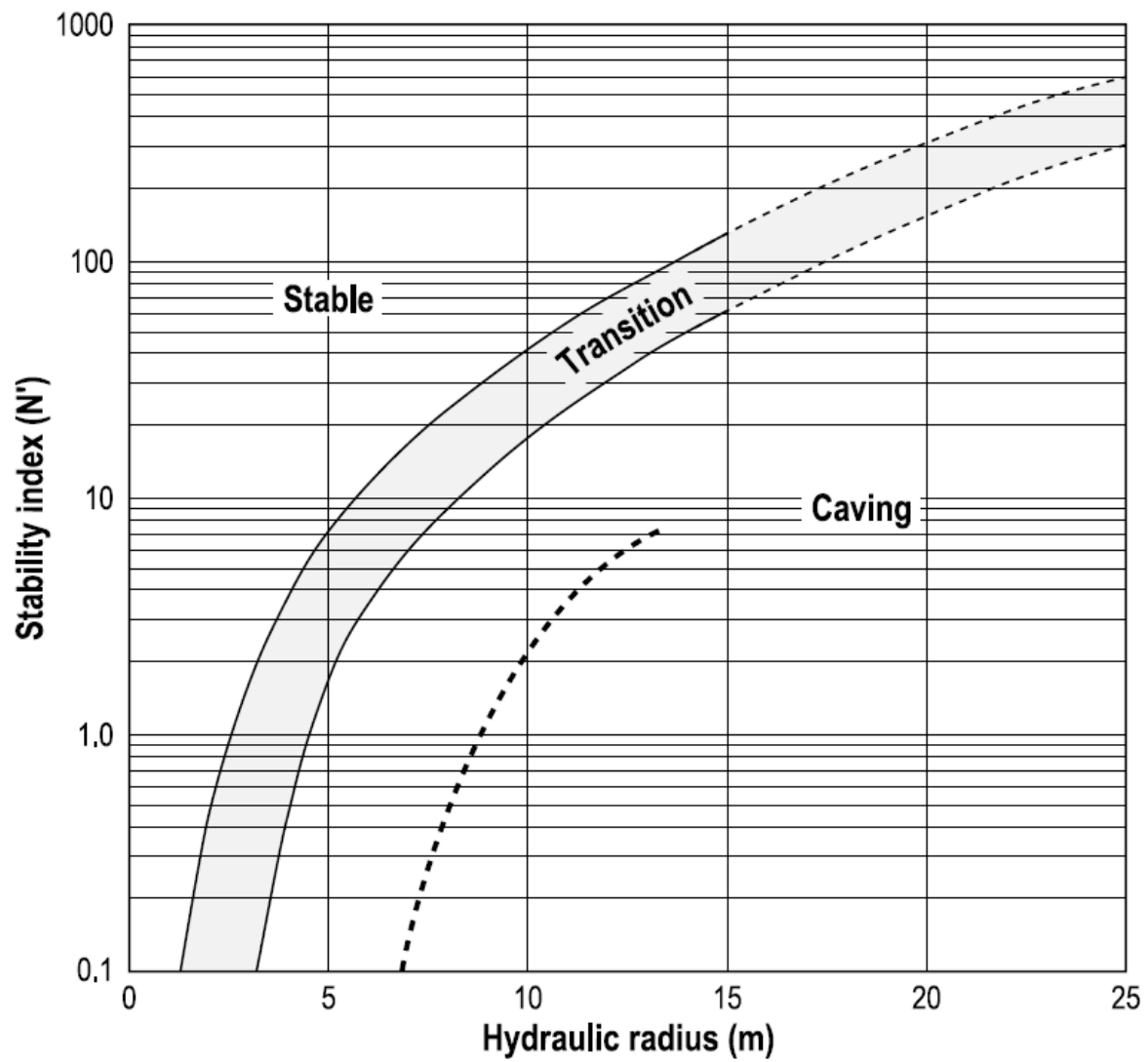


Figure 3.12. Modified Stability Graph (After Potvin, 1988, From Capes 2009)

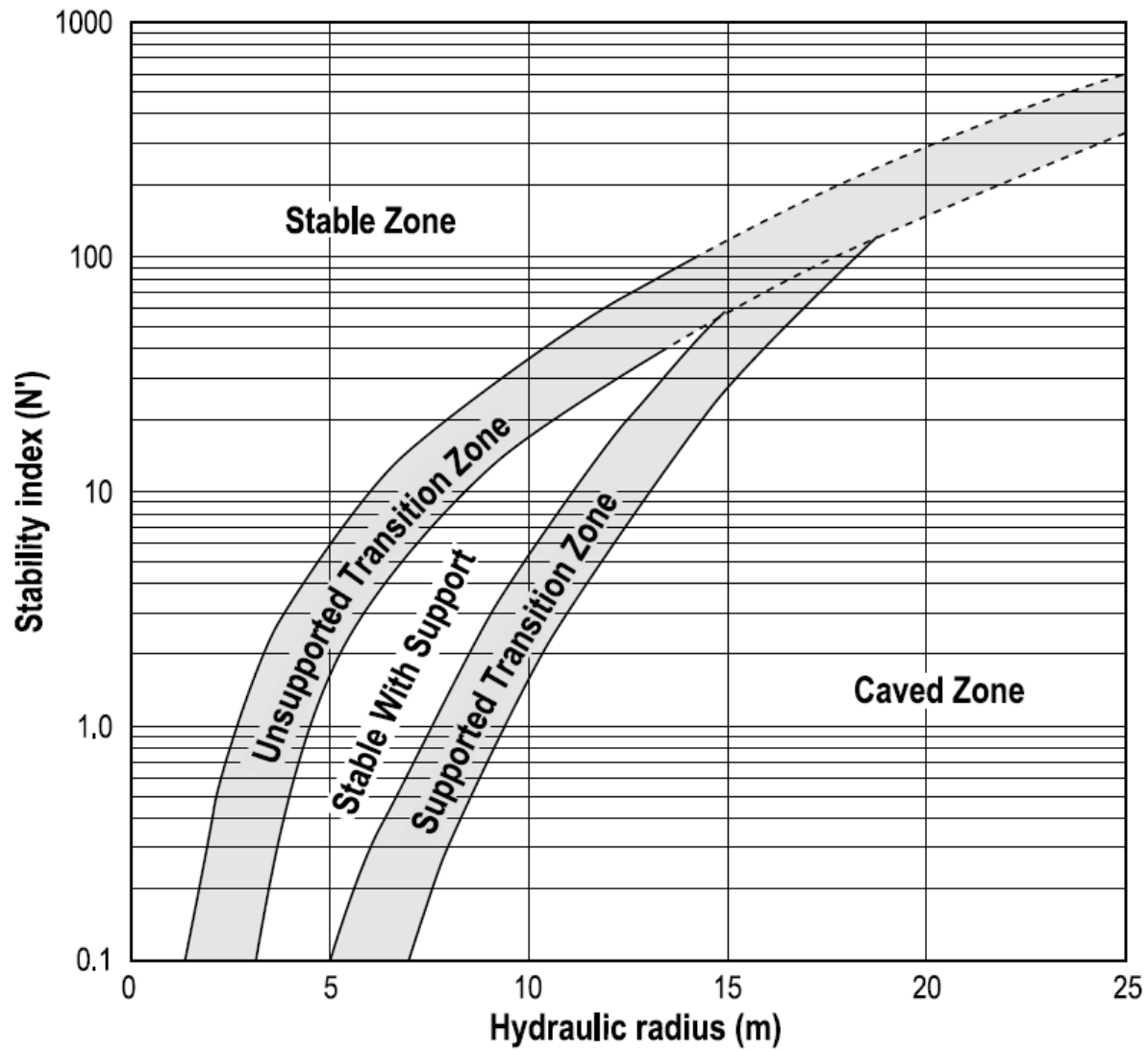


Figure 3.13. Modified Stability Graph (After Nickson, 1992, From Capes 2009)

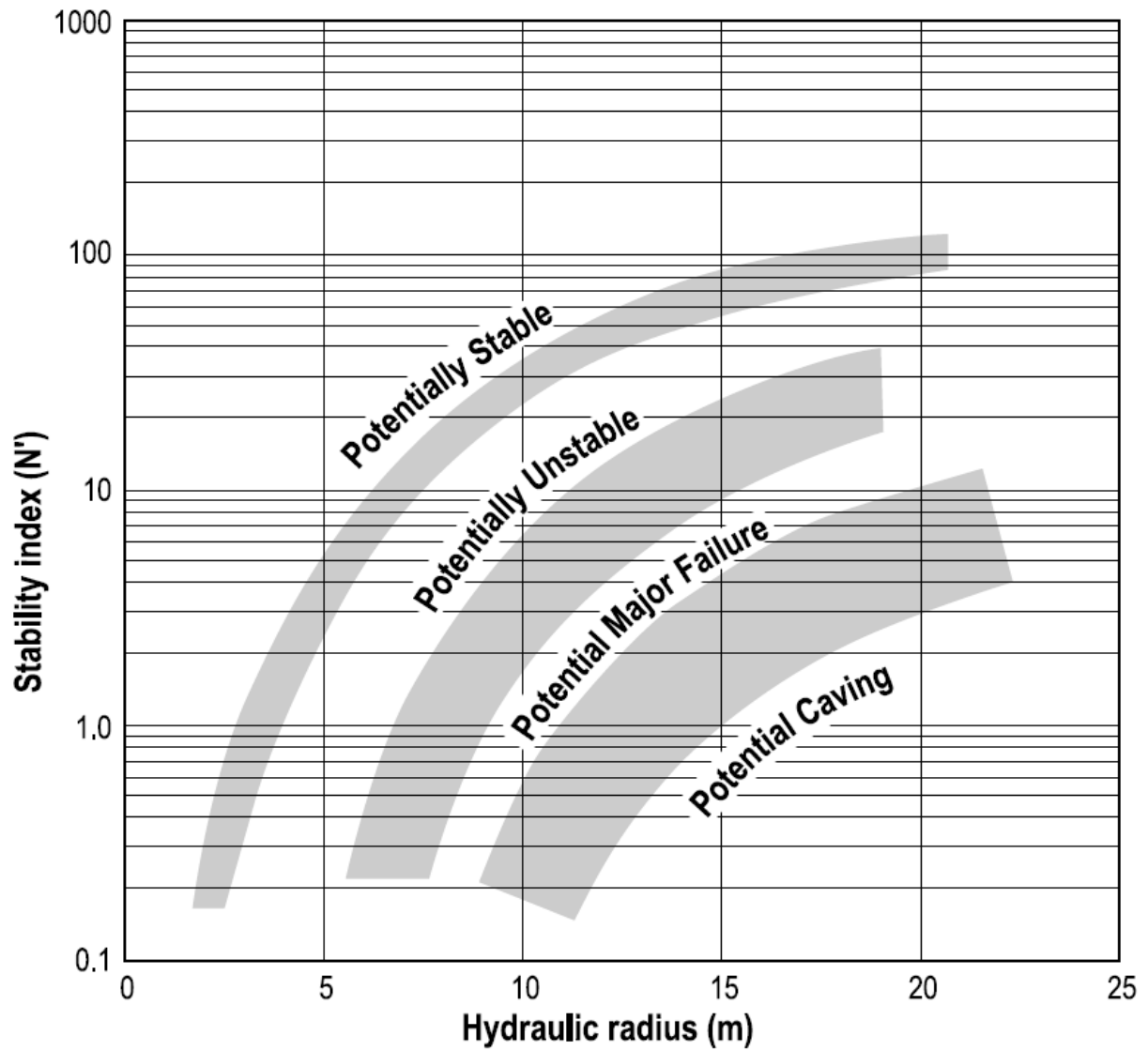


Figure 3.14. Modified Stability Graph (After Stewart and Forsyth, 1993, From Capes 2009)

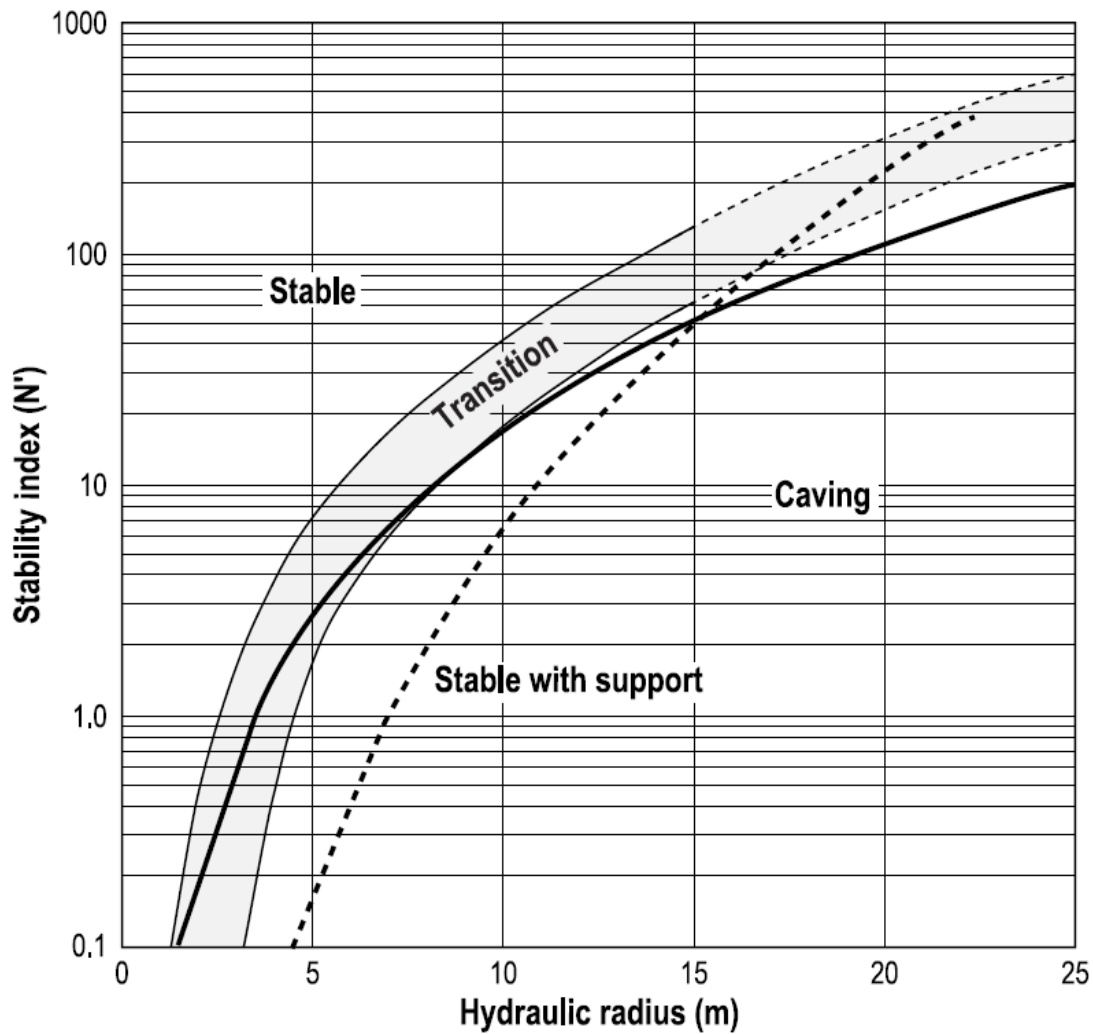


Figure 3.15. Modified Stability Graph (After Hadjigeorgiou et al. 1995, From Capes 2009)

The development of various graphs aimed to enhance accuracy in predicting hanging wall (HW) stability by considering influential factors specific to each database. However, these methods have limitations. Applicability outside the particular conditions and data used to develop the empirical system may lead to inaccuracies. Additionally, limited databases from weak rock masses may result in less reliable predictions. These methods may overlook factors in mining such as undercutting, stress, faulting, blasting effects, time-dependent behavior, and other considerations. Assessing irregular opening shape also creates challenges. Furthermore, the graphs

are usually developed for a specific rock type conditions and may not incorporate adjustments for rock contacts and structural complexities (Capes 2009).

3.3. Statistical analysis and numerical modeling

Suorineni et al. (2001) introduced the Bayesian likelihood method as a powerful tool for statistically interpreting the stability graph. They employed an extended database based on the Potvin-calibrated stability graph factors to illustrate the method's advantages. The Bayesian likelihood discrimination proved to be an optimal approach for statistically interpreting the stability graph due to its capability to reveal substantial overlap among the defined stability graph zones (stable, unstable, and caving). It also allowed for error rate estimation in the stability graph, delineated general transition boundaries between stable, unstable and caved stopes, estimated inherent predictive errors in stability graphs, evaluated the risk associated with using the stability graph for predictions, and introduced a multiple design curves stability graph based on the probability. The Bayesian likelihood discrimination's ability has been harnessed to provide deeper significance to the class boundaries in the stability graph and individual stope walls plotted within each class.

Numerical modeling stands as another widely adopted approach that has proven effective in addressing stope stability concerns. Henning and Mitri (2007), for instance, crafted a series of three-dimensional numerical models to explore the impacts of field stress, mining depth, stope configuration, and orientation on stope wall overbreak. Similarly, Purwanto et al. (2013) harnessed numerical modeling to establish the correlation between stope design and the stability of hanging wall. Hu and Cao (2009), by employing visual numerical simulation software, simulated and computed stress distribution and displacement variations within stopes during mining operations. They conducted an analysis of the stability of stope roofs and adjacent rock, as well as the alterations in sound emission associated with the mining process.

3.4. Machine learning methods

While classical stability assessment methods have been effective in many cases, the emergence of machine learning techniques has opened up new possibilities for enhancing their accuracy and efficiency. Machine learning algorithms can analyze large volumes of mining data, including shape of the opening, the properties of surrounding rock mass, underground conditions and historical stability records, to identify patterns and correlations that may not be easily discernible through traditional methods. By training machine learning models on a dataset of known stability outcomes, engineers can develop predictive models that can assess the stability of new stopes.

Several studies have explored the integration of machine learning models for the prediction of stope stability. Erdogan Erten et al. (2021) introduced a hybrid artificial neural network (ANN) approach optimized through grid search. This method was compared with conventional techniques including Naive Bayes (NB), Decision Tree (DT), k-Nearest Neighbors (kNN), Support Vector Machine (SVM), as well as the traditional stability graph method. The findings of this study revealed that the performance of the stability graph method falls short of the capabilities exhibited by machine learning algorithms. Notably, the ANN model with hyper-parameters tuned using the grid search technique showcased superior performance in terms of accuracy, precision, recall, f-measure, and g-mean compared to other machine learning algorithms. Saadaari et al. (2020) investigated the viability of employing Ensemble Learning methods to categorize and predict the stability condition of stope surfaces. They introduced and evaluated four techniques - Random Forest (RF), Gradient Boosting (GB), Bootstrap Aggregating Classifier (BAC), and Adaptive Boosting (AB) - using widely accepted and effective assessment metrics. Upon analyzing the performance outcomes, it was evident that among the four machine learning models, Gradient Boosting (GB) and Bootstrap Aggregating Classifier (BAC) demonstrated the highest efficacy in accurately classifying and predicting the stability state of stopes, encompassing categories of caved, stable, or unstable. In a comparative investigation conducted by Qi et al. (2018b), five distinct artificial intelligence strategies based on machine learning and metaheuristic algorithms were explored for their potential in predicting the stability of open stope hangingwalls (HW). The assessed algorithms encompassed logistic regression (LR), multilayer perceptron neural networks (MLPNN), decision tree (DT), gradient boosting machine (GBM), and support vector machine

(SVM). The optimization of hyper-parameters was facilitated using the Firefly algorithm (FA), which yielded successful results for this purpose. Across the testing phase, the most favorable performance was exhibited by the optimized GBM model, closely followed by the SVM model and the optimized LR model. The study highlighted the remarkable predictive capabilities of these three machine learning models in forecasting HW stability. Several researchers have directed their efforts toward the refinement and customization of specific machine learning models for assessing stope stability. Qi et al. (2018a), in their study, concentrated on optimizing the Random Forest model for enhanced efficiency, while Santos et al. (2020) shifted their attention toward the utilization of Artificial Neural Networks.

3.5 Summary

Numerous methods have been proposed for evaluating the stability of underground openings. Analytical design tools offer a range of options, provided that simplifications in geometry and rock mass properties, as well as assumptions about failure modes, are made. These tools enable the estimation of rock mass stability, offering valuable insights into the potential challenges and risks associated with underground excavations.

Continuing research in the field of stope stability in underground mining is crucial due to ongoing advancements in mining engineering aimed at maximizing exploitation and optimizing operations. The industry trend towards larger openings to increase extraction and profits may compromise stope stability. While this approach enhances efficiency and profitability, it also heightens the risk of caving. Exceeding maximum dimensions can lead to stope instability, increasing the likelihood of shifting towards unstable conditions or even catastrophic failure. Therefore, ongoing research is essential to develop strategies for maintaining stope stability while accommodating larger openings for mining operations.

3.6. References

- Barton N, Lien R, Lunde J (1974) Engineering classification of rock masses for the design of tunnel support. *Rock Mechanics* 6(4):189–236. <https://doi.org/10.1007/BF01239496>
- Beer G, Meck JL (1982) Design curves for roofs and hanging-walls in bedded rock based on “voussoir” beam and plate solutions. *Trans - Inst Min Metall, Sect A; (United Kingdom)* 91:Pages: A18-A22
- Brady BHG, Brown ET, Brown ET (2006) *Rock mechanics for underground mining*, 3. ed., repr. with corr. Springer, Dordrecht
- Capes GW (2009) Open stope hangingwall design based on general and detailed data collection in unfavourable hangingwall conditions, NR62618 Ph.D. The University of Saskatchewan (Canada)
- Clark LM (1998) Minimizing dilution in open stope mining with a focus on stope design and narrow vein longhole blasting. <https://doi.org/10.14288/1.0081111>
- Diederichs MS, Kaiser PK (1999) Tensile strength and abutment relaxation as failure control mechanisms in underground excavations. *International Journal of Rock Mechanics and Mining Sciences* 36(1):69–96. [https://doi.org/10.1016/S0148-9062\(98\)00179-X](https://doi.org/10.1016/S0148-9062(98)00179-X)
- Erdogan EG, Bozkurt KS, Yavuz M (2021) Grid Search Optimised Artificial Neural Network for Open Stope Stability Prediction. *International Journal of Mining, Reclamation and Environment* 35(8):600–617. <https://doi.org/10.1080/17480930.2021.1899404>
- Evans WH (1941) *The strength of underground strata.*, 50th edn. Trans. Instn. Min. Metall.
- Goodman RE (1989) *Introduction to rock mechanics*, 2nd ed. Wiley, New York
- Hackston A, Rutter E (2016) The Mohr–Coulomb criterion for intact rock strength and friction – a re-evaluation and consideration of failure under polyaxial stresses. *Solid Earth* 7(2):493–508. <https://doi.org/10.5194/se-7-493-2016>
- Hadjigeorgiou J, Leclair J, Potvin Y (1995) An update of the stability graph method for open stope design. *CIM Rock Mechanics and Strata Control session*, Halifax, Nova Scotia :14–18

- Henning JG, Mitri HS (2007) Numerical modelling of ore dilution in blasthole stoping. *International Journal of Rock Mechanics and Mining Sciences* 44(5):692–703. <https://doi.org/10.1016/j.ijrmms.2006.11.002>
- Hoek E, Brown ET (1980) *Underground Excavations in Rock* (1st ed.). CRC Press, London. <https://doi.org/10.1201/9781482288926>
- Hu H, Cao Y (2009) Numerical Simulation Modeling and Calculation Analysis on Stope Roof Stability under the Complex Geological Conditions in Deep Mining. In: 2009 International Conference on Engineering Computation. IEEE, Hong Kong, China, pp 175–177
- Hutchinson DJ, Diederichs MS (1996) *Cablebolting in Underground Mines*. Bi Tech Publishers, Richmond.
- Mathews KE, Hoek E, Wyllie DC, Stewart, S (1980) *Prediction Of Stable Excavation Spans for Mining at Depths below 1000 Metres in Hard Rock*. Ottawa, ON
- Nickson SD (1992) Cable support guidelines for underground hard rock mine operations. <https://doi.org/10.14288/1.0081080>
- Obert L, Duval W (1967) *Rock mechanics and the design of structure in rock*. John Wiley & Sons, Inc., New York
- Pakalnis RT (1986) Empirical stope design at the Ruttan Mine, Sherritt Gordon Mines Ltd. <https://doi.org/10.14288/1.0081095>
- Potvin Y (1988) Empirical open stope design in Canada. <https://doi.org/10.14288/1.0081130>
- Purwanto, Shimada H, Sasaoka T, Wattimena RK, Matsui K (2013) Influence of Stope Design on Stability of Hanging Wall Decline in Cibaliung Underground Gold Mine. *IJG* 04(10):1–8. <https://doi.org/10.4236/ijg.2013.410A001>
- Qi C, Fourie A, Du X, Tang X (2018a) Prediction of open stope hangingwall stability using random forests. *Nat Hazards* 92(2):1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>
- Qi C, Fourie A, Ma G, Tang X, Du X (2018b) Comparative Study of Hybrid Artificial Intelligence Approaches for Predicting Hangingwall Stability. *J Comput Civ Eng* 32(2):04017086. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000737](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000737)

Saadaari FS, Mireku-Gyimah D., Olaleye BM (2020) Development of a Stope Stability Prediction Model Using Ensemble Learning Techniques - A Case Study. *GM* 20(2):18–26. <https://doi.org/10.4314/gm.v20i2.3>

Santos AEM, Amaral TKM, Mendonça GA, Silva DDFSD (2020) Open stope stability assessment through artificial intelligence. *REM, Int Eng J* 73(3):395–401. <https://doi.org/10.1590/0370-44672020730012>

Sofianos AI (1996) Analysis and design of an underground hard rock voussoir beam roof. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* 33(2):153–166. [https://doi.org/10.1016/0148-9062\(95\)00052-6](https://doi.org/10.1016/0148-9062(95)00052-6)

Suorineni FT, Kaiser PK, Tannant DD (2001) Likelihood statistic for interpretation of the stability graph for open stope design. *International Journal of Rock Mechanics and Mining Sciences* 38(5):735–744. [https://doi.org/10.1016/S1365-1609\(01\)00033-8](https://doi.org/10.1016/S1365-1609(01)00033-8)

Tishkov M (2018) Evaluation of caving as a mining method for the Udachnaya underground diamond mine project. In: *Proceedings of the Fourth International Symposium on Block and Sublevel Caving*. Australian Centre for Geomechanics, Perth, pp 835–846

Villaescusa E (1996) Excavation design for bench stoping at Mount Isa. *Transactions of the Institute of Mining and Metallurgy, Section A Mining Industry*

CHAPTER 4: PREDICTING THE STABILITY OF OPEN STOPES USING MACHINE LEARNING

The chapter examines how Machine Learning (ML) models can be employed to predict open stope stability in underground mining. It starts by outlining challenges in traditional stability assessment methods. Highlighting the limitations of conventional approaches, it emphasizes the potential of ML algorithms in stope stability prediction. Machine Learning models such as Logistic Regression and Random Forest were employed to learn from a Potvin database and make accurate predictions. Analyzing the results, it discusses the implications for mining safety and productivity, suggesting ML integration into existing practices for improved outcomes in underground mining. This Chapter is based on paper Szmigiel A, Apel DB. Predicting the stability of open stopes using Machine Learning. J Sustain Min. 2022 Nov 22;21(3):241–8.

4.1 Introduction

The stability of underground excavations in open stope methods is one of the mining industry's major concerns. Three essential aspects need to be considered while designing open stopes. The first one concerns the properties of the rock mass and its mineral components that directly impact the behavior of the surrounding rocks. The second aspect is how stress fields are impacting the rock mass. Those stresses might develop zones of relaxation or increased compressive stresses in stope walls. The last aspect is the underground openings' geometry, size and orientation. Those three crucial features interact together and directly impact the complexity of underground stopes design (Potvin 1988).

The effectiveness of open stope mining methods depends on safety and high productivity. Usage of very large and non-entry excavations and mechanized mining equipment is necessary. However, the development of each stope is associated with large investment costs, which is the main reason for industries to reduce the number of stopes by increasing their dimensions. The significant difficulty facing that approach is that the consequences may be catastrophic when stopes are exceeded to their maximum dimensions. Another challenge for industries is that dilution in rock mass needs to be considered when designing open stopes. In addition, the dimensions need to be specifically adjusted to geotechnical conditions.

Original stability graphs developed by Matthews were based only on 50 history cases. That number was later extended to 176 cases by Potvin. In addition, Matthews's graph took three distinct separated by transition zones: stable, unstable and caved. Potvin (1988) modified that, and the number of zones was reduced to stable and caved, separated by transition.

The beginnings of the Artificial Intelligence and Machine Learning concepts date back to the mid-20th century. In 1943 the first mathematical model of a neural network was presented by Warren McCulloch and Walter Pitts, where the concepts of the neurophysiology of brain cells and calculus were combined (McCulloch and Pitts 1943). This research was a foundation that triggered the interest of scientists in further investigation. The first Artificial Intelligence model in the modern sense has its origin in the research presented by psychologist Frank Rosenblatt. He created a machine to recognize letters which became a prototype of an Artificial Neural Network known today (Rosenblatt 1957). Although, in the beginning, Machine Learning was used as a training

program for Artificial Intelligence, in the late 1970s, research focused on using knowledge-based and logical methods, which caused the separation of AI and ML (Carbonell et al. 1983). From that point, computer programs, and more precisely Machine Learning, started to be more present, expanded and applied in various tasks.

Using Machine Learning might be considered a new approach to determine the stability of open stopes. In previous research, ML models presented promising effectiveness in various research disciplines in mining engineering. For example, those algorithms were applied in mineral processing to predict the outcome values recovered from various beneficiation processes, such as flotation (Pu et al. 2020). In addition, the classification algorithms have been successfully applied in areas of mining engineering such as rockburst liability prediction (Pu et al. 2018) or an image recognition of coal (Pu et al. 2019).

Researchers have approached the problem of open stopes stability assessment with computing sciences methods. Most popular include numerical modelling, presented by Vallejos and Diaz (2020), which applies a new criterion for numerical modelling to evaluate a hangingwall overbreak.

Some of the Machine Learning and Artificial Intelligence models were employed to predict the stability of open stopes, such as Random Forest (Qi et al. 2018) and Artificial Neural Network (Santos et al. 2020). Both of those studies presented promising capabilities of those models. However, a smaller database was investigated (115 and 35 examples, respectively). Extending that databases could have a crucial impact on predicting the capabilities of those models.

The Potvin database was passed to two ML algorithms in this research – Logistic Regression and Random Forest. Both showed satisfying predicting capability, with an average accuracy of 0.68 for *LR* and 0.71 for *RF*. However, the latter performed better, especially with predicting unstable zone, which was the most challenging for both algorithms to predict because of similar values to other classes.

4.2. Open stope mining method

An open stope mining method extracts an enormous block of material using the drill or blast method. Then, tunnels are mined underground to access that orebody. After the material is removed by heavy machinery, the open void or stope is created. Later, it's usually backfilled, which allows for the extraction of adjacent deposits by opening new stopes. The walls of rock mass surrounding the stopes are called hanging walls (HW), and their properties vary depending on the geology and mining constraints. The appearance of a hanging wall causes the stope to be less stable (Capes 2009).

4.2.1. Matthew stability graph method review

The Mathews stability graph design method was established explicitly for deep underground mining excavations open stope surfaces. It was developed and presented for the first time in 1981.

This widely used method relates two calculated factors: shape factor (S) or hydraulic radius (HR) and stability number (N'). The primary principle theory behind the Mathews stability graph is that the dimensions of an excavation surface can be associated with the rock mass conditions and indicate either instability or stability of the opening (1981).

Stability Number N' is explicitly developed for designing span dimensions and support, and it yields the physical conditions of the stopes. To calculate Stability Number specific rating systems are being applied. N' is defined as follows:

$$N' = Q' \cdot A \cdot B \cdot C \quad (4.1)$$

Where:

Q' – Modified Q value

A – Rock stress factor

B – Joint orientation adjustment factor

C – Surface orientation factor

Q' value was first presented in 1974 by Barton et al. of the Norwegian Geotechnical Institute (NGI) to evaluate rock mass characteristics and ground condition (Barton et al. 1974). The purpose of calculating the Q' value was to determine the support required in mining excavations, tunnels and rock caverns. The formula is as follows:

$$Q' = \frac{RQD}{J_n} \cdot \frac{J_r}{J_a} \quad (4.2)$$

The first quotient ($\frac{RQD}{J_n}$) represents the structure of the rock mass, and the latter one ($\frac{J_r}{J_a}$) represents the roughness and frictional properties of the joint wall or filling material.

Rock Quality Designation (RQD) is a system developed by Deere (1963). It is widely used as a factor in classification systems and as a primary parameter for tunnel support selection. It quantifies the competence of a drill core, and it is defined as a ratio between the total lengths of entire pieces (larger than 10 cm) and the total length of a core. The values of J_n , J_r , and J_a are determined using NGI Classification Charts presented in Hoek and Brown (Hoek and Brown 1980).

Values of A , B and C factors are determined by graphs developed by Matthews and can be found in Matthews et al. (1981) and Potvin (1988). Factor A is a function represented by a ratio of the intact rock strength, determined by the Uniaxial Compressive Strength (USC) test, and the induced stress, maximum tangential stress acting parallel to the exposed surface at the boundary of a slope. Factor B accounts for the orientation of the geological structures (joint sets) concerning the investigated plane. It is determined by the angle of intersection between the exposed surface and the most predominant structure. Finally, factor C stands for surface inclination, assuming that slopes backs are naturally less stable than walls. The reason for that is the impact of gravity.

4.2.2 Database

Potvin's database was collected from 34 mines using the open-stoping method between 1986 and 1987. The provided data includes the characteristics of the rock mass and the physical and stress conditions. In some situations, circumstances did not allow to estimate all the necessary conditions confidently caused by lack of access to the site or missing background information. Consequently, the entire database was divided into the main database that contained accurate and complementary data that was less accurate. The main data consisted of 84 cases, and the remaining 92 cases were complementary data created using the same parameters and principles as the main data. Potvin's database consists of several parameters regarding each open stope investigated. These parameters include: block size factor (RQD/J_n), stress conditions, the difference in the dip between the designed stope surface and the critical joint, the relative difference in the strike, the indication of anisotropy of the rock mass, the shear strength of the critical joint (J_r/J_a) and effect of the gravity. These features were used to calculate the input parameters necessary for the analysis – stability number and hydraulic radius (shape factor) (Potvin 1988). The distribution of the data is shown in Figure 4.1. The green dots are examples of stable cases, the blue ones unstable, and the red ones caved. Table 4.1 shows the entire Potvin database.

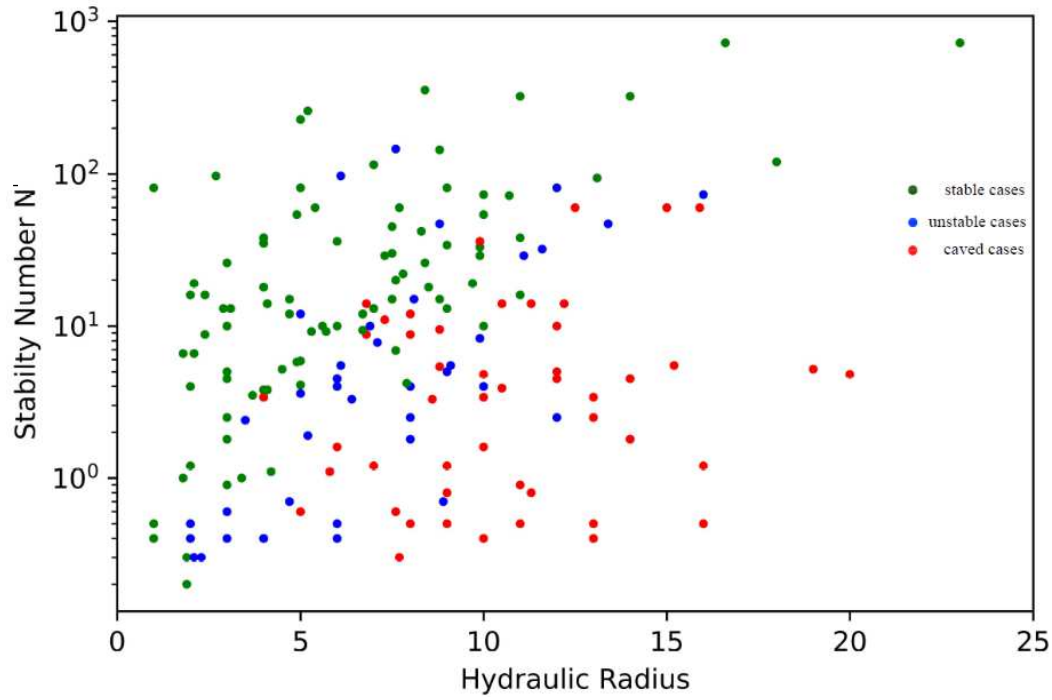


Figure 4.1. Distribution of the Potvin database.

Table 4.1. Complete Database from Potvin, 1988

	BLOCK SIZE	STRESS FACTOR	JOINT ORIENTATION FACTOR		EFFECT OF GRAVITY (C)		HYD. RADIUS	N	ASSES.
	RQD/J_n	(A)	CRITICAL JOINT (B)	J_r/J_a	SLIDING	FREEFALL/SLABBING			
1	18	1	0.64	3	6.5		5	228	STABLE
2	6	0.2	0.25	1	2.5		8.9	0.7	UNSTABLE
3	6	0.1	0.2	1	2.5		7.7	0.3	CAVE
4	7	1	0.2	1.5		3.7	7.1	7.8	UNSTABLE
5	40	1	1	1		8	14	320	STABLE
6	40	1	1	1		8	11	320	STABLE
7	40	1	1	1		6.5	5.2	260	STABLE
8	6	1	0.4	1.5	5		8.5	18	STABLE
10	4	0.3	0.2	0.8	3.5		4.7	0.7	UNSTABLE

12	7	1	0.2	0.6		6.5	9.1	5.5	UNSTABLE
13	15	1	0.2	2		7	8.3	42	STABLE
16	25	0.1	0.85	0.25		2	5.8	1.1	CAVE
17	25	0.1	0.85	0.25		2	4.2	1.1	STABLE
18	30	1	0.6	1		8	8.8	144	STABLE
19	30	0.1	0.4	1		2	3.5	2.4	UNSTABLE
20	11	1	0.2	1.5		2	1.8	6.6	STABLE
21	11	1	0.2	1.5		4.5	4.7	15	STABLE
22	11	1	0.2	1.5		4.5	8.8	15	STABLE
23	11	1	0.2	1.5		2	2.1	6.6	STABLE
24	17	1	0.2	2		2	10.5	14	CAVE
25	17	1	0.2	2		2	11.3	14	CAVE
26	17	1	0.2	2		2	12.2	14	CAVE
27	17	1	0.2	2		2	4.1	14	STABLE
28	8	1	0.3	1.5		2	7.6	6.9	STABLE
29	17	1	0.2	2	3		7.6	20	STABLE
30	17	1	0.2	2		5	9	34	STABLE
31	90	1	1	1		8	16.6	720	STABLE
32	90	0.1	1	1		2	4	18	STABLE
33	90	1	1	1		8	23	720	STABLE
34	90	0.4	1	1		2	10.7	72	STABLE
35	6	0.6	0.3	1.5		2.3	10.5	3.9	CAVE
36	6	0.9	0.3	1.5		5	9	13	STABLE
53	29	0.5	0.2	1.5		2	2.4	8.8	STABLE
54	29	0.5	0.2	1.5		2	6.8	8.8	CAVE
55	29	0.5	0.2	1.5		2	8	8.8	CAVE
56	4	1	0.3	0.5		8	19	5.2	CAVE
57	29	0.2	0.2	1.5		2	3.7	3.5	STABLE
58	29	1	1	1.5		8	8.4	352	STABLE
59	4	1	0.3	0.5		8	4.5	5.2	STABLE
61	17	1	0.3	1.5		6	7.5	45	STABLE
62	17	1	0.3	1.5		4	7.5	30	STABLE
132	6	1	0.2	1		8	5.6	10	STABLE

133	6	1	0.2	1		8	6.7	9.4	STABLE
134	5	0.1	0.2	1		2	1.9	0.2	STABLE
135	13	0.6	0.6	2		2	2.1	19	STABLE
136	13	0.5	0.6	2		2	2.4	16	STABLE
137	13	0.4	0.6	2		2	2.9	13	STABLE
138	13	0.4	0.6	2		2	3.1	13	STABLE
139	13	0.3	0.6	2		2	3	10	STABLE
140	8	1	0.3	1		6	7.5	15	STABLE
141	8	1	0.3	1		6	8.1	15	UNSTABLE
142	8	1	0.2	1		5.5	5.3	9.2	STABLE
143	8	1	0.2	1		5.5	5.7	9.2	STABLE
144	8	0.1	0.2	1		2	1.9	0.3	STABLE
145	8	0.3	0.2	1		2	1.8	1	STABLE
146	8	0.1	0.2	1		2	2.1	0.3	UNSTABLE
147	8	0.1	0.2	1		2	2.3	0.3	UNSTABLE
148	11	0.7	0.2	2		2	5	5.9	STABLE
149	5	1	0.2	0.1		6	9	0.8	CAVE
150	5	1	0.2	0.1		6	11.3	0.8	CAVE
151	15	0.4	0.2	2		2	10	4.8	CAVE
152	15	1	0.2	2		2	6.7	12	STABLE
153	15	1	0.5	2		8	18	120	STABLE
155	16	1	0.2	2	3		9.7	19	STABLE
156	16	0.1	1	2	3		5.6	10	STABLE
157	9	1	0.2	1.8		8	8.4	26	STABLE
158	10	0.1	0.2	2.5		2	3.4	1	STABLE
159	8	0.1	0.2	2		2	7.6	0.6	CAVE
161	3	1	0.2	1		8	20	4.8	CAVE
164	14	0.1	0.8	1.5		2	8.6	3.3	CAVE
165	14	1	0.2	1.5		8	9.9	33	STABLE
166	9	1	0.2	1.5	3		9.9	8.3	UNSTABLE
170	18	1	0.8	1.5		2.8	12.5	60	CAVE
171	18	1	0.8	1.5		2.8	15	60	CAVE
172	18	1	0.8	1.5		2.8	15.9	60	CAVE

173	18	1	0.8	1.5		2.8	7.7	60	STABLE
174	18	1	0.8	1.5		2.8	5.4	60	STABLE
175	18	0.5	0.3	1.5		8	11.6	32	UNSTABLE
176	18	0.5	0.85	1.5		2.5	7.3	29	STABLE
177	18	0.5	0.85	1.5		2.5	9.9	29	STABLE
178	18	0.5	0.85	1.5		2.5	11.1	29	UNSTABLE
180	6	1	0.3	1		5	6.9	10	UNSTABLE
183	16	0.1	0.3	1.5		8	4.9	5.8	STABLE
184	6	1	0.3	1		7	6.7	12	STABLE
64	4	1	0.3	1.5		5.5	6	10	STABLE
65	4	1	0.3	1.5		5.5	12	10	CAVE
66	3	1	0.3	0.8		7	3	5	STABLE
67	3	1	0.3	0.8		7	9	5	UNSTABLE
68	3	1	0.3	0.8		7	12	5	CAVE
69	18	1	0.3	3		4.5	16	73	UNSTABLE
70	6	1	0.3	0.8		8	5	12	UNSTABLE
71	6	1	0.3	0.8		8	8	12	CAVE
72	1	1	0.3	0.25		6.5	16	0.5	CAVE
73	16	1	0.3	3		8	7	115	STABLE
74	8	1	0.3	1.5		4.5	2	16	STABLE
75	8	1	0.3	1.5		4.5	11	16	STABLE
76	18	1	0.3	3		5	5	81	STABLE
77	3	1	0.3	0.25		8	14	1.8	CAVE
78	3	1	0.3	0.25		7	6	1.6	CAVE
79	3	1	0.3	0.25		7	10	1.6	CAVE
80	1	1	0.3	0.25		6.5	11	0.5	CAVE
81	18	1	0.3	3		5	9	81	STABLE
82	3	1	0.3	0.8		5.5	6	4	UNSTABLE
83	1.5	1	0.3	1.5		5	13	3.4	CAVE
84	18	1	0.3	3		4.5	10	73	STABLE
85	7	1	0.3	3		5.5	4	35	STABLE
86	20	1	0.3	3	4.5		1	81	STABLE
87	20	1	0.3	3	4.5		12	81	UNSTABLE

88	20	1	0.3	0.8		8	4	38	STABLE
89	20	1	0.3	0.8		8	11	38	STABLE
90	3	1	0.3	0.25		4	3	0.9	STABLE
91	3	1	0.3	0.25		4	11	0.9	CAVE
92	3	1	0.3	0.25		5.5	2	1.2	STABLE
93	3	1	0.3	0.25		5.5	7	1.2	CAVE
94	3	1	0.3	0.25		5.5	9	1.2	CAVE
95	3	1	0.3	0.25		5.5	16	1.2	CAVE
96	1	1	0.3	0.25		7	8	0.5	CAVE
97	1	1	0.3	0.25		8	3	0.6	UNSTABLE
98	1	1	0.3	0.25		8	5	0.6	CAVE
99	8	1	0.3	2		5.5	3	26	STABLE
100	3	1	0.3	1		5	3	4.5	STABLE
101	3	1	0.3	1		5	6	4.5	UNSTABLE
102	3	1	0.3	1		5	14	4.5	CAVE
103	6	1	0.3	0.25		5.5	3	2.5	STABLE
104	6	1	0.3	0.25		5.5	8	2.5	UNSTABLE
105	6	1	0.3	0.25		5.5	13	2.5	CAVE
106	15	1	0.3	2		6	10	54	STABLE
107	2	1	0.3	0.8		7	4	3.4	CAVE
108	2	1	0.3	0.8		7	10	3.4	CAVE
109	3	1	0.3	1		5	6	4.5	UNSTABLE
110	3	1	0.3	1		5	12	4.5	CAVE
111	2	1	0.3	0.5		6	3	1.8	STABLE
112	2	1	0.3	0.5		6	8	1.8	UNSTABLE
113	2	1	0.3	0.5		6	14	1.8	CAVE
114	3	1	0.3	0.8		5.5	2	4	STABLE
115	3	1	0.3	0.8		5.5	8	4	UNSTABLE
116	3	1	0.3	0.8		5.5	10	4	UNSTABLE
117	4	1	0.3	1.5		5.5	10	10	STABLE
118	1	1	0.3	0.25		6.5	6	0.5	UNSTABLE
119	1	1	0.3	0.25		6.5	9	0.5	CAVE
120	1	1	0.3	0.25		5	1	0.4	STABLE

121	1	1	0.3	0.25		5	2	0.4	UNSTABLE
122	1	1	0.3	0.25		5	13	0.4	CAVE
123	1	1	0.3	0.25		5.5	6	0.4	UNSTABLE
124	1	1	0.3	0.25		5.5	10	0.4	CAVE
125	1	1	0.3	0.25		6	1	0.5	STABLE
126	1	1	0.3	0.25		6	2	0.5	UNSTABLE
127	1	1	0.3	0.25		6	13	0.5	CAVE
128	12	1	0.3	0.25	4.5		7	13	STABLE
129	6	1	0.3	0.25		5.5	12	2.5	UNSTABLE
130	1	1	0.3	0.25		5	4	0.4	UNSTABLE
131	1	1	0.3	0.25		5.5	3	0.4	UNSTABLE
9	12	0.3	0.2	2		8	4.7	12	STABLE
11	5	1	0.2	0.6		7	7.9	4.2	STABLE
14	9	1	0.2	0.5		6	8.8	5.4	CAVE
15	9	1	0.3	0.5		7	8.8	9.5	CAVE
154	16	0.1	0.3	2		2	5.2	1.9	UNSTABLE
167	9	1	0.2	1.5		8	7.8	22	STABLE
168	15	1	0.2	1.5		8	6	36	STABLE
169	15	0.3	0.3	1.5		2	5	4.1	STABLE
179	15	0.1	0.85	1.5		2	4.1	3.8	STABLE
181	15	0.1	0.85	1.5		2	4	3.8	STABLE
182	15	1	0.3	1.5		8	4.9	54	STABLE
37	45	0.4	1	2.7		2	2.7	97	STABLE
38	45	0.4	1	2.7		2	6.1	97	UNSTABLE
39	45	0.6	1	2.7		2	7.6	146	UNSTABLE
40	30	0.6	1	1.3		2	8.8	47	UNSTABLE
41	15	0.6	1	2.6		2	13.4	47	UNSTABLE
42	14	0.5	0.3	1.3		2	6.1	5.5	UNSTABLE
43	14	0.5	0.3	1.3		2	15.2	5.5	CAVE
44	14	0.3	0.3	1.3		2	6.4	3.3	UNSTABLE
46	30	1	0.3	1.3		8	13.1	94	STABLE
47	9	0.3	1	2		2	7.3	11	CAVE
48	9	0.1	1	2		2	5	3.6	UNSTABLE

49	9	1	1	2		2	9.9	36	CAVE
50	9	0.4	1	2		2	6.8	14	CAVE

4.3. Model development, evaluation, and results

The two most popular Machine Learning algorithms were investigated in this research and applied to Potvin’s database. In order to achieve the most satisfying performance of the model, a few evaluation methods were applied to maximize efficiency and obtain the highest accuracy. These methods were also beneficial for eliminating the risk of overfitting, a common problem for ML algorithms applied for data with few training examples.

4.3.1. K-fold Cross-Validation

The k -fold Cross Validation is sometimes called also rotation estimation. It is a powerful tool to measure the success rate of our models used for classification (Marcot and Hanea 2021). The dataset is randomly split into a chosen number (k) of mutually exclusive sets (the folds) approximately equal in size. The main advantage of using CV is that each observation can be tested, which means that in every run, the testing set consists of different cases from the provided data set. In k -fold CV, we iterate over our data set k -times. In every round, our data is split into k sets, one part is being used for validation, and the remaining $k-1$ sets are merged into a training set (Kohavi 1995). Figure 4.2 presents the process of cross-validation where $k = 5$. That approach results in 5 different models fitted on partially overlapping training sets and tested on a non-overlapping validation fold. For every run, accuracy is established, and then the model performance is calculated as the arithmetic mean of all the accuracies.

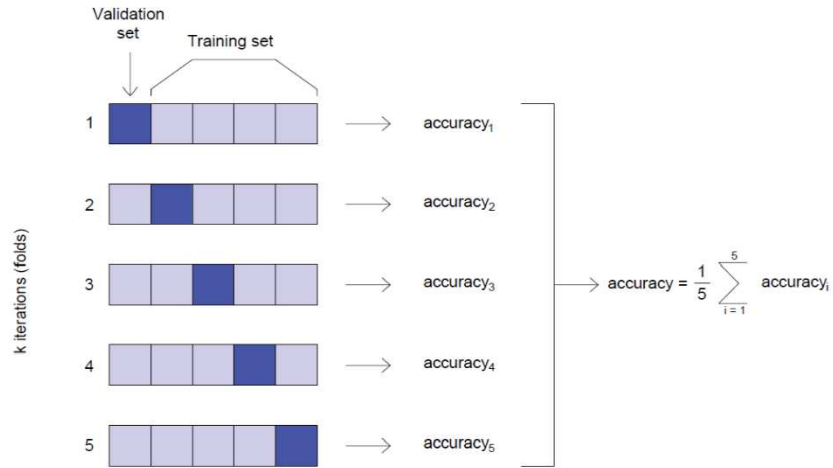


Figure 4.2. The process of k-fold Cross-Validation.

4.3.2. Confusion Matrix and ROC AUC score

A Confusion Matrix is one of the methods to evaluate the model's performance. It is a valuable tool to visualize how our model classifies the testing set. A Confusion Matrix is a size $n \times n$ where the predicted class and actual class are compared Visa et al. (2011). Table 4.2 shows a Confusion Matrix where $n = 2$ and entries have the meaning as follows:

TN – True Negative, the number of correct negative predictions

TP – True Positive, the number of correct positive predictions

FN – False Negative, the number of incorrect negative predictions

FP – False Positive, the number of incorrect positive predictions

Based on the confusion matrix, important metrics can be derived that are useful to evaluate the model's performance. The Recall (Sensitivity/True Positive Rate) shows the proportion of the

positive class that was correctly classified. False Negative Rate (*FNR*) is the proportion of the positive class that the model classified incorrectly. Specificity (True Negative Rate) indicated the proportion of correctly classified negative class and False Positive Rate (*FPR*), which shows the proportion of incorrectly classified negative class.

Table 4.2. The Confusion Matrix.

	Predicted Negative	Predicted Positive
Actual Negative	<i>TN</i>	<i>FP</i>
Actual Positive	<i>FN</i>	<i>TP</i>

$$Recall = \frac{TP}{TP+FN}, \quad (4.3)$$

$$FNR = \frac{FN}{TP+FN}, \quad (4.4)$$

$$Specificity = \frac{TN}{TN+FP}, \quad (4.5)$$

$$FPR = \frac{FP}{TN+FP}, \quad (4.6)$$

The ROC (Receiver Operator Characteristic) curve is a probability measurement tool that plots *TPR* vs *FPR*. AUC (Area Under the Curve) is a measure of the model's performance for classification problems, it shows the capability of the model to distinguish differences between classes, and it is defined as the summary of the ROC curve (Bradley 1997). The higher the AUC score, the better the performance. An excellent model should have AUC close to 1, indicating satisfying separability (Janitza et al. 2013). According to Hosmer et al. (2013), a model with an

outstanding performance would achieve an AUC score above 0.9. Values between 0.8–0.9 would be considered excellent, and 0.7–0.8 indicate acceptable classification.

4.3.3 Logistic Regression

Logistic Regression is one of the most popular machine learning algorithms used for classification problems. It is a predictive algorithm based on the concept of probability, which uses a cost function defined as the Sigmoid function.

4.3.3.1 Development of the model

As the input parameters to the logistic regression algorithm, the stability number (N') and shape factor (HR) were used. In this case, a multiclass classification problem was considered. The model predicted if the slope would be classified as Stable, Unstable, or Caved based on input parameters.

The data were randomly split into training and test sets, 80 and 20%, respectively. After that, a built-in sklearn library Standard Scaler was used to scale our data, so the distribution has a mean equal to 0 and a standard deviation equal to 1. The purpose of that is to standardize our features in cases when some of them have a larger magnitude and might dominate the estimation function, causing it to be unable to learn the features as correctly as we would expect (Pedregosa et al. 2011).

4.3.3.2 Results

The Logistic Regression algorithm from the sklearn library was fitted to our model. Figure 4.3 shows the decision boundary fitted to our data separating the unstable, stable and caved zone. It is noticeable that the model had the most problems with plotting decision boundaries for unstable cases. This is because these cases have similar values to either stable or caved ones and are hard to separate.

A confusion matrix was plotted to evaluate our model more carefully (Figure 4.4). Stable cases were mapped as 0, unstable as 1 and Cave as 2. We can see that Logistic Regression did a very

satisfying prediction performance for stable and caved cases. Unfortunately, it did not perform well for unstable zones, which supports the results obtained by plotting decision boundaries.

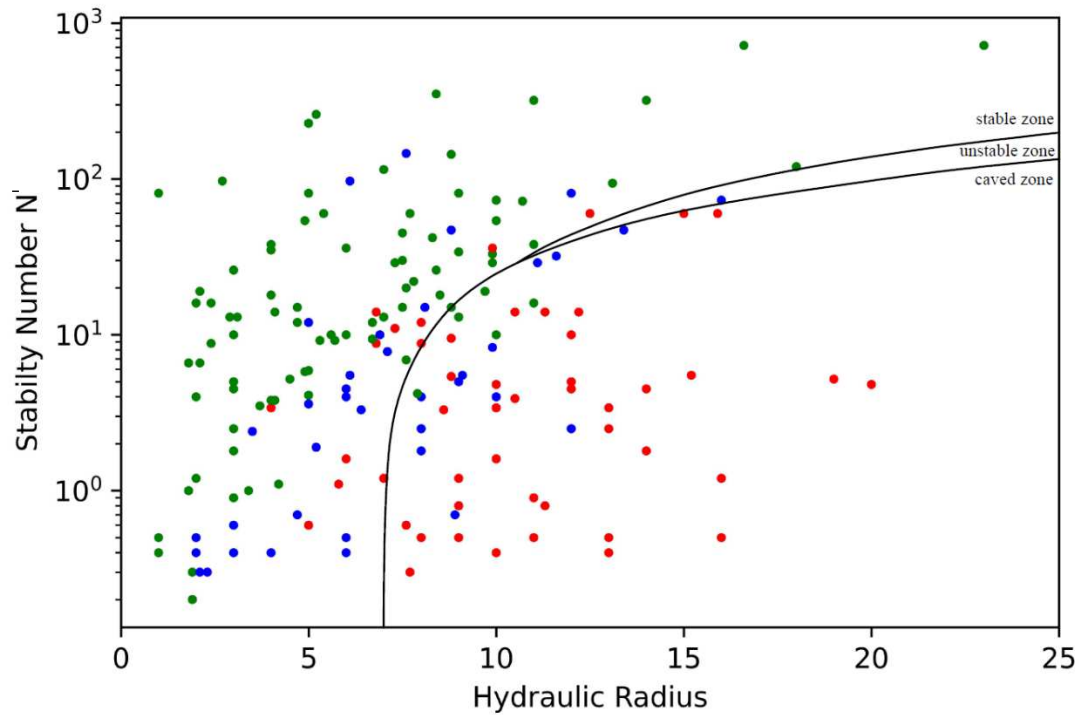


Figure 4.3. Decision boundaries for Logistic Regression.

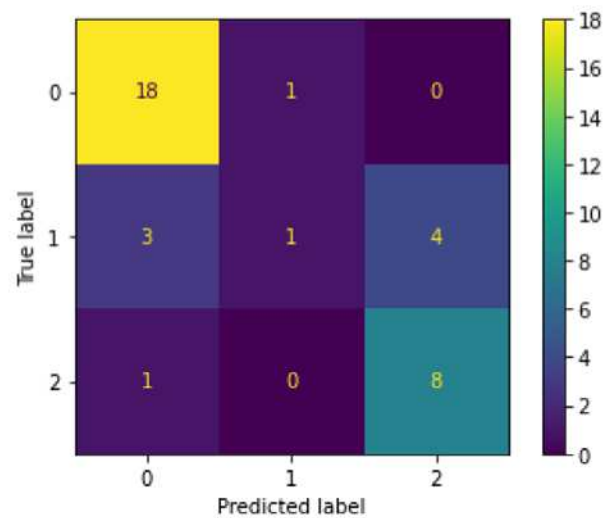


Figure 4.4. Confusion Matrix for Logistic Regression.

The Area under the ROC Curve was also calculated for the training and testing set. The obtained values were 0.81 and 0.78, respectively, which according to Hosmer et al. (2013), would be considered excellent for the training set and acceptable for the test set.

Cross-Validation was performed to determine the average accuracy of our model. The number of folds for our data was set to be equal to 5, which means five different accuracies were calculated for five different validation sets. The number of folds is determined by the size and characteristics of the datasets. It needed to be ensured that the training set and validation set are taken from the same distribution and that both sets include acceptable variation. For our dataset, the number of folds equal to 5 is sufficient, and it means that in every run the model was validated on 20% of the data (Marcot and Hanea 2021). The accuracies for our data were: 0.68, 0.64, 0.71, 0.71, 0.64. The average is 0.68, with a standard deviation of 0.03.

4.3.4. Random forest

A Random Forest is a powerful meta-estimator that can be applied to solve classification problems. The *RF* algorithm was presented by Breiman (2001) and has been successfully applied in many fields by researchers. The *RF* algorithm consists of a group of decision trees operating as a committee. The output is the value predicted by a larger number of decision trees. An example of Random Forest architecture is shown in Figure 4.5. The n represents the number of estimators (decision trees) that create the *RF*, and k_1, k_2, \dots, k_n are the results obtained by each decision tree (Qi et al. 2018).

The Random Forest algorithm allows us to regulate some of the parameters that directly impact the model's performance and are helping to control the overfitting or underfitting of the model. For example, we can change one of the parameters to control the algorithm's performance in several decision trees. The higher the number of decision trees, the better the chance for the model to properly learn the data. However, including too many estimators may slow down the process and increase the risk of overfitting.

Another parameter that helps to improve our model is the maximum depth of each tree in the forest. Developing deeper trees means that each tree would have more splits to better capture information about the data. However, if the decision trees are too deep for provided data, it might cause the same problems as too many estimators, slow processing time and overfitting of the data

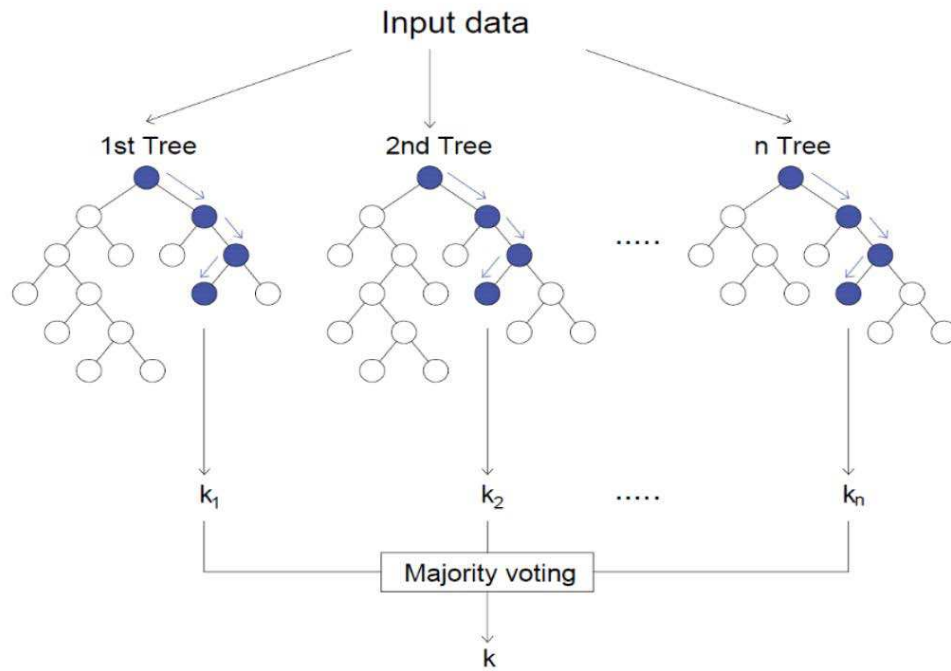


Figure 4.5. The architecture of Random Forest algorithm (Modified from Qi et al., 2018)

4.3.4.1 Development of the model

Potvin's database was passed to the Random Forest algorithm using the sklearn library in Python language. The input parameters were stability number (N') and shape factor – hydraulic radius (HR). The output was the range of stabilities labeled to each slope – either stable, unstable, or caved. The data was separated into training sets – used to train the data, and the test set – used to test the algorithm's performance. The size of the test set was equal to 36 examples, representing 20% of the whole data set.

4.3.4.2 Hyper-parameters tuning

In order to achieve the best performance of the *RF* model, the optimum number of estimators and the depth of each tree need to be established. These Hyper-parameters are necessary because the different values have special predictive performances. The AUC (Area Under the ROC Curve) score was used as the evaluation metric to find the optimum values. For the multiclass classification problem, the value of the AUC score was calculated using the One-vs-Rest scheme, and the macro average was reported.

AUC score was calculated for several estimators to find the most favorable number of decision trees for our data. Figure 4.6 shows the AUC score vs the number of estimators for the training and test sets. We can notice that the model achieves the best performance for the number of estimators, around 15. The performance decreases for more decision trees, and our model overfits the training data. The number of estimators was set to 15.

The same method was used to choose the optimum depth of each decision tree. The AUC score vs max depth was plotted for several values. In Figure 4.7, we can see that the model's performance decreases for depths higher than 4. We can also see that it overfits for large depth values. It predicts training data perfectly. However, it fails to generalize the findings for the test set. For our data, the tree depth was set to 4.

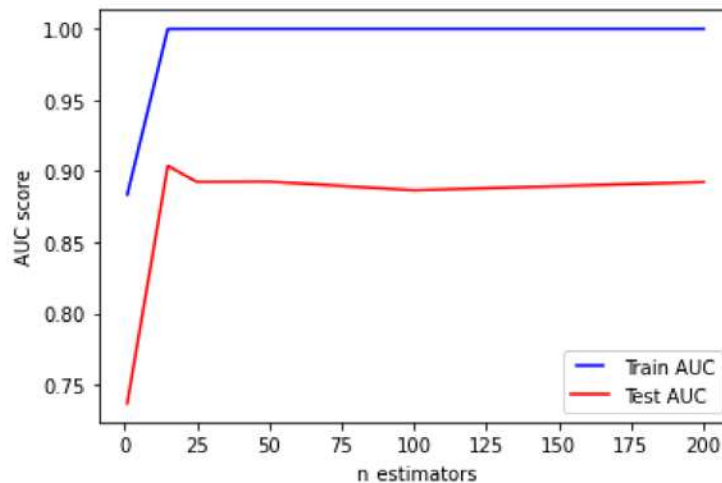


Figure 4.6. The AUC score for several numbers of decision trees

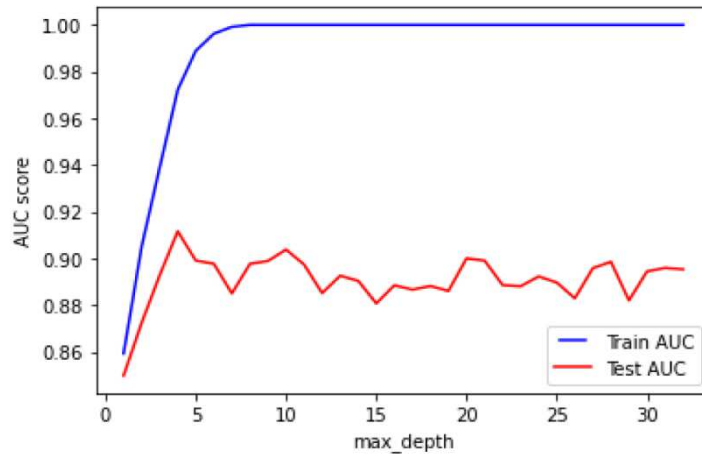


Figure 4.7. The AUC score for different values of decision trees depth.

4.3.4.3 Results

Random forest showed a favorable classification capability on Potvin's database. After choosing the number of decision trees and the depth of each tree, the model predicted the stability of stopes with satisfying accuracy results. At first, the data were randomly separated into training and testing sets. The accuracy for training data was 0.84, and for the test set, 0.75. The Confusion Matrix (Figure 4.8) was plotted to illustrate the model's performance. The label Stable was mapped as 0, Unstable as 1 and Cave as 2. It is noticeable that the model primarily has the most difficulties assigning the Unstable class. As shown in Figure 4.3, the plot of the distribution of our data, Stability Number vs Hydraulic Radius, the examples marked as Unstable are challenging to distinguish, and it's challenging for Machine Learning models to assign that class correctly similar values to either stable or Cave condition. The model shows the best performance for the Stable class; 17 out of 19 cases were predicted correctly.

In the next step, Cross-Validation was performed. The number of folds was chosen to be 5, and the obtained output was five accuracy values for five different training and validating sets. The values for each set were: 0.61, 0.75, 0.71, 0.78, 0.68. The average accuracy equals 0.71 with the standard deviation of ± 0.06 .

Then the Area Under the ROC Curve score for 15 decision trees with depth equal to 4 was calculated. For training data, the AUC score was 0.96 and 0.83 for test data. According to Hosmer et al. (2013), that score can be considered excellent for the test set and outstanding for the training set.

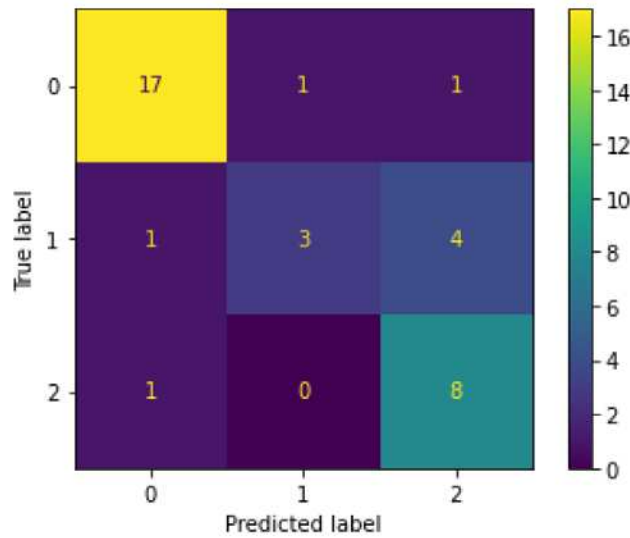


Figure 4.8. Confusion Matrix for Random Forest multiclass classification.

4.4 Summary and conclusions

The two most popular Machine Learning algorithms, Logistic Regression and Random Forest, were presented in this research to predict open stopes' stability. The total number of 176 history cases investigated was collected from Potvin (1988). Two variables from the data were selected, Stability Number (N') and the shape factor (HR), all of them with condition (label) assigned: stable, unstable or Cave.

Both models were evaluated using k -fold Cross-Validation, Confusion Matrix and ROC – AUC score to obtain the most satisfying results. Random Forest performed slightly better than Logistic Regression, mainly predicting the unstable class. The reason for that is that an unstable class has values of N and HR similar to other classes, and it is hard to separate them with a line.

In Radom Forest, hyper-parameter turning was performed to develop our model and achieve the best performance. First, the Area under the ROC curve was calculated for several estimators (decision trees) and different values of tree depth. Then the AUC vs number of estimators/tree depth were plotted to find the most optimum values for our model.

Confusion Matrix was plotted for both algorithms. It helped conclude that the model has the most difficulties predicting unstable classes. All three classes were investigated in the original Matthew stability graph method, but Potvin later reduced the number of classes to stable and Cave separated by a transition zone. That approach might be good for future investigation to increase the model's accuracy and achieve better performance since the unstable class is most difficult for ML models to predict.

In general, both algorithms showed satisfying capabilities and could be used in further investigation and have great potential in predicting the stabilities of open stopes. In future research, the data could be expanded to more historical cases to obtain even better results, as well as other ML or AI algorithms might be investigated.

4.5 References

- Barton N, Lien R, Lunde J (1974) Engineering classification of rock masses for the design of tunnel support. *Rock Mech Felsmech Mec Roches* 6(4):189–236. <https://doi.org/10.1007/BF01239496>
- Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit* 30(7):1145–1159. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2)
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
- Capes GW (2009) Open stope hangingwall design based on general and detailed data collection in unfavourable hangingwall conditions
- Carbonell JG, Michalski RS, Mitchell TM (1983) Machine Learning: A Historical and Methodological Analysis. *AI Mag* 4(3):69–69. <https://doi.org/10.1609/aimag.v4i3.406>
- Deere DU. (1963) Technical Description of Rock Cores for Engineering Purpose. *Rock Mech Eng Geol* 1:16–22
- Hoek E, Brown ET (1980) Underground excavations in rock, Rev. Institution of Mining and Metallurgy, London
- Hosmer DW, Lemeshow S, Sturdivant RX (2013) Applied logistic regression, Third edition. Wiley, Hoboken, New Jersey
- Janitza S, Strobl C, Boulesteix A-L (2013) An AUC-based permutation variable importance measure for random forests. *BMC Bioinformatics* 14(1):119. <https://doi.org/10.1186/1471-2105-14-119>
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp 1137–1143
- Marcot BG, Hanea AM (2021) What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? *Comput Stat* 36(3):2009–2031. <https://doi.org/10.1007/s00180-020-00999-9>

- Mathews, K. E. Hoek, D. C. Wyllie and S. B. V. Stewart (1981) Prediction Of Stable Excavation Spans for Mining at Depths below 1000 Metres in Hard Rock. Ottawa, ON
- McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5(4):115–133. <https://doi.org/10.1007/BF02478259>
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12(85):2825–2830
- Potvin Y (1988) Empirical open stope design in Canada. <https://doi.org/10.14288/1.0081130>
- Pu Y, Apel DB, Szmigiel A, Chen J (2019) Image Recognition of Coal and Coal Gangue Using a Convolutional Neural Network and Transfer Learning. *Energies* 12(9):1735. <https://doi.org/10.3390/en12091735>
- Pu Y, Apel DB, Wang C, Wilson B (2018) Evaluation of burst liability in kimberlite using support vector machine. *Acta Geophys* 66(5):973–982. <https://doi.org/10.1007/s11600-018-0178-2>
- Pu Y, Szmigiel A, Apel DB (2020) Purities prediction in a manufacturing froth flotation plant: the deep learning techniques. *Neural Comput Appl* 32(17):13639–13649. <https://doi.org/10.1007/s00521-020-04773-2>
- Qi C, Fourie A, Du X, Tang X (2018) Prediction of open stope hangingwall stability using random forests. *Nat Hazards* 92(2):1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>
- Rosenblatt F (1957) The Perceptron, a Perceiving and Recognizing Automaton Project Para. Cornell Aeronautical Laboratory
- Santos AEM, Amaral TKM, Mendonça GA, Silva D de FS da (2020) Open stope stability assessment through artificial intelligence. *REM - Int Eng J* 73(3):395–401. <https://doi.org/10.1590/0370-44672020730012>
- Visa S, Ramsay B, Ralescu A, Knaap E (2011). Confusion Matrix-based Feature Selection. Proceedings of The 22nd Midwest Artificial Intelligence and Cognitive Science Conference 2011, Cincinnati, Ohio, USA.

CHAPTER 5: ENHANCING UNDERGROUND EXCAVATIONS STABILITY PREDICTION IN MINING ENGINEERING: OPTIMAL CONFIGURATION OF AN ARTIFICIAL NEURAL NETWORK MODEL

This chapter focuses on advancing the accuracy of underground excavations stability predictions in mining engineering by optimizing an artificial neural network model. Analyzing Potvin's database, which consists of 175 historical cases, I explored the impact of different ANN model configurations; it was discovered that normalizing the data with Standard Scaler and implementing Swish as the activation function in all layers produced the most accurate predictions for this specific case. Furthermore, employing the SHAP (Shapley Additive exPlanations) tool allowed us to analyze the importance of the features and determine the factors with the highest influence. My findings reveal that the shape factor has the most significant impact on the stability of the underground openings, followed closely by the Q value. This research contributes to the optimization of predictive models for underground mining excavations stability and reveals the critical role played by specific parameters impacting the stability of open stopes.

5.1 Introduction

Mining engineering plays a crucial role in extracting valuable resources from the Earth's surface. Underground excavations, such as drifts and stopes, are vital components of mining operations. However, ensuring the stability of these excavations is essential to prevent accidents and maintain productivity and the well-being of workers. Traditionally, stability assessment in underground mining heavily relied on empirical approaches and engineering expertise. However, with the development of advanced computational techniques and the rise of machine learning, there is an opportunity to enhance the accuracy and efficiency of excavation stability predictions. This study explores the application of machine learning, specifically Artificial Neural Networks (ANNs), in predicting underground excavations stability in mining engineering.

Recent literature has shown a growing interest in employing machine learning and artificial intelligence for stope design and excavation stability assessment in underground mining. While previous studies such as Adoko et al. (2022), Santos et al. (2020) have focused on specific parameters, mostly stability number (N') and shape factor (HR), my research expands upon this by incorporating different factors contributing into stability assessment. By considering parameters like block size factor, shear strength of the critical joint, and others as separate model inputs, I aim to provide a more comprehensive analysis of their individual impact on stability predictions, marking an advancement in stope stability assessment.

The study focuses on Potvin's database consisting of 175 real-life cases, each with various features describing the characteristics of the surrounding rock mass, geological conditions, and the dimensions ratio of the opening. I selected the Potvin (1988) database for its capacity, thorough documentation, and completeness, making it a reliable resource for my study. With its collection of case histories, the Potvin database offers comprehensive insights into stope stability assessments, thereby providing a solid foundation for my research. Additionally, its open access availability ensures transparency and accessibility, aligning with my commitment to scientific accuracy and reproducibility. The dataset encompasses a range of stability outcomes: stable, unstable, and caved. By training an ANN model on this dataset, we can exploit the power of machine learning to discover complex relationships and patterns that may not be apparent through traditional methods. I will explore different model architectures, activation functions, and optimization algorithms to identify the optimal configuration that yields the highest accuracy and reliability. The primary

objective of this study is to investigate the impact of different factors on the stability of underground excavations. This research will contribute to the growing body of knowledge in applying machine learning techniques to mining engineering, offering insights into how those approaches can improve decision-making and safety in mining operations.

5.2 Engineering background and literature review

In underground mining operations, the stability of excavations is influenced by a variety of geological and mining factors, each contributing to the overall integrity and safety of the workspace. While open stope mining offers the advantage of increased productivity with reduced exposure to unsafe conditions, the potential for significant overbreak exists, influenced by the characteristics of the hangingwalls. Overbreak refers to the unintended displacement of unstable rock beyond the intended stope design, often caused by sloughing (Capes 2009). Such occurrences result in substantial operational expenses, production disruptions, and most significantly, create safety hazards. Therefore, it becomes crucial for mining operations to thoroughly evaluate the stability of stope walls to facilitate effective production planning and scheduling. Several important factors control and influence the overall stability in underground mining operations. This includes stress relaxation, which significantly influences mining excavation by affecting rock mass behavior (Diederichs and Kaiser 1999), mining technique and extraction rate, the mechanical and geological properties of surrounding rock and dimensions of the stopes (Stewart and Trueman 2004). Significant incidents in mining operations are faults, which refer to fractures or zones of discontinuity within the Earth's crust where there has been movement along the plane of the fracture. These movements can occur in various directions, including vertical, horizontal, or diagonal, and can result in displacement of the rock layers on either side of the fault. Faults are classified based on the direction of movement, the angle of the fault plane, and the type of stress responsible for the movement (Zhou et al. 2022). Faults are commonly present, especially in metalliferous mines, and have been recognized as the cause of instability for both mining and civil underground excavations (Suorineni 1998a). In open stope mining, critical issue is the method of liquidation of post-mining space, which involves managing the voids left underground after mining activities cease. This process may include backfilling voids with various materials or leaving them unfilled, with each approach presenting unique challenges and considerations. Additionally, the

impact of neighboring workings, such as adjacent stopes, must be carefully assessed to control risks related to stability and stress changes induced by nearby mining activities. Furthermore, managing mining edges—the boundaries of active excavation areas—is essential to prevent overbreak and maintain precise stope dimensions. Techniques such as controlled blasting and the installation of support systems play a crucial role in minimizing overbreak and ensuring worker safety.

5.2.1 Stability Parameters

The artificial neural network model was fitted to the Potvin database gathered between 1986 and 1987 from mining operation sites that were extracting material using the open stoping method. The parameters of the surrounding rock mass of each stope investigated were determined by the Mathews empirical method. These parameters included:

- Block size factor (RQD/J_n),
- The shear strength of the critical joint (J_r/J_a)
- Rock stress factor (factor A)
- Joint orientation adjustment factor (factor B)
- Effect of the gravity (factor C)
- Shape Factor (Hydraulic Radius) HR

The RQD (Rock Quality Designation) system was developed by Deere in 1964. It is widely used not only as a factor in classification systems but also as a basic parameter for tunnel support selection. It quantifies the competence of a drill core and is defined as a ratio between total lengths of an intact piece (larger than 10 cm) and the total length of a core (Deere 1963).

The J_n , J_r and J_a values represent the joint set number, joint roughness number and joint alteration number, respectively. Those values are determined using NGI (Norwegian Geotechnical Institute) Classification Chart, that can be found in Grimstad and Barton, (1980).

Factor A quantifies the effect of stresses acting on the open stopes exposed surface. This factor is a function represented by a ratio of the intact rock strength, determined by the Uniaxial Compressive Strength (UCS) test, and the induced stress, maximum tangential stress acting

parallel to the exposed surface at the boundary of a stope. Factor B accounts for the orientation of the geological structures (joint sets) concerning the investigated plane. It is determined by the angle of intersection between the exposed surface and the most predominant structure. The last factor, C, considers gravity's effects on stope stability, such as sliding, falling, and slabbing. The details about that classification procedure can be found in Potvin et al. (1988).

All these parameters are used to calculate stability number N, which is specifically developed for designing span dimensions and support. It yields the physical conditions of the stopes. N is calculated as follows:

$$N' = Q' \cdot A \cdot B \cdot C \quad (5.1)$$

Q value was first presented in 1974 by Barton et al. of the Norwegian Geotechnical Institute (NGI) to evaluate rock mass characteristics. The Q value was calculated to determine the support requirement in mining excavations, tunnels and rock caverns (Barton et al. 1974).

Originally, six parameters were used to calculate the Q value. However, the Mathew stability graph method utilized that number with four parameters. The formula is as follows:

$$Q' = \frac{RQD}{J_n} \cdot \frac{J_r}{J_a} \quad (5.2)$$

Another key element for effectively assessing the stability of underground mining openings is the shape factor, also known as the hydraulic radius (HR). The hydraulic radius, a crucial geometric parameter, defines the opening's shape and plays an important role in determining its stability. The term hydraulic radius is commonly defined as the ratio of the wall exposed area to its perimeter. In the case of inclined stopes, where the stope deviates from a vertical position, the most crucial factor for HR calculation is the hanging wall's exposure. The HR calculation considers the spans of the stopes along the dip (h) and along the strike (w), as illustrated in Figure 5.1 and equation 5.3.

$$\text{Hydraulic Radius (HR)} = \frac{W \cdot h}{2W + 2h} \quad (5.3)$$

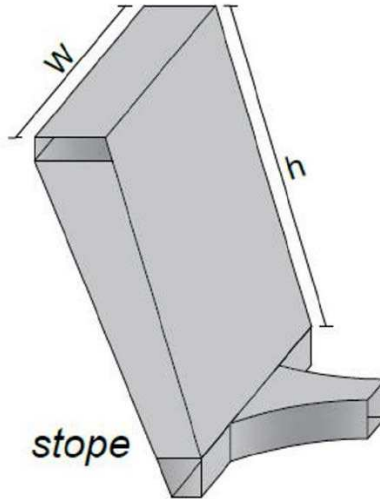


Figure 5.1 Open stope basic geometry.

The Matthew stability graph method was developed specifically for evaluating the stability of open stope surfaces in underground mining excavations. This popular and widely used method relates two factors, the stability number N' , and shape factor HR. The stability of the opening is then determined from a graph that plots stability number N vs. shape factor HR.

5.2.2 Literature review

The Matthews stability graph is an empirical method designed in 1980 to assess the stability of open stopes at depths exceeding 1000 m (Mathews et al., 1980). That research was followed by several authors who have collected new data, extended that method, and verified its effectiveness. Most of those adaptations and developments are focused on expanding the database of more cases and adjusting the position and total number of the stability zones employed. Empirical methods became widely used as a simple approach to generating guidelines for designing open stopes. The original Matthews database, consisted of 26 case studies from three mines, and 29 historical cases

from literature. Later the database was expanded by Potvin in 1988 to 175 cases, followed by Mawdesley et al. (2001) increasing the database to 485 cases.

Several studies have focused on adjusting and modifying the values of A, B, and C, resulting in different stability numbers of N' values and zones. Those modified factors were proposed by Diederichs and Kaiser (1999) and Mitri et al. (2011), Stewart and Trueman (2004) proposed an adjustment for the stress factor A. Some authors have considered the time effect in the excavations as a factor to calculate stability number N (Pakalnis 1986) or the impact of faults (Suorineni 1998b). However, there have been no general applications of those proposed factors.

Much research has focused on evaluating the influence of dimensions of the stope on its stability. Henning (2007) presented that increasing the height of a stope causes an increase in the overbreak of the stope walls. Hughes (2011) investigated the stope's dimensions and showed that larger strike lengths also increase the overbreak. The hydraulic radius, which defines the size and shape of the hanging wall, was also examined by Wang et al. (2007). That study showed a high correlation between the increasing value of HR and decreasing the stability of the stope.

The cavity monitoring system (CMS) allowed the creation of a three-dimensional analysis of the stope, which allowed the development of a tool to estimate the dilution of an opening. The concept of equivalent linear overbreak/slough (ELOS) presented by Clark (1998) was applied by several authors in the stability graph method (Papaioanou and Suorineni, 2016, Suorineni et al. 2015). Numerical modelling also significantly influenced the stope stability assessment presented by Idris et al. (2011). The study showed that the probabilistic approach might better capture the spatial variability of the surrounding rock mass properties than the classic empirical method. Numerical modelling was also applied by Heidarzadeh et al. (2019) to evaluate the effect of the geometrical parameters on the stope stability, and it showed that the hydraulic radius, stope span width and hangingwall dip have a significant influence on the stability of the stope. Henning and Mitri (2007,2008) presented a parametric numerical modeling study followed by a case study, to evaluate the influence of variety of factors on hangingwall ore dilution. Mitri et al. (2010) investigated the issue of unplanned ore dilution and its significant economic implications for mining operations. By analyzing surveyed stope profiles and utilizing numerical modeling, the study reveals that unplanned ore dilution is linked to stope design and construction, orebody geometry, and rock mass characteristics.

Some authors have approached the stope design with applications of machine learning and artificial intelligence techniques. The literature shows a variety of applications of ML and AI models that focus on designing open stopes and assessing the stability of underground excavations. In previous studies by Santos et al. (2020) and Adoko et al. (2022), the application of artificial neural network models to assess the stability of open stopes was investigated. Szmigiel and Apel (2022) compared random forest and logistic regression models. However, these studies were focusing only on stability number (N') and shape factor (HR) as input parameters. Erdogan Erten et al. (2021) considered a broader range of input parameters including stope surface, strike length, exposed height, ore width, HR, N' , and depth of excavation. A hybrid grid search-based ANN model was employed in this study, followed by comparison with k-nearest neighbour (kNN), naive bayes (NB), support vector machine (SVM) and Decision tree (DT) machine learning models. Wang et al. (2002) also evaluated the application of a neural network, where rock-mass rating (RMR) and spans of the openings were utilized as input parameters. Adoko et al. (2019) considered parameters such as hydraulic radius (HR), rock quality designation (RQD), joint set number (J_n), joint roughness number (J_r), joint alteration number (J_a), stress factor (A), joint orientation adjustment factor (B), gravity factor (C), design code, stress category, undercut area, ELOS (Equivalent Linear Overbreak\Slough) and maximum failure depth (FDmax). This study evaluated a mine stope performance in unfavorable rock mass conditions, by employing artificial neural network. However, the database for this investigation consisted only of 115 cases of hangingwall stability. In response to this, my study takes a step forward by considering all the parameters that combine into stability number (N') as distinct model inputs. These parameters are block size factor, the shear strength of the critical joint, rock stress factor (factor A), joint orientation adjustment factor (factor B), and effect of gravity (factor C), alongside shape factor (HR). This approach will allow us to examine the individual influence of each parameter on the predictive capabilities of the model and stability conditions.

5.3 Artificial Neural Network Model Overview

There is no one specific artificial neural network model that would be suitable for each problem and data. All published research papers and books come from various fields, such as engineering, psychology, social studies, mathematics, health sciences etc. Before approaching the problem we

would like to solve with ANN model, an evaluation of the dataset is necessary to determine the proper structure of the algorithm.

The database investigated in this research consists of 175 historical case studies of open stopes with information about the properties of the surrounding rock mass (factors Q' , A, B, C), the dimensions of the opening (HR) and the stability evaluation. The snapshot of the data set and its structure is presented in Table 5.1. The first 15 examples are shown, with factors Q' , A, B, C and HR determined, and the last column presents the assessment of each open stope: stable, unstable or caved. The number of stable cases in the dataset is equal to 88, the unstable cases are equal to 39, and the remaining 49 examples are classified as caved.

The pair plot (Figure 5.2) was used to investigate the data further to evaluate the correlation between all attributes in the data set. A pair plot (also called a scatterplot matrix) is a great tool that allows us to see the distribution of a single variable and the relationship between two variables. It is a popular method to establish trends for the analysis (Ahsan et al. 2021). The pairs plot function has built two basic figures, the scatter plots and histograms. The histograms on the diagonal show the distribution of one variable, and the scatter plots around allow us to see the relationship between the variables. It can be noticed in the data that only Q' value and HR are slightly correlated, whereas the rest of the features do not show any significant correlation with each other. The lack of linear correlation between the features indicates that an artificial neural network is the appropriate choice for the data since these models are particularly suited for handling non-linear relationships and complex patterns. Unlike traditional statistical methods, ANN models can capture complicated interactions and dependencies among input variables without requiring assumptions about the underlying data distribution. This flexibility enables ANN models to effectively learn from diverse and unstructured datasets, enhancing their performance in predictive tasks. Research by Hagan et al. (2014) and Bishop (1995) underscores the effectiveness of ANN models in handling non-linear relationships and noisy data, making them a preferred choice for analyzing not-correlated databases. ANN is also appropriate for this problem because those models have proved to be powerful in classification tasks.

Table 5.1 Potvin's data set investigated – first 15 examples (complete database can be found in Chapter 4)

CASE NUMBER	Q'	A	B	C	HR	STABILITY ASSESSMENT
1	54	1	0.64	6.5	5	
2	6	0.2	0.25	2.5	8.9	UNSTABLE
3	6	0.1	0.2	2.5	7.7	CAVE
4	10.5	1	0.2	3.7	7.1	UNSTABLE
5	40	1	1	8	14	STABLE
6	40	1	1	8	11	STABLE
7	40	1	1	6.5	5.2	STABLE
8	9	1	0.4	5	8.5	STABLE
9	3.2	0.3	0.2	3.5	4.7	UNSTABLE
10	4.2	1	0.2	6.5	9.1	UNSTABLE
11	30	1	0.2	7	8.3	STABLE
12	6.25	0.1	0.85	2	5.8	CAVE
13	6.25	0.1	0.85	2	4.2	STABLE
14	30	1	0.6	8	8.8	STABLE
15	30	0.1	0.4	2	3.5	UNSTABLE

**Figure 5.2** Pairs plot visualization of Potvin database.

5.3.1 Preprocessing the data

The accelerating development of data science has provided great opportunities for expanding and applying data-driven solutions for challenges in engineering. Data preprocessing is a significant foundation for proper and valid data analysis. It refers to a set of techniques for increasing the quality of the investigated raw data and has been generally acknowledged as an essential task and may account for even 80% of the total model development effort (Fan et al. 2021). In practice, it is almost always advantageous to employ preprocessing techniques on the input data before it is presented to a model (Bishop 1995, Nawi et al. 2013).

The first preprocessing technique that was applied to the data was feature scaling (standardization). There are several data scaling techniques. In the present study, two scaling methods were investigated to compare their influence on the model and to employ the most effective one. The artificial neural network is a model that learns the mapping from input variables to output variables. The input variables of the data have different ranges, which causes each variable to have a different scale, and those differences across the dataset may significantly increase the complications in proper model development. The pairs plot presented before can also help to visualize the necessity of applying standardization methods. It can be noticed that different features have different value ranges. Factors A and B have the same range from 0 to 1, and factor C is between 2 and 8. However, the Q' value is higher than 60 for most cases. Scale generally refers to the change in the range of the values and not the distribution shape. The two most popular approaches, MinMaxScaler and StandardScaler, were applied to the input data as scaling techniques.

MinMaxScaler is one of the most popular and widely recognized methods to standardize the dataset variables. For each element, the base estimation of that component is replaced with 0, the most extreme value (maximum value) is changed into 1, and then each other element is replaced with a decimal value from 0 to 1 (Patro and Sahu, 2015). MinMaxScaler preserves the original shape of the data distribution, and it doesn't modify the information embedded in the dataset. The formula for MinMaxScaler is as follows:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (5.4)$$

Where:

X_{max} – maximum value

X_{min} – minimum value

StandardScaler is another data standardization tool widely applied to several datasets. StandardScaler follows standard normal distribution; therefore, it expects the information to be ordinarily appropriated inside of each element and then scales those values to such an extent that the distribution revolves around 0 with a standard deviation of 1 (Raju et al. 2020). The values for StandardScaler are on a similar scale as MinMaxScaler, but the range of values is more significant. For each element, the mean and standard deviation are determined, and then the scaled feature is calculated as follows:

$$X_{scaled} = \frac{X - \mu}{\sigma} \quad (5.5)$$

Where:

μ - mean, σ – standard deviation

Artificial Neural Network is a powerful algorithm that requires input and output variables to be numbered, which means that categorical data should be encoded to numbers before training and fitting the model. The Potvin database is a multiclassification problem. The labels for each observation are the stability assessment: stable, unstable, or caved. To ensure the labels are properly adjusted for the model expectations, a common and determined method called one-hot encoding was applied, which is a common and determined method that transforms categorical values into vectors, requiring minimal processing. It is expressed as follows: let x be a categorical random variable with n distinct values x_1, x_2, \dots, x_n . The one-hot encoding of that value x_i is a vector where every element equals 0, except the i^{th} component, which is equal to 1 (Hancock and Khoshgoftaar 2020). In the Potvin database, a one-hot encoding technique was applied on stability categories: $S = \{\text{stable, unstable, caved}\}$, let $x_1 = \text{stable}$, $x_2 = \text{unstable}$ and $x_3 = \text{caved}$, a one-hot

encoding for the variable x is: (1, 0, 0), (0, 1, 0), and (0, 0, 1). Previous research has shown that one-hot encoding on categorical variables is sufficient to design an artificial neural network that outperforms other machine learning models (Duan 2019).

The next step to be implemented is splitting data into training, validation, and testing sets to evaluate and adjust the model to maximize the performance properly. Figure 5.3 shows the data splitting process, which allows a proper model adjustment. The size of each set should be adapted to the size of the database investigated and the model that will be employed.

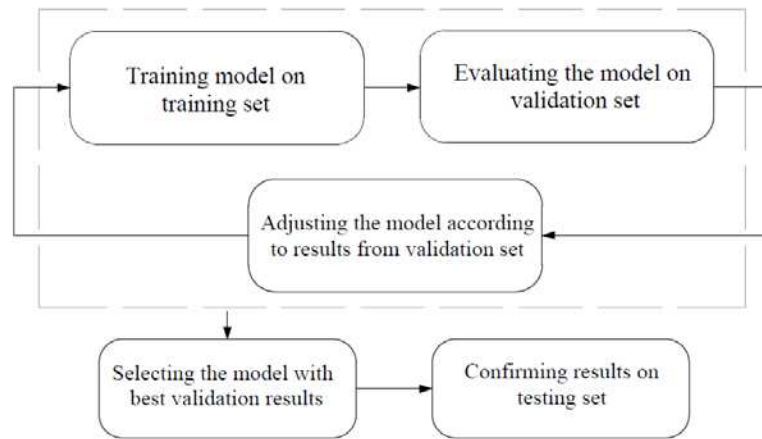


Figure 5.3 The steps of dividing the database into subsets

The decision to utilize the Potvin database consisting of 175 cases, is driven by considerations of computational efficiency, while ensuring access to a reliable and well-described dataset. Furthermore, the Potvin database, known for its completeness and availability, has not yet been explored using artificial neural network models. Previous research has demonstrated the effectiveness of employing relatively small datasets when evaluating stope stability (Qi et al. 2018; Adoko et al. 2019; Santos et al. 2020; Szmigiel and Apel 2022). This suggests that even with a limited number of cases, meaningful conclusions can still be drawn regarding the stability of underground mining excavations.

To ensure that my model has enough data to be properly trained on, I decided to split it into a 10% testing set and a 10% validation set, and the remaining 80% was used to train the ANN model. The first validation steps helped tune the model to increase the performance by choosing the best preprocessing technique and overall structure after the successful model adjustments were evaluated on an unseen testing set to confirm the results.

5.3.2 ANN model structure

Recent research interest in ANN has proved this is a powerful tool for capturing and classifying patterns in challenging data.

An Artificial Neural Network is a mathematical model that reflects human reasoning, can learn from previous experience and can cope with complex and non-linear characteristics of a certain problem. It is also specifically suited to reflect particular behavioural patterns where the relation between input and output features is difficult or impossible to define by other mathematical methods (Zhang et al. 1998). ANN was originally developed to resemble the human brain's structure— composed of interconnected neuron nodes. Each node has its own input, receiving information from another node or environment, and output, communicating and sending the information to another node/environment. Every node has a particular function that transforms the input into output. Figure 5.4 shows a diagram of a single processing element containing a neuron with weighted dendrites and axons. Dendrites are thin processes designed to receive information from another cell; axons are part of neurons that carry information away from the cell (Chavlis and Poirazi 2021).

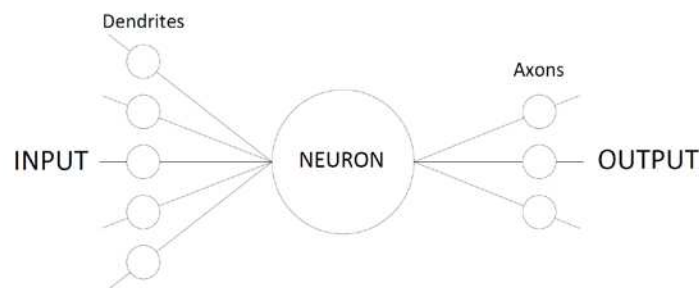


Figure 5.4 Diagram of one single processing element

The accuracy of an Artificial neural network model is defined by three fundamental elements: the architecture of the network (pattern and relationship between neurons), learning methods (determining weights) and activation functions (Arifin et al. 2019). ANN is built of three main layers: input, hidden and output. The input layer receives information from outside of the model. All neurons in the input layer are connected with the hidden layer, which consists of several neurons defined by the user. Hidden layers have an essential influence on the performance of the ANN model, and if the hidden layers were not implemented, the model would have limited capabilities, resulting in insufficient performance and poor predictions. Depending on the data structure and its complexity, the user defines the number of hidden layers and can vary from one to any number justified by a satisfying performance. The same rule applies to the number of neurons in each hidden layer. There is no definitive rule about what number of hidden layers or neurons in each layer should be applied. Every problem and every database are different and require careful investigation.

The shape of the investigated database directly determines the input layer, and it brings the data into the whole system for further processing by following layers. The input layer does not take any information from previous layers, making it the first initial step in the entire workflow for the model. The number of neurons in the input layer equals the number of features in the database. For my case, Potvin's database consists of five features, Q', A, B, C and HR, that are intended to assess the overall stability of each open stope, which means that my model would start with an input layer with five neurons.

One of the most significant and complex challenges in developing an ANN model is properly determining the number of hidden layers and neurons in each hidden layer. The literature and previous research activities don't clearly state one uniform and optimal solution for choosing the number of hidden layers and nodes. Traditional and one of the most popular methods is based on trial and error, which allows us to try different approaches to yield the most accurate results. The other methods that help to determine the number of layers and nodes include heuristic search, where we can gain knowledge from previous similar experiments where a close-to-optimal method might exist, and exhaustive search, which includes all possible topologies. Both of those methods include many alternatives, and it is very time-consuming to evaluate each of them. Those

techniques are commonly used as a starting point for consecutive searches by trial and error (Stathakis 2009).

In order to achieve the most satisfying results, the neural network requires an application of proper activation functions to perform various computations between hidden and output layers. The most common challenge for a Learning network is how the gradient flows within it, and some tend to be sharp in some directions and then slow or even equal to zero in others, creating an issue for an optimal selection. The gradient is the cause of the main problems of activation functions, such as exploding or vanishing gradients (Pascanu et al. 2012). Activation functions, referred to as transfer functions in some literature, are functions applied in neural networks to compute the weights of input variables and then decide whether the neuron can be activated. Activation functions can be linear or non-linear; the main purpose is to control the outputs of the network, and it has been proven that applying a proper function significantly improves the results of neural networks (Karlik and Olğaç 2011). In the classic linear model, the affine transformation usually gives the mapping performed from an input to the final prediction for each case. The formula for input vectors x transformation is given as:

$$f(x) = w^T + b \quad (5.6)$$

Where:

x = input

w = weights

b = biases

The neural network produces results from linear equation (5.7), necessitating activation functions to convert those linear outputs to non-linear and to learn patterns within the data. The following formula defines the outputs for those linear models:

$$y = (w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b) \quad (5.7)$$

These outputs from one layer are fed to the following layer until the final output is determined. However, since those outputs are linear by default, the non-linear activation functions should be implemented as transfer functions. The formula for non-linear output produced with an activation function is as follows:

$$y = \alpha(w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b) \quad (5.8)$$

Where α is the implemented activation function, the position of each AF in the model structure depends on its function; located after hidden layers, it transforms outputs from linear to non-linear, while in the last output layer, the activation function produces final predictions (Goodfellow et al. 2016). Deeper networks usually generate better performance of the models, although the deeper it is, the higher the risk of common issues such as vanishing or exploding gradients. With each multiplication, the values of the derivative terms that are less than one become smaller and tend to zero, causing the gradient to vanish. On the other hand, the values greater than one tend to increase with each multiplication, causing the gradient to explode. Hence, the activation functions maintain those gradients to specific and controlled limits (Elliott 1993).

5.3.3 Model evaluation

For every developed model, an evaluation of performance is a significant step to measure and determine the efficiency and accuracy of the proposed method. For this classification problem, I considered a few popular and widely applied methods to assess model predictions.

In the first place, I considered two evaluation metrics: accuracy and loss. Accuracy is a popular method for classification models, and it is a count of correct predictions usually expressed as a percentage. Accuracy is a significant tool for monitoring the model during training, allowing us to adjust some parameters to improve performance. On the other hand, loss is a cost function and a more challenging metric to interpret. Loss considers the probability and uncertainty of the prediction depending on how much the predicted value varies from the actual value. Loss summarizes the model's errors for each example in the training and validation sets data. Similar to

accuracy, it is mostly useful during model training to help adjust the algorithm's parameters. For accuracy, a higher score indicates better performance, whereas for loss, the goal is to minimize its value. It is common to observe a decrease in loss when accuracy increases, but this is not always true. Those metrics have different definitions and measure different characteristics; they may appear correlated, but there is no mathematical relationship between them. Both metrics give information about the model performance and help assess the risk of the most common issue in ANN models: overfitting and underfitting.

For multiclass classification problems, the confusion matrix proved to be one of the best and most popular methods to evaluate the performance and analyze the model's behavior by tuning some of the parameters (Szmigiel and Apel 2022). The confusion matrix compares predicted and actual classes and allows us to determine precision, recall and F1-score metrics. Precision indicates the proportion of examples classified as positive by the model and actual positives within the database. It measures how much we can trust the model and its positive predictions. Precision is defined as follows:

$$Precision = \frac{TP}{TP+FP} \quad (5.9)$$

Where TP is a true positive rate (correct positive predictions), and FP is false positive (examples incorrectly classified as positive).

Recall is a metric that measures the predictive accuracy for the positive class of the proposed model; in other words, it determines the ability of the algorithm to find all the positive features in the database. Recall is defined as:

$$Recall = \frac{TP}{TP+FN} \quad (5.10)$$

FN means a false negative rate, representing all the positive values incorrectly classified as negative.

Another important metric that needs to be established is the F1 score, which combines precision and recall under the concept of harmonic mean. The formula can be interpreted as a weighted average between recall and precision; and the best value for the F1 score is 1, and the worst is 0.

$$F1 - score = \frac{2}{precision^{-1} + recall^{-1}} = 2 \cdot \left(\frac{precision \cdot recall}{precision + recall} \right) \quad (5.11)$$

The F1- score could refer to both binary and multiclass classification problems and is significantly useful for finding the best trade-off between precision and recall (Grandini et al. 2020). Determining those values should help us better understand the model's performance and what improvements should be considered to increase prediction accuracy.

5.4 Model development and results

The results of artificial neural network predictions performed on data from underground mining excavations to assess the stability of openings are presented below. The Potvin database was passed to the model developed in Python, and the results were evaluated using several methods that have proved to be mostly effective. The database cases from real-life mines were collected to assess their stability, and the condition was indicated using the classic graph method developed by Matthews (1980). The data consisted of 88 cases classified as stable, 39 cases as unstable, and 49 as caved cases (Figure 5.5). Those stability assessments were passed to the algorithm as labels for the multiclass classification model. The data was preprocessed to maximize the performance and eliminate common issues such as unstable predictions, high bias or variance, low accuracy, and increased loss. The prepared database was then divided into separate sets: training, validation and testing, where the proportion of each set was 80%, 10% and 10%, respectively. The ANN model structure was then designed specifically for the provided database, considering the characteristics of the cases, the labels, the size, the complexity of the data, and the correlation between features. The model ought to learn, memorize and understand the data from 141 examples passed as a training set.

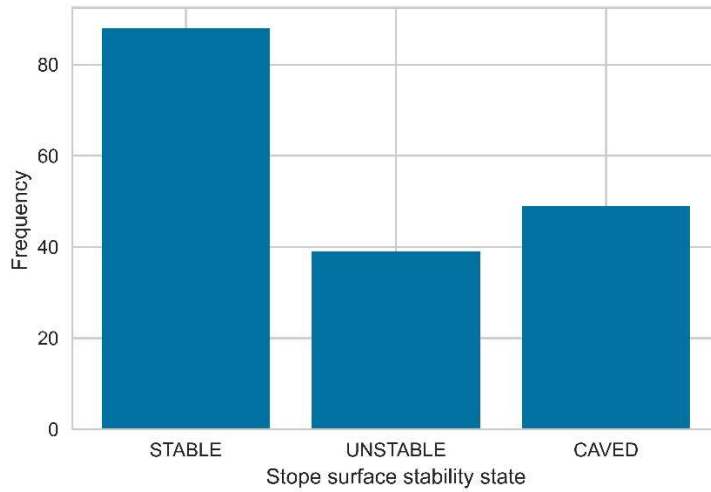


Figure 5.5 Bar plot of stability assessment count

The lack of significant correlation between the features generated the need for a more complex structure; the created feedforward ANN architecture is composed of four layers, with decreasing nodes in each hidden layer. The first input layer, constructed of 5 nodes, is directly established by the number of features defined in the data set. The second layer, and the first hidden layer within the model, comprises 20 nodes directly connected with an input layer and following the second hidden layer. Careful investigation of the prepared data, literature, trial and error runs, and scientific judgment allowed us to decide on the best activation functions for the problem. The activation function employed for the first hidden layer is Rectified Linear Units (ReLU). Introduced in 2011, ReLU was specifically designed to enhance the performance of deep neural networks, and it works by thresholding values at 0, defined as $f(x) = \max(0, x)$. In simpler terms, when $x < 0$, the model outputs 0, and when $x \geq 0$, the model outputs a linear function (Figure 5.6). ReLU became a well-known and widely applied activation function for classification problems, and it proved to be more computationally efficient than other standard functions. Because it does not saturate nor cause the vanishing gradient problem, it accelerates the convergence of gradient descent towards the global minimum (Agarap 2018). The simplicity and reliability that ReLU represents made it excessively popular and often favored by researchers and practitioners. However, ReLU does bear some issues because it is not zero-centred. It can suffer from a "dying ReLU" when all outputs for negative values are equal to zero. This causes some nodes to

completely disappear and not learn any patterns. Simultaneously, another problem is exploding activations, since there is no specific limit for ReLU it can reach infinity, which causes unstable nodes.

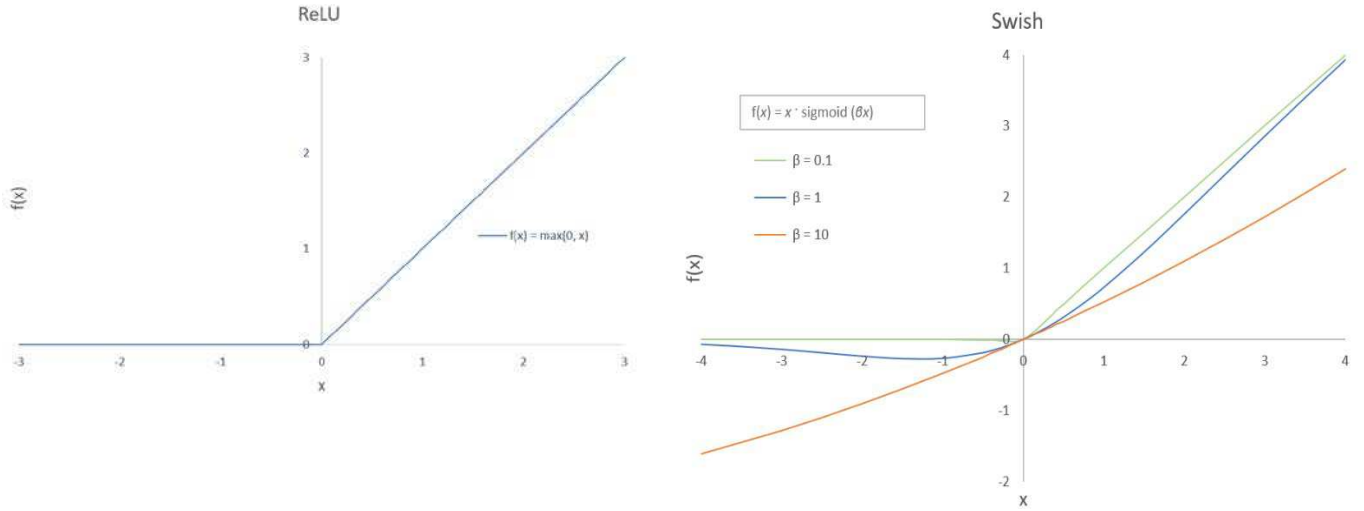


Figure 5.6 ReLU and Swish activation functions

Those issues prompted us to consider different activation functions for the training examples that were scaled using a standard scaler. Since the distribution of the standard scaler values revolves around 0 with a standard deviation of 1, causing some of the features to be less than 0, I have decided to apply swish as an activation function for that data. Presented for the first time by Ramachanran et al. (2017), the Swish activation function demonstrated better results on various deep learning models developed for challenging datasets. The formula for Swish is as follows:

$$f(x) = x \cdot \text{sigmoid}(\beta x) \quad (5.7)$$

Where $\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$ and β is a trainable parameter. As shown in Figure 6.6, Swish has some noticeable properties that can make it a better solution for certain data. Most

importantly, Swish allows a small number of negative values to be propagated to the next layer, unlike ReLU, where those values are directly changed to zero. Another advantage of the swish function is the trainable parameter β , which allows tuning the function to boost its information propagation and smooth the gradients. Softmax is the last activation function applied to the last layer for both data (scaled with StandardScaler and MinMaxScaler) in my model. Softmax activation function, also called multinomial logistic regression, takes in a vector of raw outputs from a neural network and transforms them into decimal probabilities where the values are between 0 and 1. It is a generalization of the sigmoid function that can be applied to multiclass classification problems when classes are mutually exclusive (Ren et al. 2017). Since softmax converts and scales real output values into a normalized probability distribution, it is mostly implemented in the neural network in the last layer just before the output layer and must consist of the same number of nodes as the number of classes/output layer. Softmax is defined as follows:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (5.8)$$

Where all the z_i values are input vector elements that take any real value, the bottom of the formula is the normalization term that takes those real values and changes them into decimal probability distribution values that sum to 1, and K is the number of classes in a multiclass classifier.

In the next step, the proper optimizer needed to be implemented to adjust each epoch's weight and reduce the loss of the algorithm. Choosing the proper weights for the model is a challenging task, hence the necessity of choosing a proper optimizer, which is a function that adjusts the attributes of neural networks like weights and learning rates to improve the overall accuracy and minimize loss. The Adam algorithm was applied, followed by cross-entropy as a loss function. Adam was presented for the first time by Diederik P. Kingma and Jimmy Lei Ba (2014). It is an efficient stochastic gradient-based optimization algorithm that is widely recognizable and applied. It calculates the individual learning rates for every parameter from estimates of the first and second moments of the gradient. Cross-entropy loss is a common and

widely used loss function in machine learning, especially is classification tasks; it measures the difference between the predicted and actual probability distribution of the classes. For multiclass classification problems, the cross-entropy loss is defined as:

$$L(\hat{y}, y) = - \sum_k^K y^{(k)} \log \hat{y}^{(k)} \quad (5.9)$$

Where y is a one-hot encoded vector of the true label and \hat{y} is a vector of predicted probabilities for all the classes (Bishop 2006; Goodfellow et al. 2016).

The final model structure was designed based on the size of the database and the number of features and classes in this classification problem. The specific number of nodes in hidden layers and the number of epochs were determined through experimentation, literature review and validation on a separate validation set. Figure 5.7 presents the model that generated the best results on the investigated database, where three hidden layers were developed, with 20, 10 and 5 nodes accordingly.

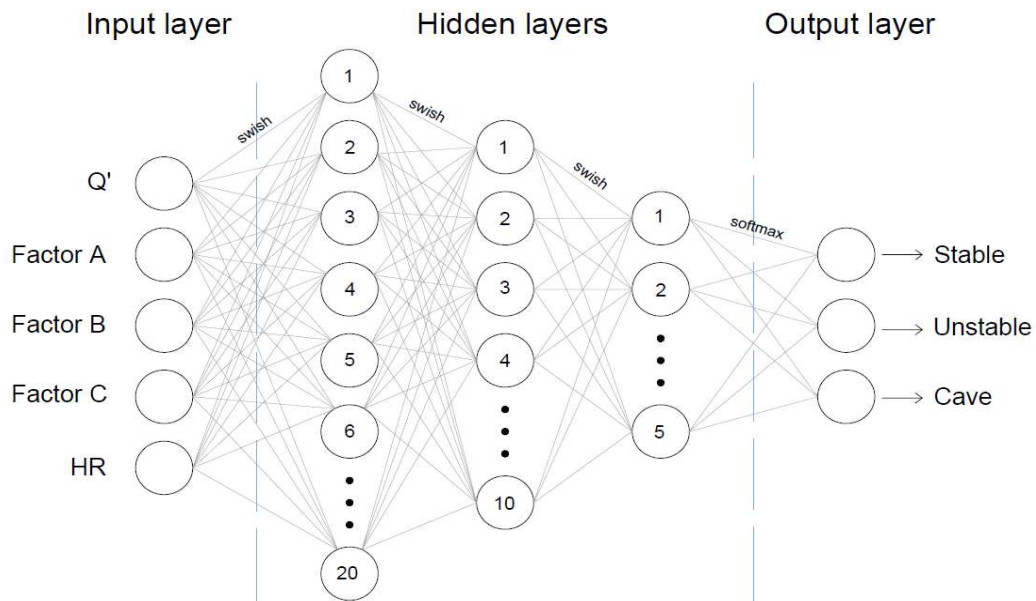
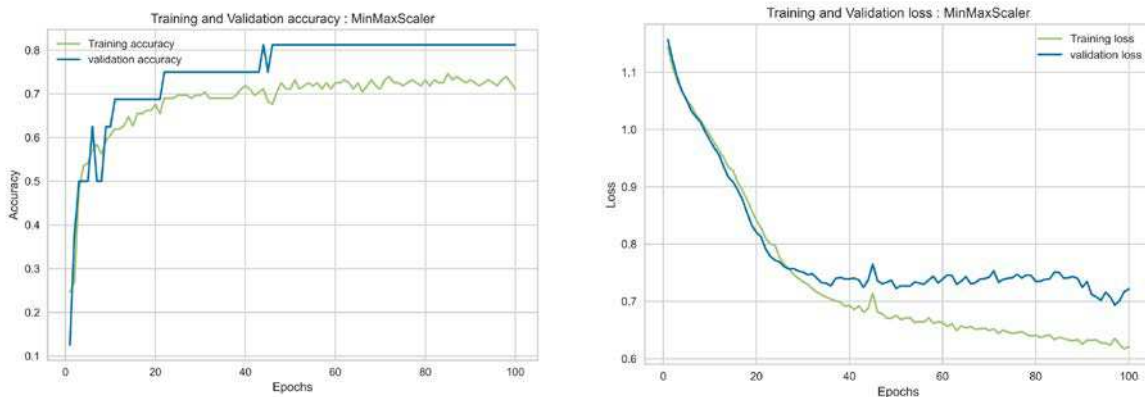


Figure 5.7 Final ANN model structure

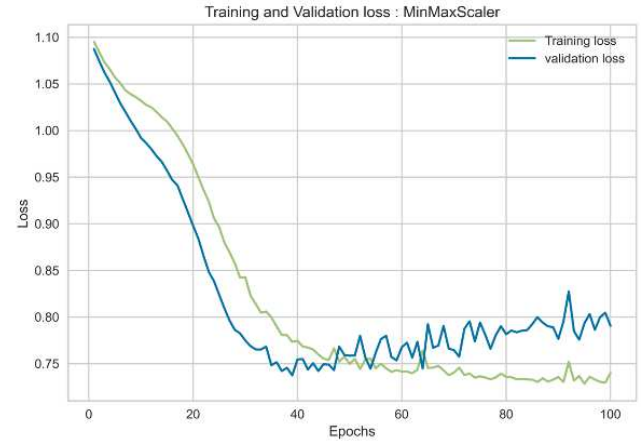
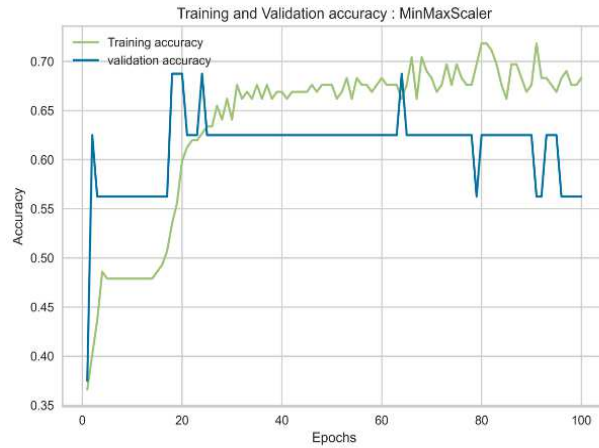
The number of epochs was set to 100, which refers to the number of times a model was trained on the training set. Each epoch consists of one forward and one backward pass of all training examples through the model. The model makes predictions on the entire training set, calculates the loss between the predicted output and the real value, and then updates the model's parameters through backpropagation. Increasing the number of epochs can result in better model performance since the model has more opportunities to learn from the training data. However, increasing epochs beyond a certain point can lead to overfitting, where the model becomes too specialized to the training data and performs poorly on unseen sets. Therefore, determining the optimal number of epochs is a crucial hyperparameter tuning step in machine learning.

5.4.1 Model results

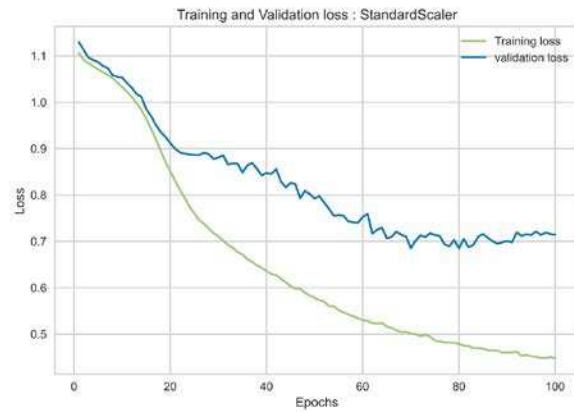
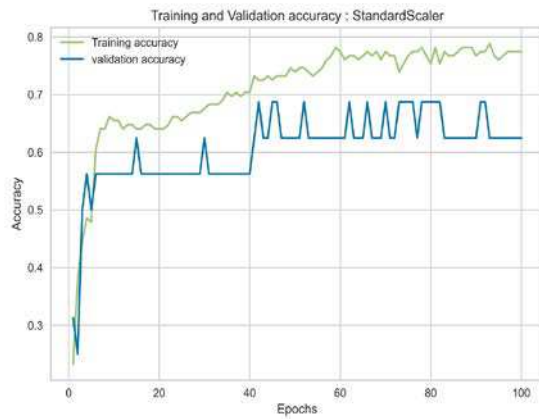
The model was trained on 80% of the database, equal to 141 data points. 10% of the data was then used for validation and tuning the model. The model was tested on an unseen set of 18 examples in the final step. In the first run, two scaling methods were compared: Standard Scaler and Min max scaler and two different activation functions: ReLu and Swish. The accuracy vs epochs and loss vs epochs for all the configurations that have generated the best performances are shown in Figures 5.8 – 5.15. These graphs were then used to establish the configuration that achieved the most satisfying results.



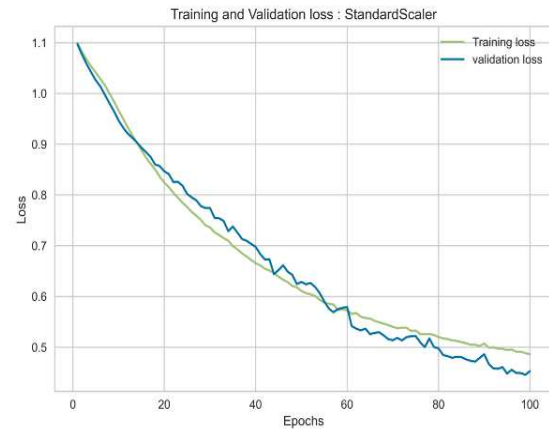
Figures 5.8 and 5.9 Accuracy and loss score vs epochs, for a data scaled with MinMaxScaler ReLu activation function



Figures 5.10 and 5.11 Accuracy and loss score vs epochs, for a data scaled with MinMaxScaler
Swish activation function



Figures 5.12 and 5.13 Accuracy and loss score vs epochs, for a data scaled with StandardScaler
Swish and ReLu activation functions



Figures 5.14 and 5.15 Accuracy and loss score vs epochs, for a data scaled with StandardScaler Swish activation function

The trade-off between accuracy and loss is important in machine learning models. Accuracy measures the overall correctness of predictions, while loss quantifies the error between predicted and actual values. While achieving high accuracy is desirable, solely focusing on accuracy may not always be optimal. In certain scenarios, a model with high accuracy may still exhibit noticeable loss, indicating poor prediction quality, hence the significance of balancing these two metrics. For instance, by increasing model complexity or allowing for more flexible decision boundaries, the model may achieve lower loss but become more prone to overfitting the training data, thus reducing accuracy on unseen cases. Taking this trade-off into account helps us decide on appropriate regularization techniques, model architecture, and hyperparameter tuning strategies, which allows us to reach the most appropriate balance between capturing complex patterns and preventing overfitting (Brownlee, 2016; Raschka and Mirjalili, 2019).

Plotting accuracy and loss versus the number of epochs has provided valuable insights into the performance of different configurations in my study. I could conclude the model's optimal preprocessing techniques and activation functions by examining the plots. Specifically, I found that scaling the data with StandardScaler and using Swish as the activation function for all hidden layers yielded the best performance. Although a configuration using MinMaxScaler with ReLU activation achieved high accuracy for the validation set (as shown in Figure 5.8), it is important to consider the corresponding loss (Figure 5.9). The high loss indicates that the model might overfit

the training data, leading to poor generalization and potentially inaccurate predictions of unseen data. I observed a drop in accuracy upon replacing ReLU with the Swish activation function (as depicted in Figures 5.10 and 5.11). Therefore, I concluded that the MinMaxScaler preprocessing option is not optimal for Potvin's stope stability database. When testing the configuration of StandardScaler and using the Swish activation function in the first hidden layer, followed by ReLU in the second and third hidden layers, I observed significant overfitting of the model. The accuracy of the training set was noticeably higher than that of the validation set (Figure 5.12).

Additionally, while the loss decreased for the training set, it remained high for the validation set (Figure 5.12). This discrepancy between the performance on the training and validation sets led us to conclude that the combination of Swish and ReLU activations for this database causes overfitting. As can be noticed in Figures 5.14 and 5.15, the combination of StandardScaler for data scaling and the Swish activation function for all the hidden layers demonstrated superior performance, reaching a better balance between accuracy and loss.

After tuning the hyperparameters, evaluating the model on the validation set, and deciding on the most satisfying configuration, testing the final model on an independent testing set is crucial. The testing set is an unbiased measure of the model's performance and provides a reliable estimate of its ability to generalize to unseen data. The model was fitted to an unseen testing set that consisted of 18 examples that were not used for training or validation. The performance parameters, such as the classification report and confusion matrix, were established.

The classification report is an important evaluation tool in machine learning, particularly for classification tasks. It provides a comprehensive summary of evaluation metrics for each class, offering valuable insights into the model's performance. Interpreting the classification report allowed us to assess the model's precision, recall, and F1 score for individual classes and overall accuracy. The classification report is presented in Table 5.2, and the confusion matrix in Figure 5.16. The model's overall accuracy is equal to 0.83, and an accuracy of this magnitude indicates that the model correctly predicts the proper class for approximately 83% of the instances in the testing set. This level of accuracy suggests that the model has learned meaningful patterns and relationships within the data, allowing it to make reasonably accurate predictions. While an accuracy of 100% is often desired, achieving such perfection is not always realistic or feasible, especially in complex real-world problems such as mining engineering. Therefore, an accuracy of

83% signifies a substantial level of success in capturing and understanding the underlying patterns in the stope stability database. High precision values for stable and cave classes indicate fewer false positives, while high recall indicates fewer false negatives predictions. The F1 score is a metric that combines precision and recall into a single value, providing a comprehensive evaluation of a model's performance. A higher F1 score is generally desirable because it indicates a better trade-off between these two metrics, it considers both false positives and false negatives, making it suitable for situations where misclassification of either positive or negative instances is equally important. The F1 score achieved by my model suggests that it can classify testing examples correctly while minimizing misclassifications, especially for Stable and Unstable classes.

Table 5.2 Classification report for testing set

	Precision	Recall	F1 - score
Stable	1.0	0.78	0.88
Unstable	0.5	1.0	0.67
Cave	0.86	0.86	0.86
Accuracy	0.83		

The confusion matrix presented in Figure 5.16 provides a visual representation of the model's predictions and facilitates a deeper understanding of its strengths and weaknesses, and it presents a summary of the model's performance across different classes.

By examining the graphical representation of the confusion matrix, we can see that my model showed promising performance in classifying test data. Seven of the nine examples of the stable class were correctly classified, demonstrating the model's ability to identify stable instances accurately. However, one stable example was misclassified as unstable and another as caved, suggesting room for improvement. Both unstable examples were correctly classified, and six were correctly classified among the seven examples representing the caved class. However, one caved example was incorrectly classified as unstable. Overall, the model's performance in classifying examples into stable, unstable, and caved classes shows promise, with scope for fine-tuning to enhance accuracy and minimize misclassifications.



Figure 5.16 Confusion Matrix of the prediction performance on the testing set

In the last step of visualizing the model's predictive capabilities, we plotted the stability Number (N) versus Hydraulic Radius (HR) for real data points and the ones predicted by the constructed ANN model. In the development process, we considered all the parameters that combine into stability number N separately to improve the performance and investigate the parameters' influence. After achieving the satisfying results, we determined the stability number N for the predicted cases to present the results and compare them visually with the real data points. Figures 5.17 and 5.18 illustrate the stability graph N vs HR for real and predicted cases, respectively.

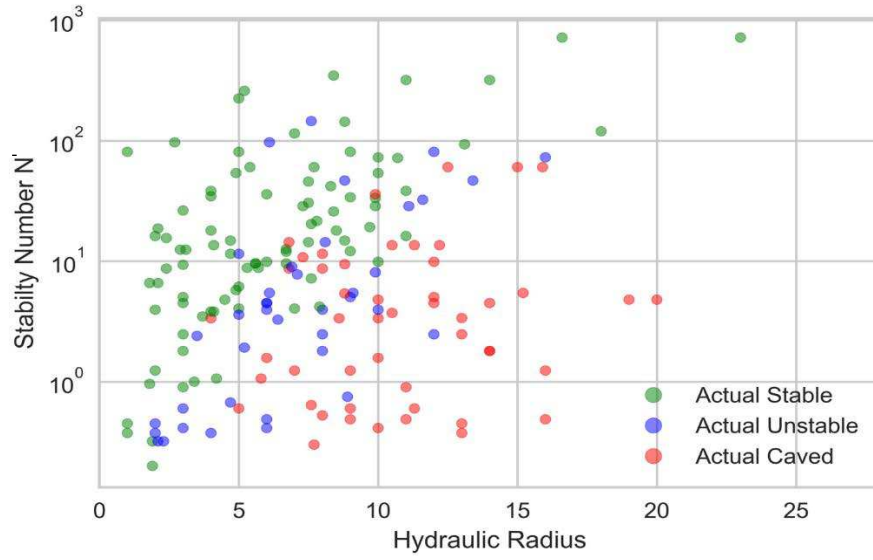


Figure 5.17 N' vs HR plot for actual classes

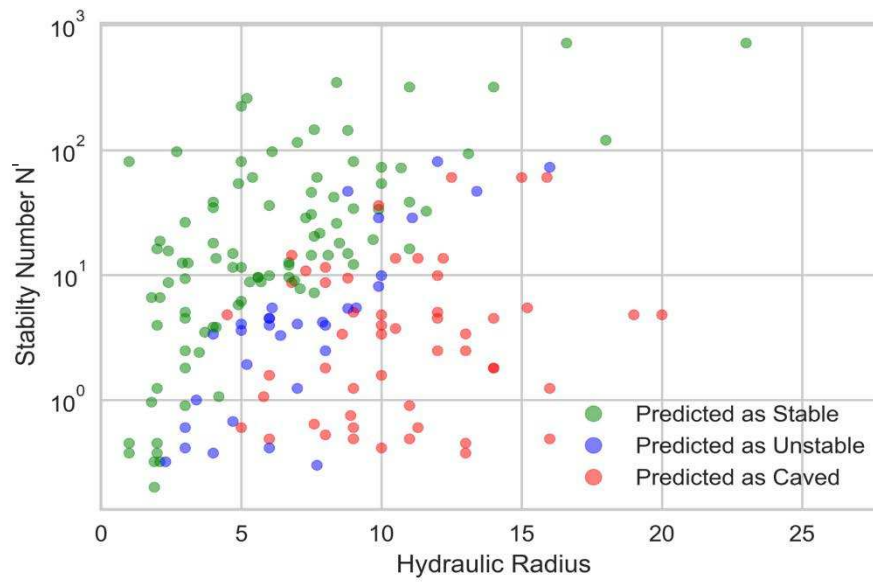


Figure 5.18 N' vs HR plot for predicted classes

Plotting those graphs confirmed that using models such as ANN allows to capture more complex relationships in data, which are difficult or impossible to classify with a classic Matthew Stability Graph method properly. ANN was able to correctly categorize points overlapping significantly, which would create a considerable challenge for classic methods based on linearly separating classes. This showcases the strength of ANN in handling complicated data structures where linear separation methods might fail. Artificial neural networks provide a more adjustable and accurate approach to classification tasks in scenarios like stability prediction in mining engineering.

Overall, the model's performance in classifying cases into the stable, unstable, and caved classes is highly promising, indicating the model's ability to make accurate predictions. This assertion is supported by the comprehensive evaluation of the classification report, the confusion matrix and stability plots N vs HR . These combined results affirm that the model exhibits promising predictive capabilities and performs exceptionally well in accurately categorizing instances into the appropriate classes.

5.4.2 Exploring feature importance with SHAP

Investigating the impact of each feature in a dataset on the final output is a crucial step in model interpretation and understanding, particularly in mining engineering, where these features can substantially influence underground stability. One powerful tool for this analysis is SHAP (Shapley Additive exPlanations) values. SHAP values comprehensively explain the importance of features and their contribution to the model's predictions. They are based on the game-theoretic concept of Shapley values, which ensure fairness and consistency in assigning feature importance. Lundberg and Lee (2017) introduced the concept of SHAP values, which has gained significant attention in machine learning interpretability. The SHAP values tool allows us to visualize and quantify the contribution of each feature, enabling an understanding of how different variables influence the model's output. By utilizing SHAP values in kernel analysis, we can derive valuable insights about feature effects, identify influential factors, and make informed decisions for model improvement and feature engineering. The features are arranged in descending order based on their importance in a SHAP summary plot. The stability estimation using the SHAP method took

approximately 20 seconds of computational time, reflecting the efficiency of the analysis attributed to the relatively small size of the dataset. The length of each bar represents the magnitude of the feature's effect.

The SHAP summary plot for the database showed that HR (Hydraulic Radius) followed by Q value has the most considerable influence on the model performance. The SHAP summary plot in Figure 5.19 revealed insightful findings regarding the importance of features in the investigated database. It indicated that the Hydraulic Radius (HR) feature, followed by the Q value, substantially influences the model's performance. The HR, representing the ratio of the underground opening dimensions, emerged as the most influential factor, highlighting its significant impact on the stability assessment of underground excavation. Variations in HR can significantly impact stress distribution within the rock mass, potentially leading to localized stress concentrations and increased susceptibility to failure.

The SHAP plot also provides valuable insights into the influence of input parameters on model predictions across different classes. Notably, the analysis reveals that the HR parameter has the greatest impact on both stable and caved classes. This suggests that the shape factor plays a significant role in determining the overall stability of excavations, regardless of their stability status. However, when examining the unstable class, the Q value emerges as the most influential factor. The observed differences in parameter importance across different stability classes highlight the nuanced interactions between various factors and the complex nature of stability predictions in underground mining operations.

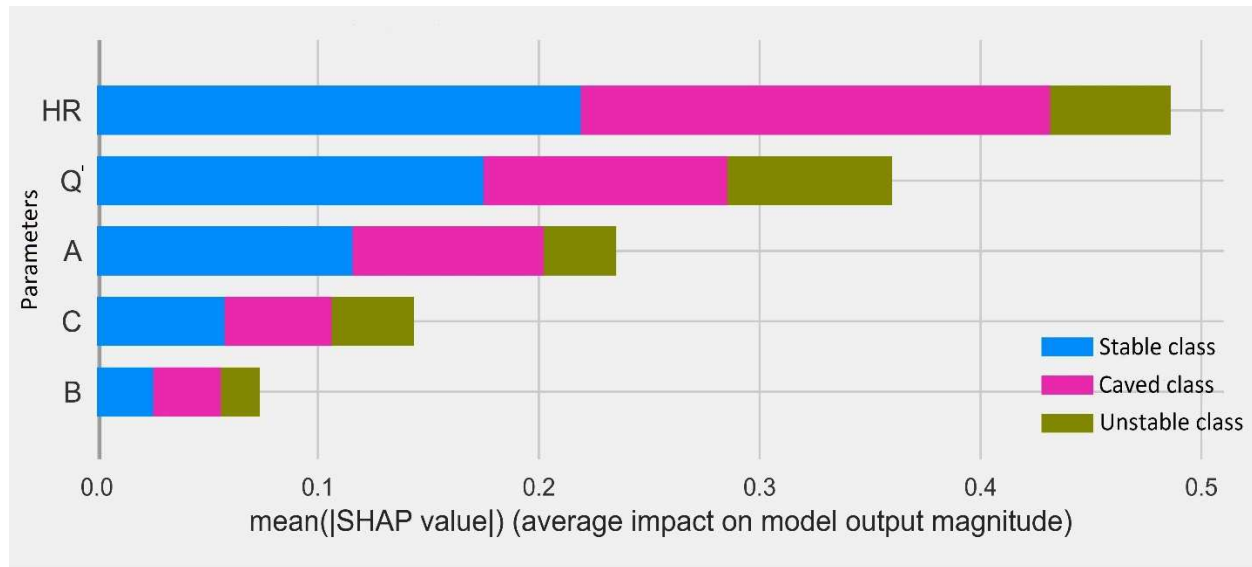
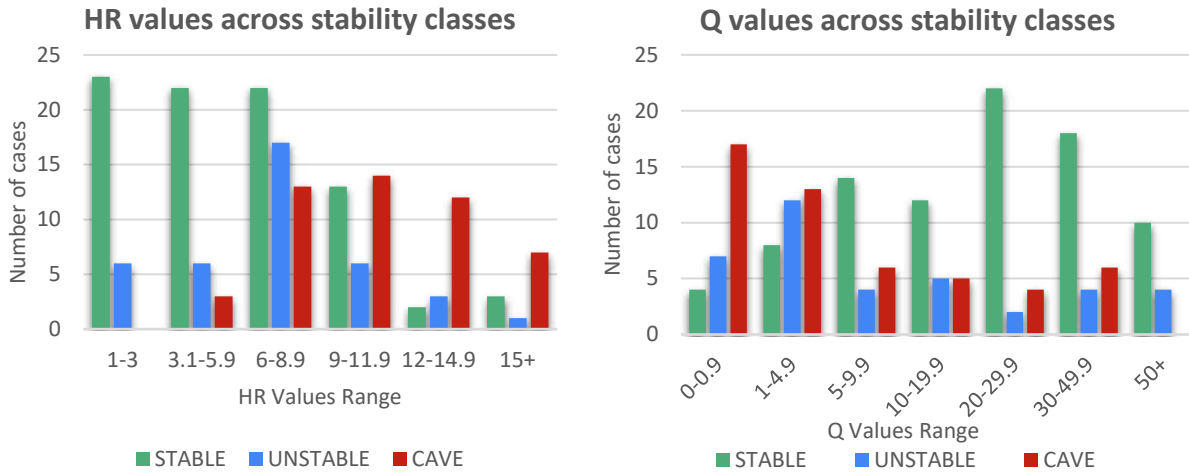


Figure 5.19 SHAP Summary Plot: Feature Importance Analysis

To conduct a more comprehensive analysis of the results, I employed plots (Figure 5.20 and 5.21) to visualize the distribution of HR and Q values across different stability classes for all cases in the database. These plots provided us with a clear representation of the frequency distribution of cases within specific ranges of HR and Q values. This approach enabled us to identify any distinct trends or patterns in the distribution of HR and Q values across different stability classes, thereby enhancing the understanding of the relationship between these parameters and SHAP plot results.



Figures 5.20 and 5.21 Distribution of HR and Q values across stability classes

Comparing the insights derived from the distribution plots of HR values with the SHAP analysis results provides a deeper understanding of the relationship between HR and slope stability. The plots reveal that a significant proportion of stable cases are concentrated in smaller HR value ranges, with approximately 75% of stable cases having HR values below 9. As HR values increase, there is a notable rise in the number of caved cases, suggesting a correlation between higher HR values and increased instability. The SHAP analysis highlights HR as the parameter with the highest impact on model predictions, particularly for stable and caved classes, which aligns with the observation from the distribution plots which show that values for these two classes are more consistent. Additionally, the unstable class exhibits a more uniform distribution across all HR ranges, and the SHAP plot indicates that this class has the lowest impact on model predictions. Contrary to HR, the distribution plots reveal that higher values of Q are associated with a better slope stability. Specifically, the absence of caved cases with Q values exceeding 50 suggests that higher Q values correspond to a lower likelihood of a slope collapse. The SHAP analysis indicates also that Q has a higher predictive impact on the unstable class compared to HR. Interestingly, the figure 5.21 illustrates a more uniform distribution of the unstable class across Q values, implying that variations in Q contribute significantly to the model's predictions for unstable cases. These contrasting trends between HR and Q underscore the complex relationship between geological, mechanical and shape parameters in influencing slope stability. The observed highest impact of

HR and Q on the stable class can be also attributed to the distribution of cases within the dataset. With approximately half of all cases classified as stable, the model is exposed to a significant proportion of instances representing this particular class. As a result, the model is better equipped to learn the characteristics associated with the stable class, leading to a higher influence. This shows the importance of data balance and quantity in training machine learning models.

These findings underscore the critical role of HR and Q value in accurately predicting and assessing the stability of underground excavations, emphasizing the need for careful consideration and detailed analysis of these features in mining engineering. This suggests that optimizing the dimensions of the mining openings can substantially impact the stability of stopes in underground excavations. On the other hand, the Q value, representing the characteristics of the rock mass, is a factor over which engineers have limited control. It serves as a reminder that while we can influence certain aspects of the mining design, the properties of the rock mass itself have challenges that may be more difficult to avoid. Therefore, understanding the influence of both HR and the Q value is essential for mining engineers to make informed decisions and seek a balance between design adjustments and the characteristics of the rock mass to ensure the stability and safety of underground openings.

5.5. Summary

My study focused on developing an Artificial Neural Network (ANN) model for predicting underground excavations stability in mining engineering. I investigated the impact of various factors, particularly the Hydraulic Radius (HR), on the model's performance. By conducting evaluations, I obtained significant insights into the criticality of HR and the optimal configuration of the ANN model.

My results demonstrated that shape factor (HR), representing the ratio of mining opening dimensions, played a crucial role in determining stability outcomes. I found that adjusting HR values directly influenced the accuracy of stability predictions. Leveraging the predictive power of the ANN, engineers can now utilize the model to assess the impact of different HR values on stability. This knowledge empowers them to make informed decisions regarding the dimensions of mining openings, ultimately improving underground stability.

Additionally, I identified that a specific model configuration, incorporating the StandardScaler for data preprocessing and the Swish activation function for all hidden layers, yielded the best-predicting results. The StandardScaler normalized the feature values, reducing the impact of varying scales and enhancing the ANN's ability to capture underlying patterns. The Swish activation function, known for its non-linearity and smoothness, facilitated the model in capturing complex relationships between input features and stability outcomes. In this model configuration, the average difference in accuracy between the training and validation sets across all epochs was approximately 7%. Moreover, the average percentage difference for loss between the training and validation sets was only 5% throughout all the epochs. These findings underscore the consistency and efficiency of my model's performance across various stages of the training process and indicate that the undesired problem of model overfitting was avoided. My study highlights the importance of investigating the impact of individual features, such as HR, on the final model output, which enables professionals to identify critical variables, leading to improved risk management and operational efficiency in underground mining. By incorporating such insights into the ANN, mining engineers can optimize HR values to improve stability. These findings underline the potential of machine learning techniques in assisting engineers to make data-driven decisions and enhance underground excavation design.

In conclusion, my study contributes to the field of mining engineering by showcasing the significance of the dimensions of the openings and providing an optimized ANN model configuration. By leveraging the model's predictive capabilities, engineers can better understand the relationship between HR and stability, leading to more informed design decisions and improved underground stability in mining operations.

5.6 References

- Adoko AC, Saadaari F, Mireku-Gyimah D, Imashev A (2022) A Feasibility Study on The Implementation of Neural Network Classifiers for Open Stope Design. *Geotech Geol Eng* 40(2):677–696. <https://doi.org/10.1007/s10706-021-01915-8>
- Adoko AC, Yakubov K, Alipov A (2019) Mine Stope Performance Assessment in Unfavorable Rock Mass Conditions Using Neural Network-Based Classifiers
- Agarap AF (2018) Deep Learning using Rectified Linear Units (ReLU). <https://doi.org/10.48550/ARXIV.1803.08375>
- Ahsan MM, Mahmud MAP, Saha PK, Gupta KD, Siddique Z (2021) Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance. *Technologies* 9(3). <https://doi.org/10.3390/technologies9030052>
- Arifin F, Robbani H, Annisa T, Ma'arof NNMI (2019) Variations in the Number of Layers and the Number of Neurons in Artificial Neural Networks: Case Study of Pattern Recognition. *J Phys: Conf Ser* 1413(1):012016. <https://doi.org/10.1088/1742-6596/1413/1/012016>
- Barton N, Lien R, Lunde J (1974) Engineering classification of rock masses for the design of tunnel support. *Rock Mechanics* 6(4):189–236. <https://doi.org/10.1007/BF01239496>
- Bishop CM (1995) *Neural networks for pattern recognition*. Clarendon Press ; Oxford University Press, Oxford : New York
- Bishop CM (2006) *Pattern recognition and machine learning*. Springer, New York
- Brownlee J (2016) *Machine learning mastery with Python: understand your data, create accurate models, and work projects end-to-end*. Machine Learning Mastery
- Capes GW (2009) Open stope hangingwall design based on general and detailed data collection in unfavourable hangingwall conditions, NR62618 Ph.D. The University of Saskatchewan (Canada)
- Chavlis S, Poirazi P (2021) Drawing Inspiration from Biological Dendrites to Empower Artificial Neural Networks. <https://doi.org/10.48550/ARXIV.2106.07490>
- Clark LM (1998) Minimizing dilution in open stope mining with a focus on stope design and narrow vein longhole blasting. <https://doi.org/10.14288/1.0081111>
- Deere DU (1963) Technical description of rock cores for engineering purposes. *Rock Mechanics and Engineering Geology* :1:16-22
- Diederichs MS, Kaiser PK (1999) Tensile strength and abutment relaxation as failure control mechanisms in underground excavations. *International Journal of Rock Mechanics and Mining Sciences* 36(1):69–96. [https://doi.org/10.1016/S0148-9062\(98\)00179-X](https://doi.org/10.1016/S0148-9062(98)00179-X)

Duan J (2019) Financial system modeling using deep neural networks (DNNs) for effective risk assessment and prediction. *Journal of the Franklin Institute* 356(8):4716–4731. <https://doi.org/10.1016/j.jfranklin.2019.01.046>

Elliott DL (1993) A Better Activation Function for Artificial Neural Networks. University of Maryland. Systems Research Center

Erdogan Erten G, Bozkurt Keser S, Yavuz M (2021) Grid Search Optimised Artificial Neural Network for Open Stope Stability Prediction. *International Journal of Mining, Reclamation and Environment* 35(8):600–617. <https://doi.org/10.1080/17480930.2021.1899404>

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. *Front Energy Res* 9:652801. <https://doi.org/10.3389/fenrg.2021.652801>

Goodfellow I, Bengio Y, Courville A (2016) Deep learning. The MIT Press, Cambridge, Massachusetts

Grandini M, Bagli E, Visani G (2020) Metrics for Multi-Class Classification: an Overview

Grimstad E, Barton NR (1993) Updating of the Q-System for NMT. *Proceedings of the International Symposium on Sprayed Concrete* :46–66

Hagan MT, Demuth HB, Beale MH, De Jesús O (2014) Neural network design, 2nd edition. Martin T. Hagan, s.L

Hancock JT, Khoshgoftaar TM (2020) Survey on categorical data for neural networks. *J Big Data* 7(1):28. <https://doi.org/10.1186/s40537-020-00305-w>

Heidarzadeh S, Saeidi A, Rouleau A (2019) Evaluation of the effect of geometrical parameters on stope probability of failure in the open stoping method using numerical modeling. *International Journal of Mining Science and Technology* 29(3):399–408. <https://doi.org/10.1016/j.ijmst.2018.05.011>

Henning JG (2007) Evaluation of long-hole mine design influences on unplanned ore dilution. Department of Mining and Materials Engineering, McGill University

Henning JG, Mitri HS (2007) Numerical modelling of ore dilution in blasthole stoping. *International Journal of Rock Mechanics and Mining Sciences* 44(5):692–703. <https://doi.org/10.1016/j.ijrmms.2006.11.002>

Henning JG, Mitri HS (2008) Assessment and Control of Ore Dilution in Long Hole Mining: Case Studies. *Geotech Geol Eng* 26(4):349–366. <https://doi.org/10.1007/s10706-008-9172-9>

Hughes R (2011) Factors influencing overbreak in narrow vein longitudinal retreat mining. Department of Mining and Materials Engineering, McGill University, Canada

Idris MA, Saiang D, Nordlund E (2011) Probabilistic analysis of open stope stability using numerical modelling. *IJMME* 3(3):194. <https://doi.org/10.1504/IJMME.2011.043849>

Karlik B, Olğaç A (2011) Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks

Kingma DP, Ba J (2014) Adam: A Method for Stochastic Optimization. <https://doi.org/10.48550/ARXIV.1412.6980>

Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*:Pages 4768-4777

Mathews KE, Hoek E, Wyllie DC, Stewart, S (1980) Prediction Of Stable Excavation Spans for Mining at Depths below 1000 Metres in Hard Rock. Ottawa, ON

Mawdesley C, Trueman R, Whiten WJ (2001) Extending the Mathews stability graph for open-stope design. *Mining Technology* 110(1):27–39. <https://doi.org/10.1179/mnt.2001.110.1.27>

Mitri H, Hughes R, Lecomte E (2010) Factors influencing unplanned ore dilution in narrow vein longitudinal mining. :448–454

Mitri HS, Hughes R, Zhang Y (2011) New rock stress factor for the stability graph method. *International Journal of Rock Mechanics and Mining Sciences* 48(1):141–145. <https://doi.org/10.1016/j.ijrmms.2010.09.015>

Nawi NM, Atomi WH, Rehman MZ (2013) The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks. *Procedia Technology* 11:32–39. <https://doi.org/10.1016/j.protcy.2013.12.159>

Pakalnis RT (1986) Empirical stope design at the Ruttan Mine, Sherritt Gordon Mines Ltd. <https://doi.org/10.14288/1.0081095>

Papaioanou A, Suorineni FT (2016) Development of a generalised dilution-based stability graph for open stope design. *Mining Technology* 125(2):121–128. <https://doi.org/10.1080/14749009.2015.1131940>

Pascanu R, Mikolov T, Bengio Y (2012) On the difficulty of training Recurrent Neural Networks. <https://doi.org/10.48550/ARXIV.1211.5063>

Patro SGK, Sahu KK (2015) Normalization: A Preprocessing Stage. <https://doi.org/10.48550/ARXIV.1503.06462>

Potvin Y (1988) Empirical open stope design in Canada. <https://doi.org/10.14288/1.0081130>

Qi C, Fourie A, Du X, Tang X (2018) Prediction of open stope hangingwall stability using random forests. *Nat Hazards* 92(2):1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>

- Raju VNG, Lakshmi KP, Jain VM, Kalidindi A, Padma V (2020) Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification. In: 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, Tirunelveli, India, pp 729–735
- Ramachandran P, Zoph B, Le QV (2017) Searching for Activation Functions. <https://doi.org/10.48550/ARXIV.1710.05941>
- Raschka S, Mirjalili V (2019) Python machine learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow 2, Third edition. Packt, Birmingham Mumbai
- Ren Y, Zhao P, Sheng Y, Yao D, Xu Z (2017) Robust Softmax Regression for Multi-class Classification with Self-Paced Learning. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Melbourne, Australia, pp 2641–2647
- Santos AEM, Amaral TKM, Mendonça GA, Silva DDFSD (2020) Open stope stability assessment through artificial intelligence. REM, Int Eng J 73(3):395–401. <https://doi.org/10.1590/0370-44672020730012>
- Stathakis D (2009) How many hidden layers and nodes? International Journal of Remote Sensing 30(8):2133–2147. <https://doi.org/10.1080/01431160802549278>
- Stewart PC, Trueman R (2004) Quantifying the effect of stress relaxation on excavation stability. Mining Technology 113(2):107–117. <https://doi.org/10.1179/037178404225004986>
- Suorineni FT, Papaioanou A (2015) Dilution-based Stability Graph for Open Stope Design. Mining Education Australia – Research Projects Review 4:27–33
- Suorineni FT (1998a) Effects of faults and stress on open stope design, Ph.D. Waterloo, Ontario, Canada
- Szmigiel A, Apel DB (2022) Predicting the stability of open stopes using Machine Learning. Journal of Sustainable Mining 21(3):241–248. <https://doi.org/10.46873/2300-3960.1369>
- Wang J, Milne D, Pakalnis R (2002) Application of a neural network in the empirical design of underground excavation spans. Mining Technology 111(1):73–81. <https://doi.org/10.1179/mnt.2002.111.1.73>
- Wang J, Milne D, Wegner L, Reeves M (2007) Numerical evaluation of the effects of stress and excavation surface geometry on the zone of relaxation around open stope hanging walls. International Journal of Rock Mechanics and Mining Sciences 44(2):289–298. <https://doi.org/10.1016/j.ijrmms.2006.07.002>
- Zhang G, Eddy Patuwo B, Y. Hu M (1998) Forecasting with artificial neural networks: International Journal of Forecasting 14(1):35–62. [https://doi.org/10.1016/S0169-2070\(97\)00044-7](https://doi.org/10.1016/S0169-2070(97)00044-7)

Zhou Q, Liu D, Lin X (2022) Pre-evaluation of fault stability for underground mining based on geomechanical fault-slip analysis. *Geomatics, Natural Hazards and Risk* 13(1):400–413.
<https://doi.org/10.1080/19475705.2022.2032401>

CHAPTER 6: EXPLORING MACHINE LEARNING TECHNIQUES FOR OPEN STOPE STABILITY PREDICTION: A COMPARATIVE STUDY AND FEATURE IMPORTANCE ANALYSIS

Study presented in this chapter investigates and compares several machine learning algorithms. By analyzing a dataset comprising stope dimensions and geomechanical properties, I explore the potential of machine learning models such as Random Forest, Support Vector Machine, AdaBoost, XGBoost, LightGBM, and Artificial Neural Network in predicting stope stability. Evaluation metrics including accuracy, precision, recall, and F1 score are employed to assess model performance, with the Artificial Neural Network emerging as the most effective. Furthermore, SHapley Additive exPlanations (SHAP) analysis enhances interpretability by explaining the contribution of individual features to model predictions.

6.1 Introduction

Understanding the stability of underground excavations is important for ensuring the safety, durability, and longevity of mining operations. Over the years, classical stability assessment methods have played a crucial role in evaluating the stability of stopes in underground mining. These methods, rooted in well-established principles and empirical formulas derived from extensive field observations and laboratory tests, have provided valuable insights into rock mass behavior. Among the various empirical approaches utilized in assessing stope stability, stability graphs, as proposed by Matthew et al. (1980), have emerged as particularly popular. These graphs are built upon rock mass classification systems such as the Q value system by Barton et al. (1974) and the rock mass rating (RMR) introduced by Bieniawski (1973). By integrating factors like the rock stress factor (A), joint orientation adjustment factor (B), and surface orientation factor (C), stability graphs facilitate the calculation of the stability number N, a crucial parameter for designing stope dimensions and support.

However, in addition to these classical methods, modern advancements in machine learning have opened up new solutions for stope stability assessment. Research conducted by Adoko et al. (2022) demonstrates the feasibility of utilizing feed forward neural network classifiers to predict stope stability. This study, based on a database obtained from Adoko et al. (2022), comprising of 225 cases from three mines in Ghana, West Africa, highlights the potential of machine learning in enhancing predictive capabilities and understanding complex relationships among various parameters influencing stope stability.

In my study, I analyze this dataset, by exploring and comparing several machine learning algorithms to further improve predictive accuracy. By incorporating parameters such as shape dimensions and rock mechanical properties, I aim to gain a comprehensive understanding of the factors influencing stope stability. The comparative analysis includes machine learning models like Random Forest, Support Vector Machine (SVM), AdaBoost, XGBoost, LightGBM, and Artificial Neural Network (ANN). Through evaluation and examination of metrics such as accuracy, precision, recall, and F1 score, I identify the strengths and weaknesses of each model in predicting stope stability. Furthermore, I employ SHapley Additive exPlanations (SHAP) analysis to gain interpretability into the predictions of each model. By analyzing the contribution of individual features to the model's predictions, SHAP analysis enhances transparency and reliability,

facilitating improvements in stope design strategies and engineering efforts aimed at ensuring mining safety and efficiency.

6.2 Background of stope stability assessment

Understanding the factors that influence stability is crucial for ensuring the safety, durability, and longevity of underground excavations. Classical stability assessment methods have long been employed to evaluate the stability of stopes in mining operations. These methods are based on well-established principles and empirical formulas derived from extensive field observations and laboratory tests.

Stability graphs, initially proposed by Matthew et al., (1980) represent one of the most popular empirical approaches utilized in assessing the stability of stopes in underground mining excavations. The method was developed based on popular and widely used rock mass classification systems such as Q value system presented by Barton et al. (1974) and rock mass rating (RMR) proposed by Bieniawski (1973). Matthew stability graph method revolves around determining crucial factors that influence the stability of the rock mass by utilizing specifically developed graphs that relate various characteristics and properties of the rock. These graphs facilitate factors such as the rock stress factor (A), joint orientation adjustment factor (B), and surface orientation factor (C). These factors are then combined to calculate stability number N, which is critical parameter developed for designing stope dimensions and support in underground mining openings. It serves as a quantitative indicator of the physical conditions and stability of the stopes and it is calculated as follows:

$$N' = Q' \cdot A \cdot B \cdot C \quad (6.1)$$

The crucial factor for the successful assessment of opening stability in underground mining is the shape factor, which relates the dimensions of the opening and is commonly referred to as the hydraulic radius (HR). The hydraulic radius is a critical geometric parameter that characterizes the shape of the opening and is essential in determining its stability.

The term hydraulic radius is commonly understood as the ratio of the area of exposure of the hanging wall to its perimeter. In the context of inclined stopes, where the stope is not in a vertical position, the most critical aspect for calculating the HR is the exposure of the hanging wall (Figure 6.1). The calculation of HR takes into account the spans of the stopes along the dip (h) and along the strike (w) (Tishkov 2018), as shown in Figure 6.1.

The final graph developed by Matthew for stope assessment plots the stability number N against the hydraulic radius HR . Each case in the dataset has been categorized as either stable, unstable, or caved based on previous evaluations of real cases. As a result, the graph is divided into three distinct zones, representing the stability assessments for the stopes.

This graph serves as a valuable tool for classifying new stopes into one of these three stability categories. When a new stope is to be assessed, its corresponding values of N and HR can be plotted on the graph. By examining the location of these values on the graph, it can easily be determined whether the new stope falls into the stable, unstable, or caved zone. This classification aids in understanding the potential stability conditions of the new stope and guides the selection of appropriate support measures to ensure its safety and structural integrity (Suorinen et al. 2001).

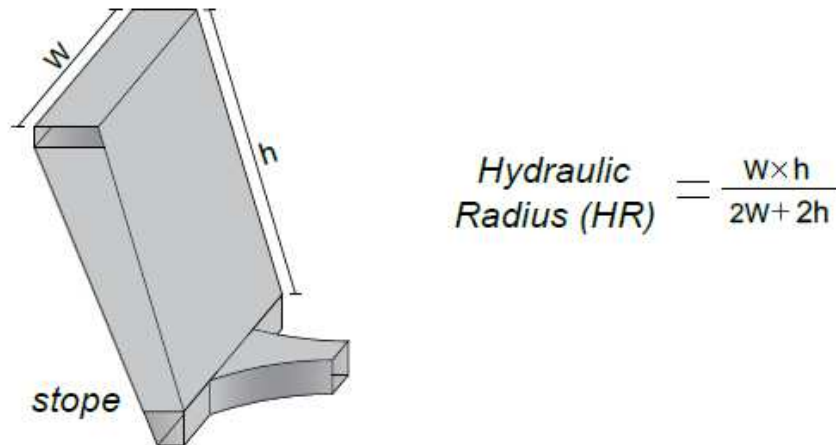


Figure 6.1. Calculation of Hydraulic radius

Visual separation of zones in stability assessment graphs can pose challenges related to subjectivity and reproducibility. Human judgment may introduce variations in categorizing stopes into stable, unstable, or caved zones, leading to discrepancies in the analysis. Additionally, the inherent risk associated with potential errors in categorization raises concerns about the reliability of the assessments. Empirical design methods have the potential to continuously improve and evolve over time as more data becomes available and engineers gain increased experience with the method. The nature of empirical design allows for flexibility and adaptability, enabling engineers to refine and update the approach based on new observations and findings.

Since 1980, significant research in mining engineering have contributed to enhancing the reliability and effectiveness of stability graphs for assessing underground mining excavations. These developments have been crucial in refining the stability graph method and making it a more robust and trusted tool in mining practice. In 1988, Potvin's approach introduced a refinement to the stability graph, departing from the traditional three-zone classification system proposed by Mathews et al. Instead, Potvin proposed the stability graph with two main zones, stable and caved. However, he also incorporated an additional transition zone, representing a critical boundary between the stable and caved regions (Potvin 1988). Potvin acknowledged that the stability graph can be affected by human bias and unknown inherent errors when visually defining its zones and suggested utilizing statistical tools for zone definition instead.

Nickson (1992) was a pioneer in attempting to establish the boundary positions through statistical methods. He employed discriminant analysis on the multivariate stability database, a three-dimensional dataset and used Mahalanobis' distance to divide the data into distinct groups. He achieved this by deriving a linear separation between stable and caved unsupported scenarios, utilizing a logarithmic transformation. Notably, his analysis excluded unstable cases, and he did not determine separation lines for zones involving unstable or caving situations. Nickson also compared his statistically determined boundary, which separates stable and caved conditions, with Potvin's proposed transition zone. Based on his findings, Nickson recommended that Potvin's transition zone should be employed for designing unsupported stope surfaces. Hadjigeorgiou et al. (1995) gathered additional stability data and conducted a repeat analysis using discriminant methods, yielding comparable outcomes. Suorineni et al. (2001) introduced the Bayesian likelihood method as a powerful tool for statistically interpreting the stability graph. They

employed an extended database based on the Potvin-calibrated stability graph factors to illustrate the method's advantages. The Bayesian likelihood discrimination proved to be an optimal approach for statistically interpreting the stability graph due to its capability to reveal substantial overlap among the defined stability graph zones (stable, unstable, and caving). It also allowed for error rate estimation in the stability graph, delineated general transition boundaries between stable, unstable and caved stopes, estimated inherent predictive errors in stability graphs, evaluated the risk associated with using the stability graph for predictions, and introduced a multiple design curves stability graph based on the probability. The Bayesian likelihood discrimination's ability has been harnessed to provide deeper significance to the class boundaries in the stability graph and individual stope walls plotted within each class.

Numerical modeling stands as another widely adopted approach that has proven effective in addressing stope stability concerns. Henning and Mitri (2007), for instance, crafted a series of three-dimensional numerical models to explore the impacts of field stress, mining depth, stope configuration, and orientation on stope wall overbreak. Similarly, Purwanto et al. (2013) harnessed numerical modeling to establish the correlation between stope design and the stability of hanging wall. Hu and Cao (2009), by employing visual numerical simulation software, simulated and computed stress distribution and displacement variations within stopes during mining operations. They conducted an analysis of the stability of stope roofs and adjacent rock, as well as the alterations in sound emission associated with the mining process.

While classical stability assessment methods have been effective in many cases, the emergence of machine learning techniques has opened up new possibilities for enhancing their accuracy and efficiency. Machine learning algorithms can analyze large volumes of mining data, including shape of the opening, the properties of surrounding rock mass, underground conditions and historical stability records, to identify patterns and correlations that may not be easily discernible through traditional methods. By training machine learning models on a dataset of known stability outcomes, engineers can develop predictive models that can assess the stability of new stopes.

Several studies have explored the integration of machine learning models for the prediction of stope stability. Erdogan Erten et al. (2021) introduced a hybrid artificial neural network (ANN) approach optimized through grid search. This method was compared with conventional techniques

including Naive Bayes (NB), Decision Tree (DT), k-Nearest Neighbors (kNN), Support Vector Machine (SVM), as well as the traditional stability graph method. The findings of this study revealed that the performance of the stability graph method falls short of the capabilities exhibited by machine learning algorithms. Notably, the ANN model with hyper-parameters tuned using the grid search technique showcased superior performance in terms of accuracy, precision, recall, f-measure, and g-mean compared to other machine learning algorithms. Saadaari et al. (2020) investigated the viability of employing Ensemble Learning methods to categorize and predict the stability condition of slope surfaces. They introduced and evaluated four techniques - Random Forest (RF), Gradient Boosting (GB), Bootstrap Aggregating Classifier (BAC), and Adaptive Boosting (AB) - using widely accepted and effective assessment metrics. Upon analyzing the performance outcomes, it was evident that among the four machine learning models, Gradient Boosting (GB) and Bootstrap Aggregating Classifier (BAC) demonstrated the highest efficacy in accurately classifying and predicting the stability state of slopes, encompassing categories of caved, stable, or unstable. In a comparative investigation conducted by Qi et al. (2018b), five distinct artificial intelligence strategies based on machine learning and metaheuristic algorithms were explored for their potential in predicting the stability of open slope hangingwalls (HW). The assessed algorithms encompassed logistic regression (LR), multilayer perceptron neural networks (MLPNN), decision tree (DT), gradient boosting machine (GBM), and support vector machine (SVM). The optimization of hyperparameters was facilitated using the Firefly algorithm (FA), which yielded successful results for this purpose. Across the testing phase, the most favorable performance was exhibited by the optimized GBM model, closely followed by the SVM model and the optimized LR model. The study highlighted the remarkable predictive capabilities of these three machine learning models in forecasting HW stability. Several researchers have directed their efforts toward the refinement and customization of specific machine learning models for assessing slope stability. Qi et al. (2018a), in their study, concentrated on optimizing the Random Forest model for enhanced efficiency, while Santos et al. (2020) shifted their attention toward the utilization of Artificial Neural Networks.

It is significant to continue research in the area of slope stability in underground mining. The development of mining engineering is constantly moving towards increased exploitation and most favorable optimization, this is causing the industry to decide on larger sized opening thus

reducing the quantity of the stopes. This approach, even though mostly efficient for rapid extraction and profit growth, could have a negative impact on the stability of the opening. When dimensions are exceeded to their maximum, the stability of the stopes is reduced causing them to shift toward unstable condition or even catastrophic failure.

At the stage of mine planning and stopes design, it is crucial to consider and evaluate all the parameters that have direct or non-direct influence on the stability condition. The integration of machine learning with classical stability assessment methods offers several advantages in that matter. It allows for more comprehensive and data-driven evaluations, considering a wider range of variables and their interactions, that might have crucial impact on the stability of each open stope. Machine learning models can also help identify complex relationships between shape factor, rock properties, and stability outcomes that may not be apparent using traditional approaches. Hence the necessity to investigate the stope stability data further, in order to determine which parameter has the most significant impact on the stability of a stope and how to use that knowledge to increase safety in mining environment.

6.3 Database analysis and pre-processing

Analyzing a database and proper preprocessing is a crucial step in the machine learning process and have significant impact on the success of any machine learning model. It is important to assess the quality and characteristics of the database before employing any predictive model. Real-world data often contains missing values, outliers, duplicates, and inaccuracies that have negative impact on the model performance. Analyzing the database helps identify and address these issues, ensuring that the data used especially for training is reliable and accurate. Proper assessment provides insights into characteristics of the database and its features, it allows us to determine the type of each attribute, whether they are numerical or categorical, and understand the distribution of values. This understanding is crucial for choosing preprocessing techniques followed by implementation of appropriate machine learning algorithms. Different machine learning models create different assumptions about the database, analysis allows us to ensure that the chosen algorithm aligns with the characteristics and complexity of the data. Presented by Crone et al. (2006) investigation, have shown a strong evidence that preprocessing techniques have critical

impact on the predictive performance of the models. The results presented encourage to implement analysis of the database and preprocessing in order to produce valid and accurate outcomes of the classification algorithms.

The database investigated, comes from research conducted by Adoko A. C., Saadari F., Mireku-Gyimah D., and Imashev A., titled “A Feasibility Study on The Implementation of Neural Network Classifiers for Open Stope Design” published in 2022. The study utilized a feed forward neural network classifier to predict the stability of open stopes in underground mining obtaining an average accuracy of 91%. The general misclassification of the model (less than 10%) showed that FFNN outperformed the classic stability graph methods, which yielded the misclassification of almost 40%.

Database consist of 225 cases and was collected from three different mines in Ghana, West Africa, in the period of over three months. It includes information such as height, span and length of each stope, as well as geomechanical properties: Q' value, rock stress factor (A), joint orientation adjustment factor (B) and surface orientation factor (C). The features that were passed to the algorithm were modified stability number N' and hydraulic radius HR determined from the database. Figure 6.2 illustrates the distribution of cases. A plot of stability number N versus hydraulic radius was created, where different colors were employed to signify whether a specific case was classified as stable, unstable, or caved. It provided us with a visual representation of how the data is spread across different classes. Analysis of the plot revealed that cases with grater hydraulic radius tend to exhibit a higher tendency for instability and caving. As for the stability number, its influence on stability is somewhat less pronounced. However, instances with exceptionally low stability number values were predominantly categorized as caved. This observation also provided us with valuable insight into the relationship between weak rock conditions and low stability number values. In such scenarios, decreasing the dimensions of the opening can have a notably favorable effect on stability.

Adoko et al. (2022) investigated the Feed Forward Neural Network algorithm, where the input parameters were stability number N and shape factor HR. While the outcomes of this investigation yielded satisfactory accuracy in predicting stope stability, there is still potential to delve deeper into the dataset. This further exploration could help in enhancing the predictive capabilities of the model. Additionally, it offers an opportunity to conduct a comprehensive examination of all the

constituent features that contribute to the computation of the shape factor HR and the stability number N. This extended analysis can provide a more in-depth understanding of the complex relationship among these features and their collective influence on slope stability. Therefore, in this study I investigated and compared several machine learning algorithms where the input data were all the parameters that combine into stability number N and HR. These parameters were: shape parameters - height of the slope (H), Span (S), Length (L) and Rock mechanical properties - Q' value, rock stress factor (A), joint orientation adjustment factor (B) and surface orientation factor (C). All the cases had a stability assessment determined: stable, unstable, caved. The complete database from Adoko et al. (2020) with all the parameters is shown in the Table 6.1.

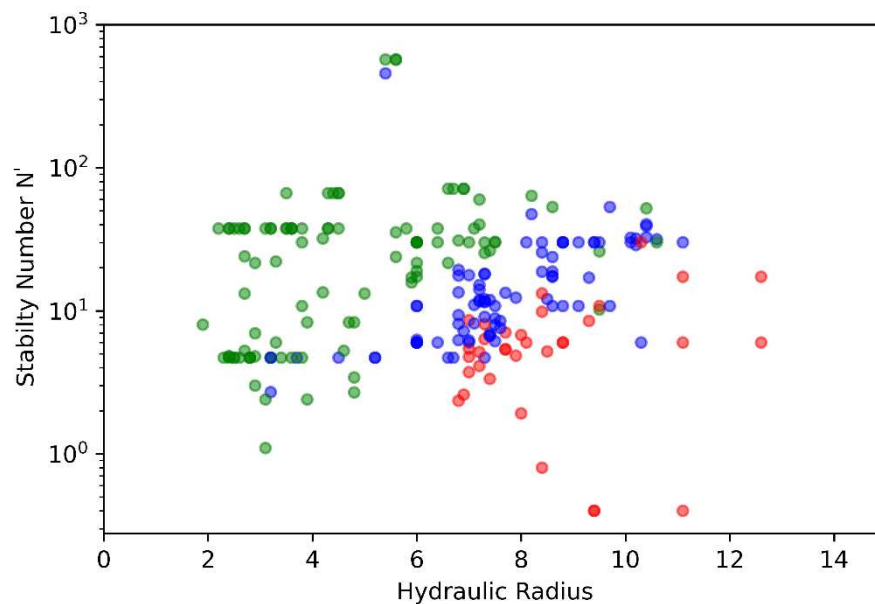


Figure 6.2 Distribution of cases in database. Green – stable cases, blue – unstable and red – caved.

Table 6.1 Complete Adoko et al. (2022) Database

Case number	Height (H)	Span (S)	Length (L)	Q'	A	B	C	Stability Asses.
1	30	11.9	30	4.7	1	1	8	STABLE
2	30	30	43	1.5	1	0.5	8	CAVED
3	30	11.9	20	4.7	1	1	1	UNSTABLE
4	30	20	30	1.5	1	0.5	8	UNSTABLE
5	30	20	30	4.7	1	0.8	8	STABLE
6	30	22.3	43	4.7	1	1	1	UNSTABLE
7	30	22.3	30	4.7	1	1	8	STABLE
8	30	30	43	4.7	1	0.8	8	UNSTABLE
9	30	10	20	4.7	1	1	4.7	STABLE
10	30	10	30	4.7	1	1	8	STABLE
11	30	20	30	1.5	1	0.5	8	UNSTABLE
12	30	20	30	4.7	1	0.8	8	STABLE
13	30	27	26	4.7	1	1	1	UNSTABLE
14	30	27	30	4.7	1	1	8	STABLE
15	30	26	30	1.5	1	0.5	8	UNSTABLE
16	30	26	30	4.7	1	0.8	8	STABLE
17	30	9.2	20	4.7	1	1	1	STABLE
18	30	9.2	30	4.7	1	1	8	STABLE
19	30	20	30	1.5	1	0.5	8	UNSTABLE
20	30	20	30	4.7	1	0.8	8	STABLE
21	30	6.4	43	4.7	1	1	1	STABLE
22	30	6.4	30	4.7	1	1	8	STABLE
23	30	30	43	1.5	1	0.5	8	CAVED
24	30	30	43	4.7	1	0.8	8	UNSTABLE
25	30	5.2	38	4.7	1	1	1	STABLE
26	30	5.2	30	4.7	1	1	8	STABLE
27	30	30	38	0.2	1	0.5	8	CAVED
28	30	30	38	4.7	1	0.8	8	UNSTABLE
29	30	6	85	4.7	1	1	1	STABLE
30	30	6	30	4.7	1	1	8	STABLE
31	30	30	85	0.1	1	0.5	8	CAVED
32	30	30	85	4.7	1	0.8	8	UNSTABLE
33	30	6.2	30.2	4.7	1	1	1	STABLE
34	30	6.2	22	4.7	1	1	8	STABLE
35	30	22	30.2	1.5	1	0.5	8	UNSTABLE
36	30	22	30.2	4.7	1	0.8	8	STABLE
37	30	6.6	20	4.7	1	1	1	STABLE
38	30	6.6	30	4.7	1	1	8	STABLE
39	30	20	30	2.7	1	0.5	8	UNSTABLE
40	30	20	30	4.7	1	0.8	8	STABLE

41	30	13	10	4.7	1	1	1	STABLE
42	30	13	30	4.7	1	1	8	STABLE
43	30	10	30	2.7	1	0.5	8	STABLE
44	30	10	30	4.7	1	0.8	8	STABLE
45	30	19	46	4.7	1	1	1	UNSTABLE
46	30	19	30	4.7	1	1	8	STABLE
47	30	30	46	2.7	1	0.5	8	UNSTABLE
48	30	30	46	4.7	1	0.8	8	UNSTABLE
49	60	7.2	60	4.7	1	1	1	STABLE
50	60	7.2	61	4.7	1	1	8	STABLE
51	60	61	35	1.5	1	0.5	8	CAVED
52	60	61	35	2.7	1	0.8	8	CAVED
53	60	8.1	43	4.7	1	1	1	STABLE
54	60	8.1	60	4.7	1	1	8	STABLE
55	60	61	43	1.5	1	0.5	8	CAVED
56	60	61	43	2.7	1	0.8	8	CAVED
57	60	7	83	4.7	1	1	1	UNSTABLE
58	60	7	60	4.7	1	1	8	STABLE
59	60	61	83	1.5	1	0.5	8	CAVED
60	60	61	83	2.7	1	0.8	8	CAVED
61	30	9.5	30	4.7	1	1	1	STABLE
62	30	9.5	30	4.7	1	1	8	STABLE
63	30	30	20	1.5	1	0.5	8	UNSTABLE
64	30	30	20	2.7	1	0.8	8	STABLE
65	30	6.6	20	4.7	1	1	1	STABLE
66	30	6.6	30	4.7	1	1	8	STABLE
67	30	30	20	2.7	1	0.5	8	UNSTABLE
68	30	30	20	4.7	1	0.8	8	STABLE
69	30	5.6	30	4.7	1	1	1	STABLE
70	30	5.6	30	4.7	1	1	8	STABLE
71	30	30	30	2.7	1	0.5	8	UNSTABLE
72	30	30	30	4.7	1	0.8	8	STABLE
73	30	9.3	43	4.7	1	1	1	STABLE
74	30	9.3	30	4.7	1	1	8	STABLE
75	30	30	43	2.7	1	0.5	8	UNSTABLE
76	30	30	43	4.7	1	0.8	8	UNSTABLE
77	30	12	35	4.7	1	1	1	UNSTABLE
78	30	12	30	4.7	1	1	8	STABLE
79	30	30	35	1.5	1	0.5	8	CAVED
80	30	30	35	4.7	1	0.8	8	UNSTABLE
81	30	18	60	71.3	1	1	1	STABLE
82	30	18	30	71.3	1	1	8	STABLE
83	30	30.5	60	8.1	1	0.5	8	UNSTABLE

84	30	30.5	60	4.7	1	0.8	8	UNSTABLE
85	30	17	60	71.3	1	1	1	STABLE
86	30	17	30	71.3	1	1	8	STABLE
87	30	31	60	8	1	0.5	8	UNSTABLE
88	30	31	60	4.5	1	0.8	8	UNSTABLE
89	30	18	60	71.2	1	1	1	STABLE
90	30	18	30	71.2	1	1	8	STABLE
91	30	33	60	7.9	1	0.5	8	UNSTABLE
92	30	33	60	4.7	1	0.8	8	STABLE
93	30	17	65	71.3	1	1	1	STABLE
94	30	17	30	71.3	0.8	1	8	UNSTABLE
95	30	30.5	65	6	1	0.8	6.8	UNSTABLE
96	30	30.5	65	13	1	0.5	8	STABLE
97	30	6.6	52	3	1	1	1	STABLE
98	30	6.6	30	3	1	1	8	STABLE
99	30	30	52	6	0.5	0.5	6.8	STABLE
100	30	30	52	13	0.5	0.5	8	STABLE
101	30	8.1	52	4.7	1	1	8	STABLE
102	30	8.1	30	4.7	1	1	8	STABLE
103	30	30	52	2.7	1	0.5	8	CAVED
104	30	30	52	4.7	1	0.8	8	UNSTABLE
105	30	13	50	4.7	1	1	1	UNSTABLE
106	30	13	30	8.3	1	1	8	STABLE
107	30	30	50	0.1	1	0.5	8	CAVED
108	30	30	50	4.7	1	0.8	8	UNSTABLE
109	30	13	50	4.7	1	1	1	UNSTABLE
110	30	13	30	8.3	1	1	8	STABLE
111	30	30	50	0.1	1	0.5	8	CAVED
112	30	30	50	4.7	1	0.8	8	UNSTABLE
113	107	15	60	2.7	1	1	8	STABLE
114	107	15	105	2.7	1	1	8	STABLE
115	107	107	60	8.3	1	0.5	8	CAVED
116	107	107	60	4.7	1	0.8	8	CAVED
117	30	7.1	65	2.7	1	1	1	UNSTABLE
118	30	7.1	30	2.7	1	1	8	STABLE
119	30	30	65	1.5	1	0.5	8	UNSTABLE
120	30	30	65	4.7	1	0.8	8	CAVED
121	30	9.2	55	8.3	1	1	1	STABLE
122	30	9.2	30	8.3	1	1	8	STABLE
123	30	30	55	2.7	1	0.5	8	UNSTABLE
124	30	30	55	8.3	1	0.8	8	UNSTABLE
125	30	12.5	40	8.3	1	1	1	STABLE
126	30	12.5	30	8.3	1	1	8	STABLE

127	30	30	40	2.7	1	0.5	8	UNSTABLE
128	30	30	40	8.3	1	0.8	8	STABLE
129	30	12.2	40	8.3	1	1	1	STABLE
130	30	12.2	30	8.3	1	1	8	STABLE
131	30	30	40	4.7	1	0.5	8	UNSTABLE
132	30	30	40	2.7	1	0.8	8	UNSTABLE
133	24	24	34	14.9	0.4	0.3	4.8	CAVED
134	26	26	33	17.1	0.5	0.3	4.5	UNSTABLE
135	27	27	38	25.4	0.6	0.3	2.7	UNSTABLE
136	26	26	30	14.4	0.8	0.2	7.7	UNSTABLE
137	25	25	34	29.8	1	0.3	6.7	STABLE
138	26	26	22	17	0.6	0.3	6.2	STABLE
139	31	31	27	36.6	0.3	0.3	4.6	UNSTABLE
140	32	32	35	17.1	0.4	0.3	4.8	CAVED
141	23	23	38	18	0.6	0.2	5.4	UNSTABLE
142	25	25	30	13.4	0.5	0.3	3.1	UNSTABLE
143	26	26	38	8.9	0.6	0.3	4.4	CAVED
144	26	26	34	11.7	0.4	0.2	7.4	UNSTABLE
145	27	27	35	18.5	0.3	0.3	5.1	UNSTABLE
146	27	27	32	24.7	0.6	0.3	5.7	STABLE
147	29	29	41	18.3	0.4	0.3	5.5	UNSTABLE
148	25	25	40	21.4	0.8	0.2	3.9	UNSTABLE
149	26	26	34	27.9	0.8	0.2	5.9	STABLE
150	26	26	31	13.1	0.8	0.3	3.5	UNSTABLE
151	30	30	30	8.7	0.8	0.2	4.4	UNSTABLE
152	32	32	28	36.4	0.8	0.2	5.2	STABLE
153	24	24	34	5.4	0.7	0.3	4.8	CAVED
154	26	26	33	6.4	0.6	0.3	5.5	CAVED
155	27	27	38	3.8	0.6	0.3	7.1	CAVED
156	26	26	30	6.8	1	0.2	3.5	CAVED
157	25	25	34	6.5	0.8	0.3	3.3	CAVED
158	26	26	22	11.6	0.5	0.3	3.6	UNSTABLE
159	31	31	27	34.7	0.7	0.3	5.5	STABLE
160	32	32	35	34.1	0.5	0.3	5	UNSTABLE
161	23	23	38	9.8	0.6	0.2	3.5	CAVED
162	25	25	30	7.4	0.6	0.3	7	UNSTABLE
163	26	26	38	5.3	0.6	0.3	5.6	CAVED
164	26	26	34	17.8	0.4	0.2	4.7	UNSTABLE
165	27	27	35	10.1	0.5	0.3	5	UNSTABLE
166	27	27	32	22.8	0.4	0.3	4.4	UNSTABLE
167	29	29	41	7.7	0.5	0.3	4.5	CAVED
168	25	25	40	9.9	0.7	0.2	3.9	CAVED
169	26	26	34	8.7	0.3	0.2	6.4	CAVED

170	26	26	31	11.1	0.7	0.3	3.5	UNSTABLE
171	30	30	30	12.4	0.7	0.2	5.1	UNSTABLE
172	32	32	28	15.4	0.7	0.2	3.7	UNSTABLE
173	25	25	30	46	0.6	0.2	3.5	UNSTABLE
174	25	25	30	32	0.6	0.2	3.5	UNSTABLE
175	25	25	30	44	0.8	0.2	4.4	STABLE
176	25	25	30	25	0.8	0.2	4.4	UNSTABLE
177	25	25	20	38	1	0.3	3.1	STABLE
178	25	25	20	25.6	1	0.3	3.1	STABLE
179	24	24	51	39	0.8	0.3	6.8	STABLE
180	24	24	51	29	0.8	0.3	6.8	UNSTABLE
181	29	29	30	32	0.3	0.2	6.2	UNSTABLE
182	29	29	30	18	0.3	0.2	6.2	UNSTABLE
183	25	25	51	14.9	0.6	0.2	7.4	CAVED
184	25	25	51	21.1	0.6	0.2	7.4	UNSTABLE
185	28	28	55	27.3	0.8	0.2	3.9	UNSTABLE
186	28	28	55	12.1	0.9	0.2	3.9	CAVED
187	26	26	41.5	12	0.6	0.2	4.7	CAVED
188	26	26	41.5	3.4	0.6	0.2	4.7	CAVED
189	25	25	35	20	1	0.3	5	STABLE
190	25	25	35	12	1	0.3	5	UNSTABLE
191	25	25	30	31	0.3	0.3	2.9	UNSTABLE
192	25	25	30	9	0.3	0.3	2.9	CAVED
193	25	25	35	10	0.9	0.2	6.4	UNSTABLE
194	25	25	35	7	0.9	0.2	6.4	CAVED
195	25	25	55	17.4	0.5	0.3	6.7	UNSTABLE
196	25	25	55	19.7	0.6	0.3	6.7	UNSTABLE
197	25	25	31	25	0.2	0.2	7.2	UNSTABLE
198	25	25	31	9	0.2	0.2	7.2	CAVED
199	26	26	30	20	0.5	0.2	3.1	UNSTABLE
200	26	26	30	12	0.5	0.2	3.1	CAVED
201	25	25	35	14	0.3	0.3	7.2	UNSTABLE
202	25	25	35	28	0.3	0.3	7.2	UNSTABLE
203	15	15	55	12	0.6	0.3	7.3	STABLE
204	15	15	55	13	0.6	0.3	7.3	STABLE
205	39	39	45	17.4	1	0.3	7.7	UNSTABLE
206	39	39	45	16.9	1	0.3	7.7	UNSTABLE
207	24.5	24.5	35	16	0.5	0.2	7.4	UNSTABLE
208	24.5	24.5	35	19	0.5	0.2	7.4	UNSTABLE
209	60	6.9	60	0.9	1	0.2	6.1	STABLE
210	60	7	58	1.5	1	0.2	8	STABLE
211	23	6.2	31	2	1	0.3	8	STABLE
212	23.4	6.1	31	2	1	0.3	8	STABLE

213	14	9.7	40	1.5	1	0.2	8	STABLE
214	13.9	10	32.2	2	1	0.3	8	STABLE
215	14	9.9	32.2	2.9	1	0.3	8	STABLE
216	23	19	15	20	1	0.2	8	STABLE
217	30	6.6	30	8.8	1	0.2	7.5	STABLE
218	30	6.6	30	3.5	1	0.2	7.5	STABLE
219	11	5.8	36	20	1	0.2	2	STABLE
220	17.3	11.3	85	8.8	1	0.2	7.5	STABLE
221	20.3	16.8	41	3.5	1	0.2	7.5	STABLE
222	20.8	7.6	48	15	1	0.2	2	STABLE
223	23	19	15	8	1	0.3	5.6	STABLE
224	14.4	26.9	15	1.6	1	0.3	5.6	STABLE
225	14.4	26.4	15	3	1	0.3	3.8	STABLE

6.3.1 Database analysis

Before proceeding with data pre-processing the investigation of data investigation is a reasonable to step to ensure data quality and to help in making informed decisions about preprocessing technique and set the foundation for building effective and accurate machine learning models.

To identify the distribution of the data points the histograms of each feature were plotted (Figure 6.3). Histograms provide a visual representation of the distribution of data across different values. This helps in understanding how the data is spread and whether it follows a normal distribution, skewed distribution, or has other patterns.

From the histograms we can notice that shape parameters are mostly concentrated around similar values. For the height it is especially prominent that most of the cases are clustered around 30 m, there are only a few instances where the height deviates significantly from this value. Similar tendency can be noticed in rock mechanics parameters, especially for Q' values, factor A and C. Addressing this issue might involve considering strategies such as standardization and/or rescaling the data to provide the model with a more varied and informative set of input features. Plotting histograms assisted us in understanding the distribution of these parameters, which is crucial for making informed decisions during the preprocessing stage and optimizing the model's performance.

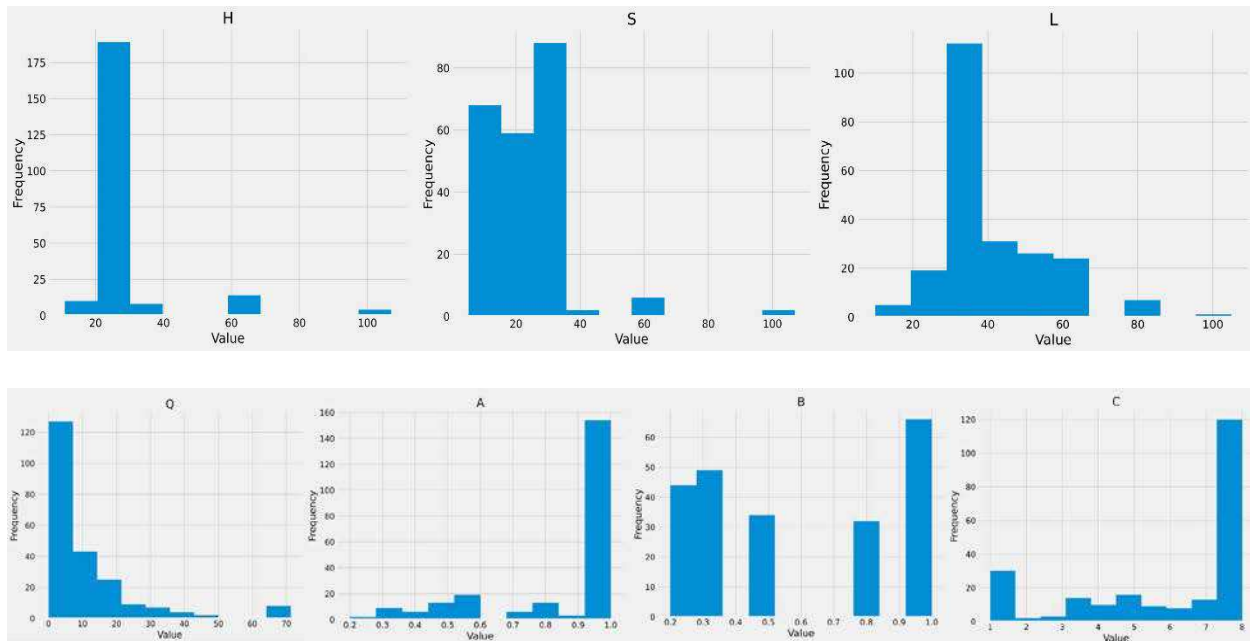


Figure 6.3 Histograms of the parameters

For further investigation the boxplots, also known as box-and-whisker plots were created (Figure 6.4). A boxplot is a graphical representation that displays the distribution and spread of a dataset. It provides a visual summary of the central tendency, spread, and potential outliers in the data. Boxplots offer a concise and informative determination of key statistical measures, making them a valuable tool in the exploratory data analysis phase before applying machine learning algorithms.

Box plots describe a sample by utilizing the 25th, 50th, and 75th percentiles—referred to as the lower quartile (Q1), median (m or Q2), and upper quartile (Q3)—along with the interquartile range ($IQR = Q3 - Q1$), encompassing the central 50% of the data. Quartiles demonstrate resilience to outliers and keep information about both the center and spread. As a result, they are favored over the mean and standard deviation for database with asymmetry or irregular shapes and for samples containing extreme outliers (Krzywinski and Altman 2014). The median is defined by a line separating the box, marking the midpoint of the dataset. This line indicates that 50% of the

data is surpassing the median. The top of the box plot represents the upper quartile (Q3), which means that 25% of the data exceeds this value, while the lower quartile (Q1) is represented at the bottom of the box, where 25% of the data is less than this value. The top "whisker" illustrates values higher than the median, and outliers are represented by dots above the top "whisker." A similar interpretation applies to the bottom "whisker" and outliers. Box plots can also show the skewness in the dataset, with the position of the median on the box indicating how much data falls above or below it.

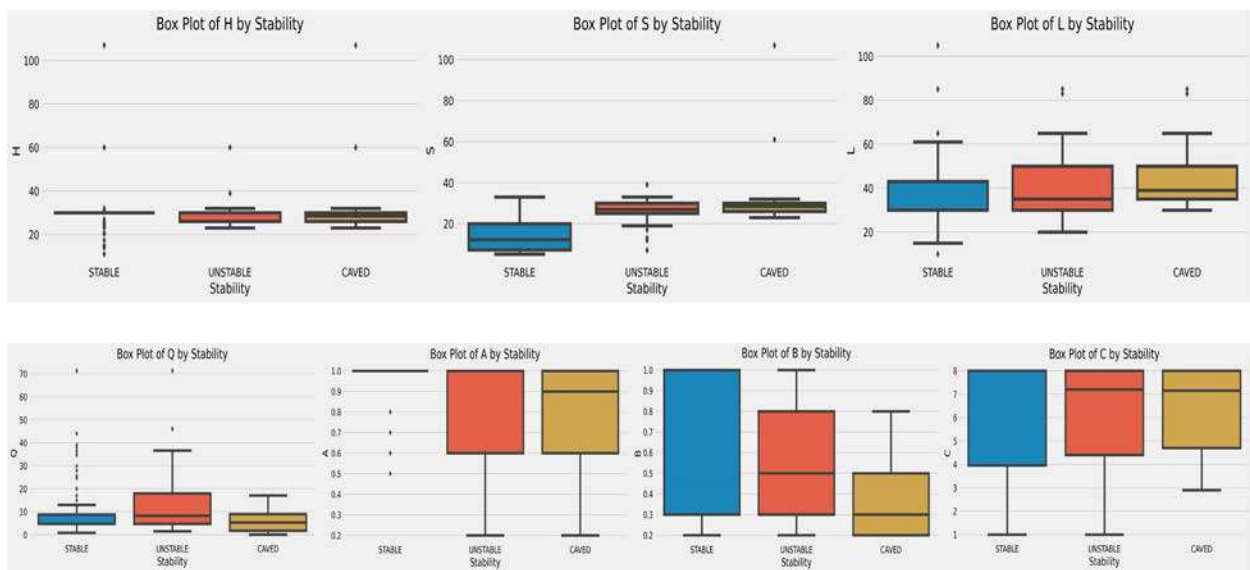


Figure 6.4. Box – plots of the parameters

Creating individual box plots for each stability assessment category—stable, unstable, and caved - offers an extensive visual representation of how the features are distributed within specific classes. This approach enables a detailed analysis of the tendencies and variations in each feature relative to the different stability conditions. Boxplots with extended lengths indicate a greater dispersion of data. Upon closer examination, it becomes evident that factors A, B, and C exhibit a more scattered distribution compared to the shape values and factor Q'. However, an exception is observed in the case of Factor A for stable cases, where a substantial number of instances cluster around similar values, contributing to a distinct pattern within this stability category. This observation shows that, for these specific factors, the data tends to vary widely, especially in

unstable and caved scenarios, with an exemption for stable cases of factor A, where a more concentrated distribution is apparent. A limited number of outliers are evident, particularly related to shape parameters, within all stability classes. These outliers show instances that noticeably deviate from the range of the central parts of the data. In the dataset, an intentional choice was made to keep these outliers, as they capture natural variations within the dataset. This decision acknowledges that these outliers contribute valuable information and reflect diversity in the data, rather than being treated as anomalies or errors.

The final step in analyzing the database involved the generation of a correlation heatmap (Figure 6.5), a graphical representation that shows the interconnections between all the features. The heatmap provides a comprehensive overview of how each feature relates to every other feature in the dataset. The color in the heatmap signifies the strength and direction of the correlation: red shades denote stronger positive correlations, while darker blue shades indicate weaker or negative correlations.

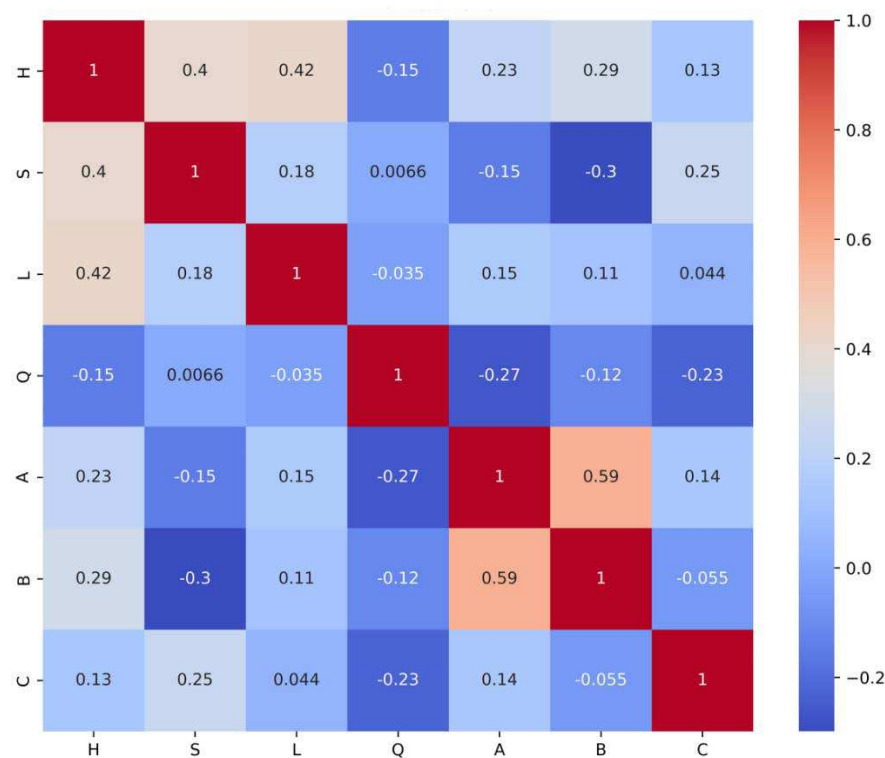


Figure 6.5. Correlation heatmap of all parameters

A higher correlation between features implies a statistical relationship where changes in one feature are associated with changes in another. In some cases, this can lead to redundant information, where one feature might provide similar insights as another. In machine learning, dealing with highly correlated features can be crucial because it may introduce bias into predictive models (Srivastava, 2023). However, in this specific dataset, I observed that the features do not exhibit a strong correlation. This lack of interconnection is a positive sign for the performance of machine learning models. When features are not highly correlated, it suggests that each feature contributes unique information and insights to the model. In such cases, the model is less likely to be influenced by unnecessary or overlapping information, which can lead to more accurate and unbiased predictions.

The presented analysis of the data allowed us to fully understand its structure, complexity, and potential. The fact that the database has three distinct classes: stable, unstable, and caved, implies that I had to specifically look for machine learning models that have ability to manage multiclass classification problems. Another factor to consider is the size of the database, which consists of 225 cases. Small databases often contain a limited range of examples, which may not fully represent the diversity of the real-world scenarios the model is meant to handle. This limited size was one of the reasons I decided not to remove the outliers in the data, as they can provide valuable information and show diversity. The lack of correlation between the features signifies that the model might be less prone to overfitting in the training data and could generalize better to new, unseen data. In order to effectively handle the complexities and challenges posed by the dataset, I made a decision to employ and compare several Machine Learning models. The objective was to assess the performance of each model and determine which one could achieve the highest accuracy and most correct predictions, while being suitable solution for my specific problem. The Machine Learning models investigated are: Random Forest, Support Vector Machine (SVM), AdaBoost, XGBoost, LightGBM, and Artificial Neural Network (ANN).

6.3.2 Database pre-processing

To prepare the data to be passed to machine learning models, I have undergone a series of processing steps. This data preparation process involves stages aimed at optimizing the dataset's format and structure to ensure compatibility with machine learning algorithms. The goal is to enhance the models' effectiveness by providing them with a well-organized and informative input, enabling them to learn and generalize patterns effectively during the training phase. Pre-processing involves employing various techniques and is widely recognized as a crucial task, establishing a significant portion, potentially up to 80%, of the overall model development effort (Fan et al. 2021). In practical terms, it is almost always beneficial to apply pre-processing techniques to input data before presenting it to a model (Nawi et al. 2013).

The initial pre-processing technique applied to the data involved feature scaling, also known as standardization. This was specifically necessary because the input variables in the dataset exhibit varying ranges, leading to each feature having a different scale. Such differences across the dataset can pose challenges in developing an accurate model. This need for standardization becomes evident after examining histograms and boxplots, showcasing the diverse value ranges of different features. For instance, Factors A and B share a range from 0 to 1, while Factor C spans from 2 to 8, the Q' value extends to 70, and shape parameters reach even up to 100. It's essential to note that scale relates to the variation in the value ranges, not the distribution shape. Standardization is essential for ensuring uniformity in scale across features, crucial in the proper development of the model.

For the database I decided to employ StandardScaler tool as a standardization method, which follows standard normal distribution. StandardScaler operates by expecting the data to be normally distributed within each feature and subsequently scales these values to center the distribution around 0, with a standard deviation of 1 (Raju et al. 2020). The process involves determining the mean and standard deviation for each variable. Subsequently, the scaled feature is calculated by adjusting each value based on these statistics. This standardized approach ensures that the features have consistent scales, facilitating a more uniform and effective analysis of the data during the modeling process.

The Machine Learning models under consideration required numerical values for all input variables. However, the feature labels in the database were categorical, classified as stable, unstable, and caved. To ensure optimal Machine Learning performance and obtain reliable outputs, a crucial step involved converting these categorical labels into numerical values. This transformation was executed as follows: stable was assigned the numerical value 1, unstable was assigned 2, and caved was assigned 0. This categorization enables the models to effectively process and interpret the labels during the learning process, ensuring compatibility with the numerical expectations of the machine learning algorithms.

Another important step in the stage of data preparation is splitting it into training, testing and validation sets. In machine learning it is a routine practice is to split the data before passing it to the predictive models. The primary reason for splitting the database is for evaluation of the performance of the model and it is especially important in supervised learning machine learning tasks (Hastie et al. 2017). The training set is used to train the machine learning algorithm, the primary purpose of it is to allow the model to learn the underlying patterns, relationships, and structures within the data so that it can make accurate predictions and classifications when presented with new, unseen data. The model adjusts its internal parameters during the training process to minimize the difference between its predicted outputs and the actual target values in the training set (Birba 2020). The validation set is an independent part of the dataset used to fine-tune the model during the training phase and make decisions about its architecture and hyperparameters. This allows researchers to experiment with different model configurations and select the combination that optimizes the model's performance. This is also crucial for avoiding overfitting to the training data and ensuring the chosen model is effective in unseen scenarios. The validation set provides an unbiased evaluation which helps to detect and reduce overfitting (Xu and Goodacre 2018). Finally, the testing set, which the model has not seen during training or validation, serves as an independent subset of the dataset to assess how well the model generalizes to new, unseen data. The testing set is used to calculate various performance metrics, such as accuracy, precision etc. depending on the nature of the machine learning task. These metrics provide informative measures of the model's effectiveness in making predictions (Jain et al. 2022).

For my specific problem, I decided to randomly split the dataset into those mentioned before three subsets, where 80% of the samples was allocated as training, 10% as validation and

10% as testing set. I applied the same splitting for all the evaluated machine learning models to maintain consistency and ensure fairness in the comparative analysis of the models.

All these pre-processing techniques allowed us to understand the database further, and properly prepare it to be passed to several machine learning models that I want to compare in this study. Figure 6.6 presents all the pre-processing steps for machine learning workflow.

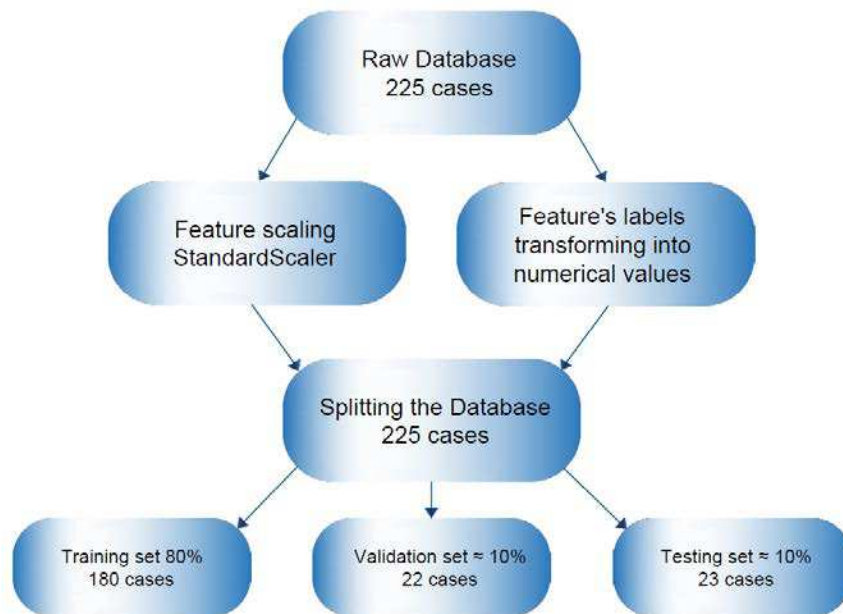


Figure 6.6 Pre-processing workflow for the database preparation

6.4 Comparative study overview

In the scope of my study, the objective was to explore and compare a selection of the most widely adopted machine learning models that have demonstrated effectiveness in addressing problems similar to ours. I specifically looked for models with established success in handling multi-class classification tasks, aligning with the nature of the problem where instances needed to be categorized into one of three classes (stable, unstable, caved). An important consideration was

the suitability of the chosen models for problems without significant correlation between the features. I wanted to ensure that they could effectively capture diverse relationships. The decision-making process also involved taking into account the number of instances in the dataset, which consists of 225 cases. Additionally, given the supervised nature of my problem, where each case is labeled with a known class, I focused on machine learning models designed for supervised learning.

Conducting a comparative study and examining different machine learning models offers several advantages and insights that contribute to the effectiveness of the overall analysis. Different models have unique strengths and weaknesses. A comparative study allows for an evaluation of their performances, helping to identify which models are more suitable and effective for the given problem. Some problems may be inherently complex, and certain models may handle it better than others. Comparative studies assist in determining which models are better equipped to deal with such relationships and patterns within the data. Different models may also emphasize different features in making predictions. Comparing different ML algorithms can shed light on which features are considered most important across various models, providing valuable insights into the factors influencing the predictions. Relying on a single model may create risks, especially if that model is not well-suited to the specific characteristics of the data. A comparative study addresses this risk by exploring a range of models and offering a more comprehensive understanding of their performance.

6.4.1 Investigated Machine Learning models

To conduct this comparative study the following machine learning models were investigated: Random Forest, Support Vector Machine (SVM), AdaBoost, XGBoost, LightGBM and Artificial Neural Network (ANN). The inclusion of these diverse models ensures a thorough examination, offering insights into their respective strengths and weaknesses in the context of my study.

6.4.1.1 Random Forest

Random Forest is a powerful and versatile machine learning model known for its satisfactory performance across a wide range of tasks. It belongs to the ensemble learning family, where multiple decision trees are combined to form a more flexible and accurate model. The strength of Random Forest lies in its ability to reduce overfitting, handle complex relationships in data, and provide insightful feature importance rankings. Breiman's innovative work on Random Forests laid the foundation for this machine learning technique, and subsequent research and practical applications have endorsed its significance in the machine learning field (Breiman 2001).

By combining the predictions from number of individual trees, each trained on a random part of the data, the model achieves a higher degree of accuracy and generalizes well. Random Forest has demonstrated excellence in both classification and regression problems, making it a popular choice in many research fields, including mining engineering. Several researchers have successfully applied RF in their study focused around area of stoping mining methods and stability predictions. Qi et al. conducted a study to predict the stability of hanging wall with random forest (2018a), Szmigiel and Apel investigated the feasibility of employing random forest to assess the stability of open stopes (2022), and Jorquera et al. applied random forest to predict the dilution in sublevel stoping mining method (2023). Random forest have been also successfully applied in related mining engineering disciplines, such as predicting mining induced stresses in underground openings (Vinay et al. 2023), rock pillar stability (Zhou et al. 2015) or ground settlement predictions (Zhou et al. 2017).

These successful applications have built my confidence in including random forest model in my study, to observe its performance on the specific data compared with other models, as well as determine the effects of particular features on its predictions.

6.4.1.2 Support Vector Machine

The Support Vector Machine (SVM) is a powerful machine learning model known for its efficacy in classification tasks, particularly excelling in scenarios with high-dimensional data. Developed by Cortes and Vapnik, SVM constructs an optimal hyperplane that efficiently separates

different classes in the feature space, maximizing the margin between them (Cortes and Vapnik 1995).

SVM's has an ability to handle complex decision boundaries, making it well-suited for non-linear classification tasks. It is possible due to the use of kernel functions that transform input data into higher-dimensional spaces. This is especially advantageous attribute, that should effectively handle the database which exhibits a low correlation index between the features. Importantly, SVM proved to be effective in multi-class classification problems by employing strategies like one-vs-one or one-vs-all, demonstrating adaptability and satisfying performance across diverse applications (Evgeniou and Pontil 2001). Some applications of SVM model in the area of mining engineering focus on open stopes, Jorquera et al. study about predicting dilution in sublevel stoping evaluates SVM along with other models (Jorquera et al. 2023), similarly few comparative open stope stability studies conducted (Erdogan Erten et al. 2021), (Qi et al. 2018b). However most of the research with SVM models focuses around different mining engineering challenges, such as predicting the stability of hard coal pillars (Li et al. 2023), slope stability analysis (Samui 2008), modeling displacement time series of geomaterials (Feng et al. 2004), or predictions of mining subsidence (Li et al. 2011). All these mentioned advantages of SVM model, as well as not that many applications in the area of stope stability, convinced us in evaluating this particular machine learning technique along others.

6.4.1.3 Adaptive Boosting - AdaBoost

Adaptive Boosting, is an ensemble learning algorithm introduced by Yoav Freund and Robert Schapire (1996). This effective method operates by combining the predictions of weak learners, typically decision trees, to form a powerful and accurate classifier. AdaBoost assigns higher weights to misclassified examples, forcing following weak learners to focus on these challenging cases. The final prediction is then determined through a weighted majority vote of the individual weak learners. AdaBoost's adaptability and capability to handle complex relationships in data, and its emphasis on learning from misclassifications and continuously elevating its prediction, make it particularly suitable for multiclass classification problems and uncorrelated features (Wang 2012). Numerous application of AdaBoost algorithm can be found in the area of

rockbursts predictions in mining engineering (Ahmad et al. 2022), (Wang et al. 2023), (Li et al. 2022), but not many applications of this particular model can be found in the area of stope stability assessment, except Saadaari et al. (2020). Given that this particular model isn't popular in mining stoping method applications, although it has demonstrated efficiency in addressing similar challenges in engineering, I made a decision to incorporate and assess its performance in my study.

6.4.1.4 Extreme Gradient Boosting - XGBoost

Extreme Gradient Boosting, is another powerful and efficient machine learning algorithm introduced by Chen and Guestrin (2016). In comparison to AdaBoost, XGBoost stands out for its scalability, fast processing, and ability to handle extensive datasets. While both are an ensemble learning methods that combine weak learners to create a successful model, XGBoost employs a more advanced optimization approach, incorporating regularization terms and parallel processing to improve performance. XGBoost is well-suited for multiclass classification, it implements a one-vs-all strategy, creating individual classifiers for each class and then combining their outputs. The algorithm's ability to capture complex relationships between the features and effectively manage them, makes it an excellent candidate to include in my study. Some mining engineering applications of XGBoost can be found in literature, including subsidence predictions (Gu et al. 2022), rock fragmentation and ground vibration predictions in blasting operations (Chandrabhas et al. 2022) or hard rock pillar and underground entry-type excavations stability predictions (Liang et al. 2020), (Zhou et al. 2023).

6.4.1.5 LightGBM

LightGBM, model developed by Microsoft, is a gradient boosting framework designed for distributed and efficient training on large datasets. Compared to other gradient boosting algorithms like AdaBoost and XGBoost, LightGBM employs a histogram-based approach, accumulating data points to accelerate the training process. Its leaf-wise growth strategy focuses on minimizing loss, resulting in a more accurate and adaptive model. LightGBM stands out for its ability to handle categorical features efficiently and it is suitable for multiclass classification

problems, employing, similarly to other models, a one-vs-all strategy (Ke et al. 2017). The area of mining engineering research doesn't show many LightGBM model applications, except study on predicting hard rock pillar stability (Liang et al. 2020) and mineral grade estimation (Kaplan et al. 2021).

Given its ability to handle complex relationships in data, LightGBM is a compelling choice for my study, where predicting the stability of open stopes involves dealing with diverse and complex patterns within the dataset.

6.4.1.6 Artificial Neural Network

The last model that I have decided to employ in my study is a popular widely used artificial neural network. This effective machine learning model is inspired by the structure and functioning of the human brain and it's built of interconnected nodes organized into layers. Unlike ensemble methods such as presented before Radom Forest, AdaBoost, XGBoost, and LightGBM, ANN operates on a fundamentally distinct architecture, leveraging connected neurons to capture complex, non-linear patterns within the input data. While boosting models excel in combining weak learners to form a strong classifier, ANN stands out for its capacity to automatically learn hierarchical representations of features. Unlike decision tree-based models, ANN does not rely on predetermined splits but dynamically adjust weights during training, making them well-suited for capturing challenging relationships. Including ANN in my comparative study offers a comprehensive exploration of diverse modeling approaches, ensuring a thorough understanding of how different models handle the challenge of predicting the stability of open stopes. There is numerous research focusing on ANN applications in mining engineering, with stability of stopes included. These include Wang's (2002) application of an ANN model for designing underground excavation spans and Adoko's (2022) feasibility study on implementing ANN classifiers for open stope design.

6.5 Evaluation and results

Evaluating machine learning models is essential for ensuring their effectiveness and reliability in real-world applications. By assessing performance metrics such as accuracy, precision, recall, and F1 score, we can determine how well a model performs on a given task and identify areas for improvement. This process helps in selecting the most suitable model for the problem and supports decision-making processes.

6.5.1 Training and evaluation

The first model subjected to training and testing on the dataset was the Random Forest (RF). RF offers an adaptable framework by allowing the adjustment of hyperparameters to optimize model performance and address overfitting concerns. Two crucial hyperparameters in RF are the number of estimators (decision trees), and the maximum depth of each individual tree. Calculating accuracy for validation and training set allowed us to notice and address any overfitting problems displayed by the RF model by adjusting those parameters. When the accuracy for training set is much higher compared to validation, the model is learning too well from the training set and struggles to properly generalize for new unseen data. When the number of decision trees and the depth of each tree were set to higher values, the accuracy for training set was reaching close to 100%, while the validation accuracy was dropping significantly. Thus, the optimal number of decision trees was set to 100, with the depth of each tree equal to 5. This hyperparameter setting achieved the accuracy of training and validation set equal to 88% and 90% respectively.

Second model that was trained was Support Vector Machine. The SVM model has two parameters, C and Gamma that serve as adjustable hyperparameters. The C parameter, often referred to as the regularization parameter, impacts the trade-off between achieving a smoother decision boundary and correctly classifying training points. A smaller C value results in a softer margin, allowing for more points to be classified correctly but potentially leading to overfitting, while a larger C value enforces a stricter margin, prioritizing a simpler separation between classes but potentially sacrificing the correct classification of some training points. The gamma parameter is associated with the radial basis function (RBF) kernel, a common choice in SVMs models for handling non-linear relationships in data. A smaller gamma value results in a wider Gaussian kernel, leading to smoother decision boundaries and potentially underfitting the training data. In

contrast, a larger gamma value narrows the Gaussian kernel, allowing the model to capture more complex patterns in the data but possibly leading to overfitting (Liu et al. 2006). Grid search approach was performed to find the optimal combination of C and gamma values for a SVM model. Grid search is an optimization technique that systematically searches through a predefined set of hyperparameter values for a model. It involves evaluating the effectiveness for each combination of C and gamma values using a cross-validation, in order to find the ones that result in the best performance without overfitting.

The same grid search method was uniformly employed across all ensemble learning algorithms—AdaBoost, XGBoost, and LightGBM—during the training phase to tune their hyperparameters. Specifically, the parameters under investigation were the number of estimators (representing weak learners or trees) and the learning rate, which controls the contribution of each weak learner to the final combined model. Cross-validation was performed to find the most efficient values of those parameters, ensuring the most accurate results while avoiding overfitting.

In the final step of the modeling process, an Artificial Neural Network (ANN) was built and tailored to the dataset. The exploration of various ANN model structures was performed, with an evaluation on the validation set. The finalized architecture consists of four layers: an input layer, two hidden layers, and an output layer. The input layer's configuration aligns directly with the number of features in the dataset. Subsequently, the first and second hidden layers consist of 32 and 16 neurons, respectively and the output layer's structure is determined by the number of classes within the dataset. Activation functions played a crucial role in shaping the model, with the first three layers employing the swish activation function, known for its non-linearity and smoothness. The output layer, responsible for producing the final predictions, employed the softmax function to ensure appropriate class probabilities. The optimization of weights during each epoch and the minimization of the algorithm's loss were facilitated by the Adam optimizer. Categorical cross-entropy was adopted as the loss function, aligning with the multi-class nature of the problem, contributing to the model's ability to determine and classify instances accurately.

6.5.2 Results

All the investigated Machine Learning models were assessed with the same evaluation metrics, to demonstrate a clear contrast between their results and conduct a proper comparative analysis. The most popular and basic metric is the accuracy score, which can be calculated for both validation and testing sets (Figure 6.7). Upon the examination of the model performance, it becomes clear that the Artificial Neural Network (ANN) stands out as the leading performer, achieving a great accuracy score of 91% on the testing set. Following closely is the Random Forest model, demonstrating an accuracy score of 86%. Notably, both the Support Vector Machine and LightGBM classifiers showed identical accuracy scores for both the validation and testing sets. However, it is crucial to highlight that, at this point of evaluation, the Adaptive Boosting classifier appeared as the least effective, displaying the lowest accuracy among all the models assessed.

In the next step of my analysis the averaged classification report metrics such as precision, recall and F-1 score were determined and compared (Figure 6.8). This provides aggregated performance evaluation across all classes, presents summary of how well the model performs on average. These metrics provide a single numerical value that represents the overall performance, simplifying the interpretation of results. Precision measures the accuracy of positive predictions. It calculates the ratio of true positive predictions to the total predicted positives, high precision values indicate fewer false positives predictions. Recall measures calculates the ratio of true positive predictions to the total actual positives, higher recall values indicate fewer false negatives predictions. The F1 is a harmonic mean of precision and recall that combines them into a single value, higher F1 score is generally desirable because it indicates a better trade-off between these two metrics. Similarly as with accuracy, the ANN model achieved the highest score for all three metrics, while RF's results declined for recall and F-1 score. Among all the models, AdaBoost is again showing the lowest results, therefore it exhibits poor predicting capabilities for slope stability data.

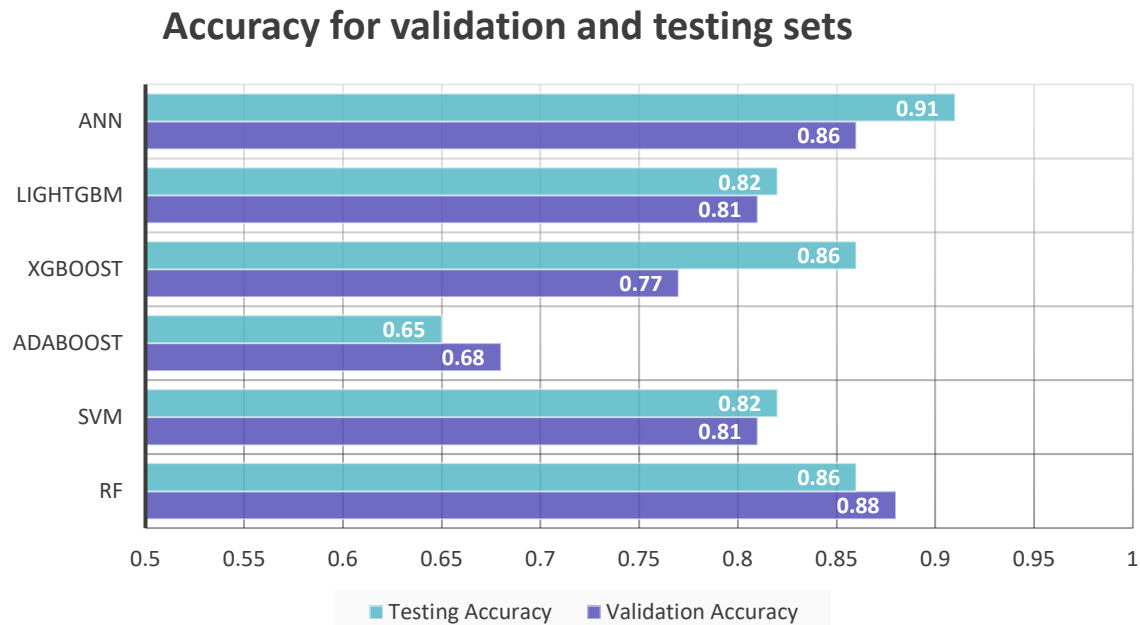


Figure 6.7 Accuracy score for all the Machine Learning models

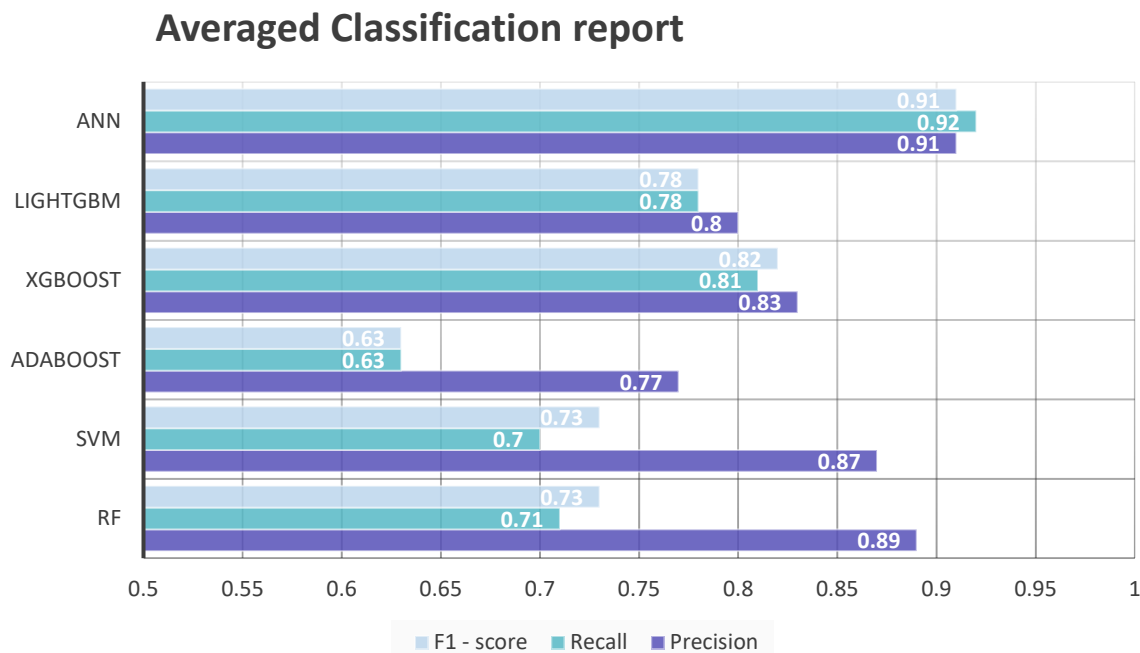


Figure 6.8 Classification report for all the Machine Learning models

In the context of slope stability predictions, precision appears as a crucial metric, especially in the case of stable class, as it directly addresses the accuracy of positive predictions, holding particular significance when the cost associated with false positives is high. These false positive predictions could potentially lead to significant safety concerns. When an unstable or caved condition is incorrectly classified as stable, it creates a major risk to the safety and integrity of the mining environment. Precision in this context becomes the crucial metric because it emphasizes the correctness of the model in identifying actual conditions. The precision scores across all classes are shown on Figure 6.9. The recall score on the other hand (figure 6.10), has a significant importance in the case of caved predictions. When the recall score for caved class declines, it means that there are instances that are incorrectly classified as stable or unstable. This mispredictions are also dangerous as they are causing some potentially dangerous conditions to be presented as either unstable or stable.

When analyzing the precision and recall graphs, a crucial realization was evident, emphasizing the limitations of the Random Forest model despite its apparent efficiency based on overall accuracy. While accuracy is a comprehensive metric, a more careful examination exposed a critical concern related to the recall score for the caved class. The recall value of 0.33, indicated a notable deficiency in the RF model's ability to accurately identify instances of caved conditions. This score showed that a substantial portion of the actual caved cases were being misclassified as either stable or unstable. Such misclassifications create a significant risk to mine safety, as they imply that potentially hazardous conditions associated with caved areas are not being properly identified.

The Artificial Neural Network once again demonstrated its exceptional capability in producing the best classification results across all classes, standing out in comparison to all other models evaluated in this study. The ANN achieved a perfect recall score of 1 for both stable and unstable cases, indicating that none of these cases were incorrectly assigned to a different class. However, a challenge appears in the classification of unstable cases, which represent conditions that are between stable and caved, posing potential risks in a mining environment. The precision score of 0.82 for caved class and 0.91 for stable class, means that some of those unstable cases were assigned to one of those two classes. However, the lower precision for the caved class indicates that a larger proportion of these unstable cases was classified as caved. This aligns with

a safer approach, as misclassifying unstable conditions as caved is a more reasonable and careful choice in terms of mine safety.

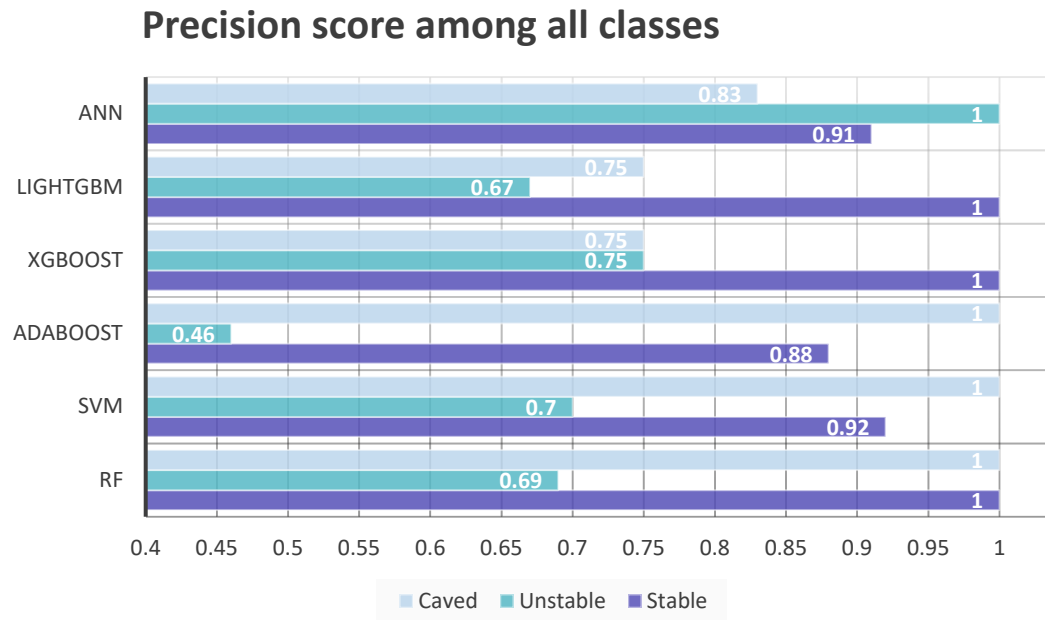


Figure 6.9 Precision scores for all Machine Learning models

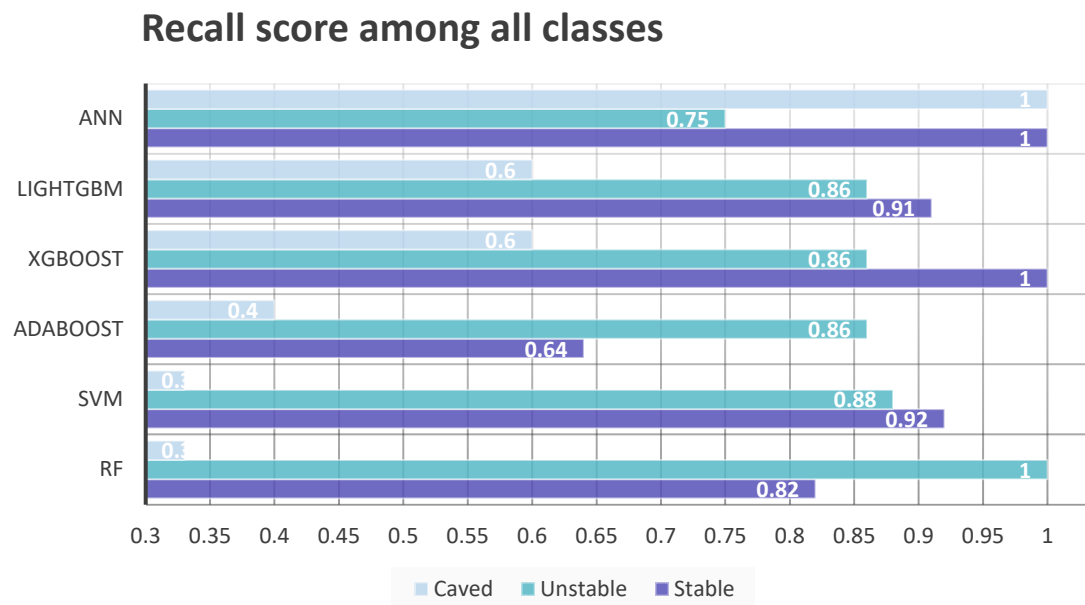


Figure 6.10 Recall scores for all Machine Learning models

6.5.3 Feature importance analysis with SHAP

In the context of my study on open stope stability, conducting feature importance analysis is important for a comprehensive understanding of the factors influencing stability outcomes. Feature importance analysis allows us to determine the relative contribution of each parameter to the predictive performance of the machine learning models. By identifying the most influential features, we gain insights into the key determinants of stope stability, enabling engineers to prioritize and focus their attention on critical factors during the design and planning stages. This knowledge not only enhances the interpretability of the model but also assists in making informed decision-making in real-world mining scenarios. Additionally, feature importance analysis helps in the optimization of data collection efforts, guiding engineers to gather more detailed information on crucial parameters. Ultimately, this process regulates the machine learning model with domain-specific knowledge, ensuring that the predictions are not only accurate but also reflect the nuances of open stope stability and safety assessments.

To perform this feature analysis a SHapley Additive exPlanations (SHAP) analysis was performed for all machine learning models investigated in this study. SHAP is a framework for explaining the output of machine learning models outcomes. It assigns a value (Shapley value) to each feature, indicating its contribution to the model's prediction (Lundberg and Lee 2017). In the context of my study on open stope stability assessment, performing SHAP analysis is crucial for several reasons. It provides interpretability to complex machine learning models, helping us understand how each input feature influences the stability predictions. This interpretability is essential for gaining insights into the driving factors behind stability conditions. The information from SHAP analysis is invaluable for improving stope design strategies and focusing engineering efforts on the aspects that have the greatest influence on stability. Moreover, SHAP analysis aids in the validation and verification of the model's predictions. By understanding the rationale behind each prediction, engineers can assess the model's reliability and identify potential areas of improvement in both the model and the underlying data. SHAP analysis enhances the transparency, interpretability, and reliability of the machine learning model, making it an essential step in ensuring the practical applicability of the model's predictions in real-world mining scenarios.

SHAP value summary plots for RF, SVM, AdaBoost, XGBoost, LightGBM and ANN are shown on Figures from 6.11 to 6.16. The chart's horizontal axis represents SHAP values, while the vertical

axis displays all the features in the data, with the most influential one being on top. Each point on the chart corresponds to a SHAP value for a prediction and a feature. Red indicates a higher feature value, while blue indicates a lower one. By observing the distribution of the red and blue dots, we can gain a general understanding of the impact of feature directionality. The positive SHAP values are indicating more positive impact on the model, in this case those positive values mean that there is higher probability of predicting that case into stable or unstable class.

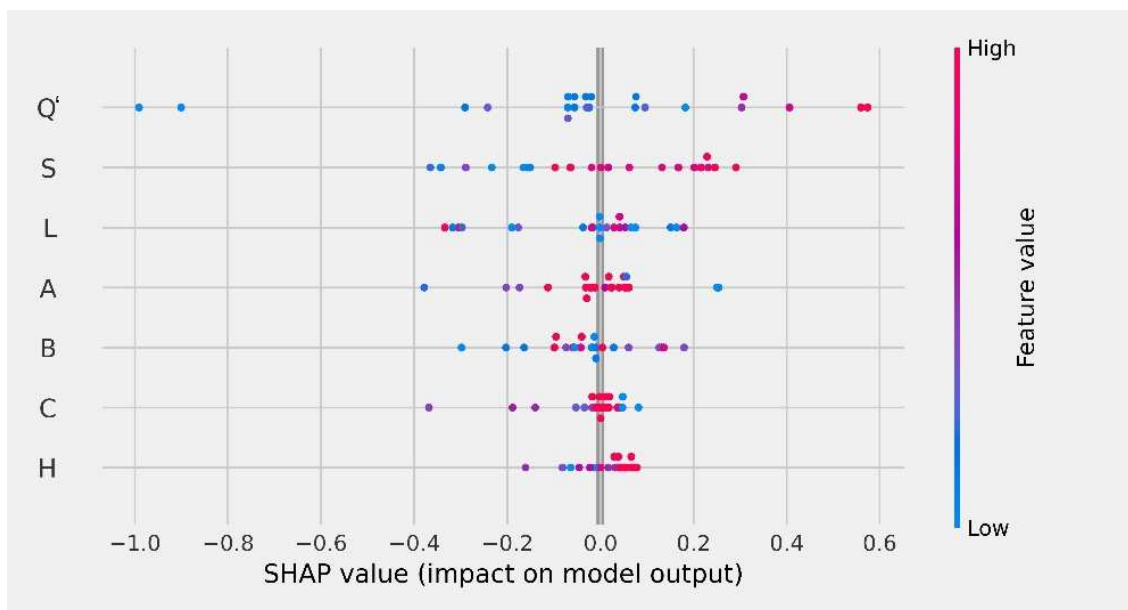


Figure 6.11 Random Forest - SHAP value summary plot

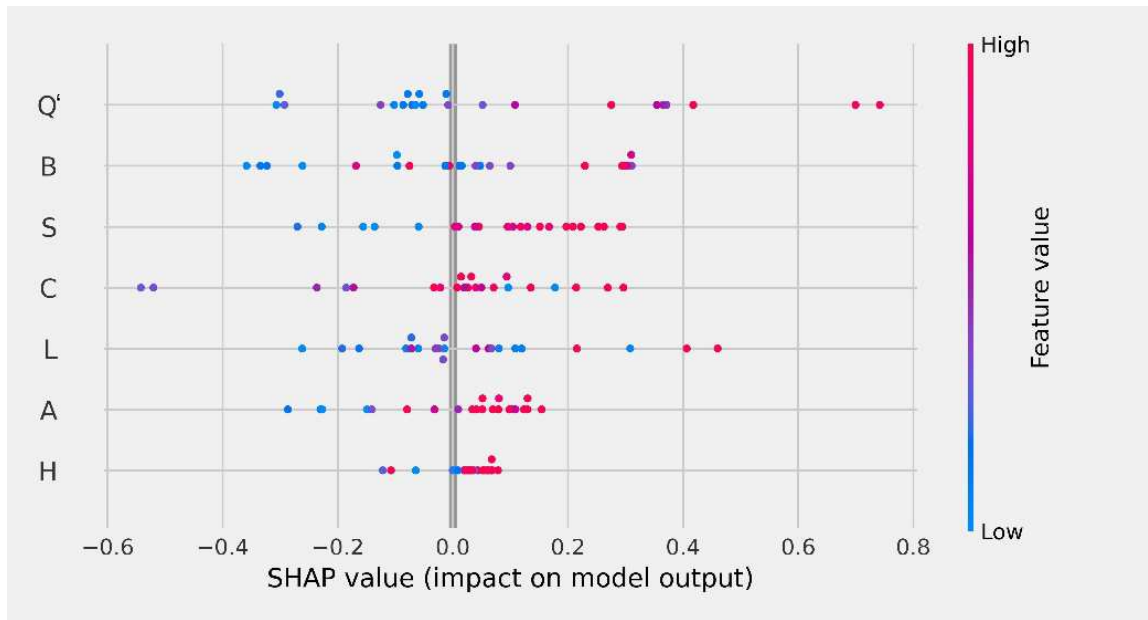


Figure 6.12. SVM - SHAP value summary plot

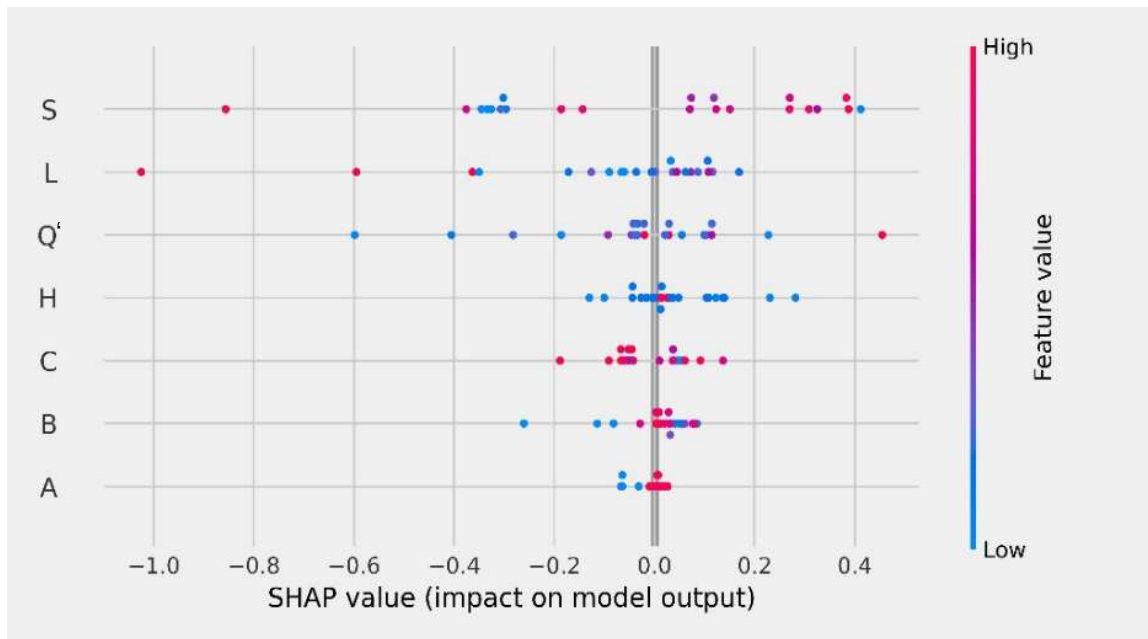


Figure 6.13 AdaBoost- SHAP value summary plot

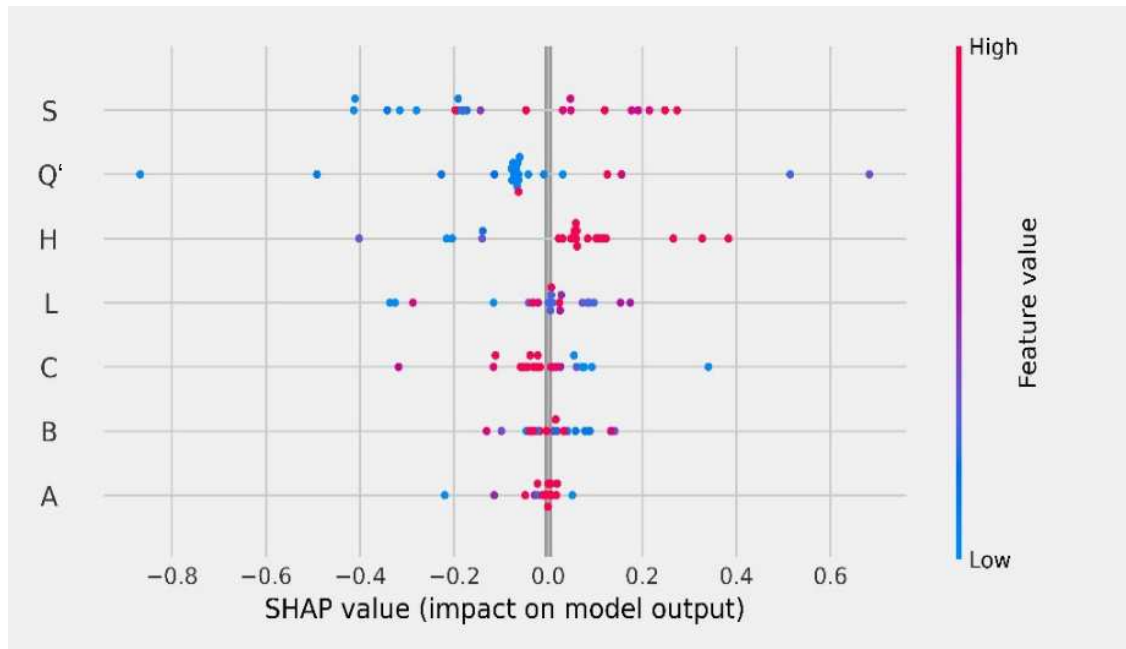


Figure 6.14 XGBoost- SHAP value summary plot

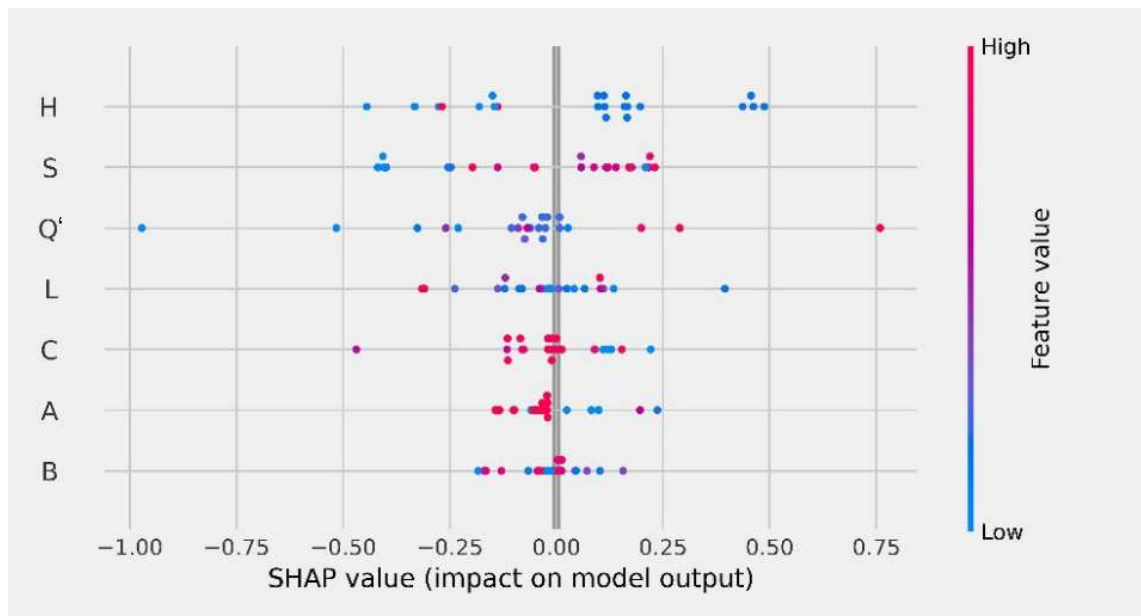


Figure 6.15 LightGBM- SHAP value summary plot

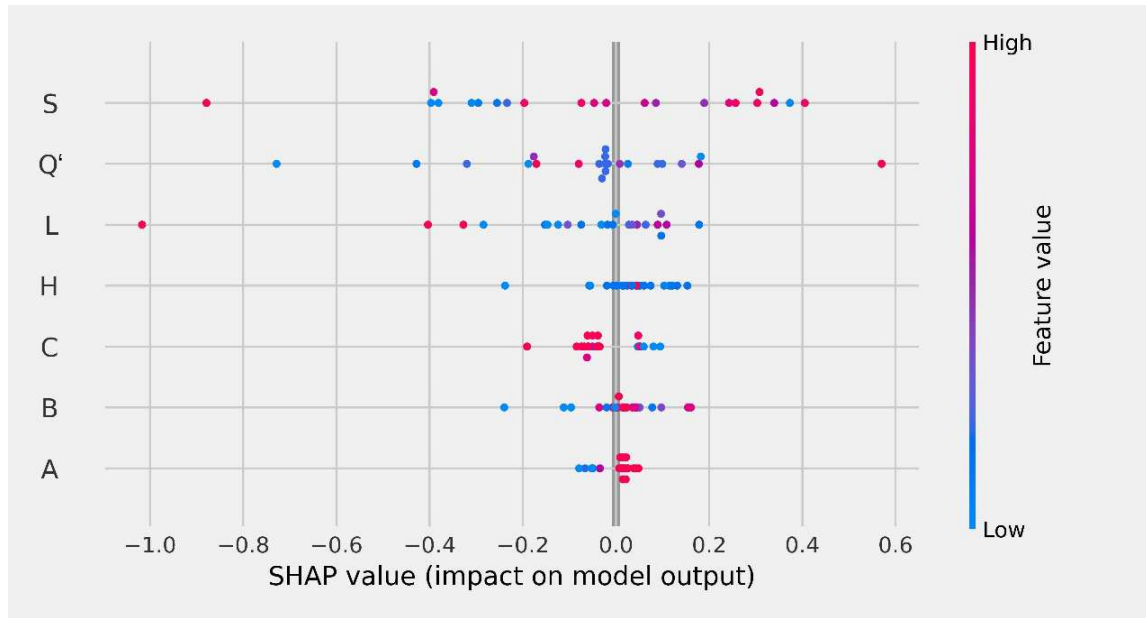


Figure 6.16 ANN model - SVM - SHAP value summary plot

In the case of a RF model, the Q' value has the most influence on the final prediction, followed by a shape value S which is the span of the opening and then L (length). The higher value of the Q factor, the higher probability that the slope will have a stable condition, similar trend can be noticed with the S factor. In case of the SVM model, it can be noticed that Q value is also the most influential, with the higher values having more positive impact on the stability of the slope. Surprisingly for the SVM model, the B factor has a second highest impact, whereas in other models its influence is less significant. For other models, the shape parameters, especially span, have the most significant impact. These results are good for a mine designing stage, as it is usually more feasible to adjust the size of the opening (such as a slope length or span) than to change the overall surrounding rock conditions. Engineers have greater control over the dimensions of the opening and can adapt them to suit specific requirements, such as equipment accommodation and stability considerations. Therefore, mine design prioritizes optimizing the size and shape of openings while working within the limitations of the surrounding rock conditions. However models such as AdaBoost, XGBoost and LightGBM, that have prioritized the shape of the opening in their predictions, are also models that exhibited less accurate overall performances. These models also

have a lack of consistency when it comes to the value of the feature impacting the predictions, with some having higher values influencing the model positively and others doing the opposite.

The Artificial Neural Network (ANN) model exhibited the best performance in terms of predictive capabilities, consistently delivering high classification scores across all stability classes. Upon closer examination of the feature importance analysis, it became apparent that the span factor (S) exhibited a high degree of influence on the model's predictions. Despite being the most influential factor, there was a notable lack of consistency in the impact of higher span values on stability predictions. While higher span values generally appeared to have a positive influence on stability predictions, there were instances where this trend was not maintained. In fact, there was one case where a high span value resulted in a significantly negative impact on stability. This inconsistency in the influence of span highlights the complexity of the relationship between this factor and stope stability.

Since the variability observed in the influence of span on stability predictions is not consistent, it is advisable to prioritize the second most influential factor, which is the rock quality factor (Q'). The analysis revealed a clearer trend with Q', where lower values consistently exhibited either highly negative or neutral influences on stability predictions. This reliability in the relationship between Q' and stability suggests that it may be a more dependable factor to consider when making predictions. Additionally, the length factor (L) was observed to have a more consistently negative impact on stability predictions for larger values. Although not as influential as factors S or Q', the consistent trend with length reinforces its importance in stope stability assessment.

6.6 Summary

In this study, I explored the application of various machine learning models to predict the stability of open stopes in underground mining operations. The investigation aimed to compare the performance of different models and provide insights into their effectiveness in addressing the complex problem of stope stability assessment.

Upon evaluating the performance of the investigated machine learning models, several key findings emerged. The Artificial Neural Network (ANN) demonstrated exceptional predictive capabilities, outperforming all other models with an accuracy score of 91% on the testing set. This result underscores the effectiveness of ANN in accurately categorizing instances into stable, unstable, or caved classes. Following closely behind, the Random Forest model exhibited promising performance, achieving an accuracy score of 86%. Despite its lower accuracy compared to ANN, Random Forest demonstrated competitive results, highlighting its potential as a reliable predictive tool for slope stability assessment. In contrast, the Support Vector Machine (SVM), LightGBM, and AdaBoost classifiers displayed comparable accuracy scores, although lower than ANN and Random Forest. Notably, the Adaptive Boosting classifier exhibited the least effective performance, demonstrating the lowest accuracy among all models assessed.

Precision and recall metrics provided further insights into the performance of the models in the context of slope stability predictions. Precision, which measures the accuracy of positive predictions, emerged as a crucial metric, especially for the stable class. The ANN model demonstrated high precision scores across all classes, indicating its capability to accurately identify stable conditions with minimal false positives. The recall score for the caved class emerged as a significant concern, particularly for the Random Forest model. Despite its high overall accuracy, Random Forest exhibited a notable deficiency in accurately identifying instances of caved conditions. This limitation poses significant safety risks in mining environments.

The SHapley Additive exPlanations (SHAP) analysis provided valuable insights into the factors influencing the stability predictions of each machine learning model. For instance, the Q' value emerged as the most influential factor in the predictions of the Random Forest and SVM models, emphasizing the importance of rock mass quality in determining slope stability. Interestingly, models such as AdaBoost, XGBoost, and LightGBM prioritized the shape parameters, particularly span, in their predictions. While this prioritization aligns with mine design principles, these models exhibited less accurate overall performances. The ANN model exhibited factor S as the most influential on the model predictions. However, in that case, there is also lack of consistency for the higher values, which seem to have rather positive influence on the stability except few cases. In the case of ANN it is better to look at the second most influential factor which is Q' , where it is more clear that lower values have either highly negative or neutral influence on

the stability predictions. It is also worth noticing that the factor L, has clearly more negative impact for the larger values. Due to the inconsistency of factor S, the analysis suggests relying on the second most influential factor Q', for more accurate stability predictions.

Despite the promising results obtained in this study, several paths for future research and improvement exist. Further exploration of the dataset and refinement of the feature selection process could enhance the predictive capabilities of the models. Additionally, investigating ensemble methods and hybrid approaches combining machine learning with classical stability assessment methods may yield even more accurate predictive models.

In conclusion, this study explores the potential of machine learning techniques in addressing the complex problem of stope stability assessment. By leveraging advanced algorithms and extensive data analysis, mining engineers can make informed decisions to enhance safety and optimize the design of underground excavations.

6.7 References

- Adoko AC, Saadaari F, Mireku-Gyimah D, Imashev A (2022) A Feasibility Study on The Implementation of Neural Network Classifiers for Open Stope Design. *Geotech Geol Eng* 40(2):677–696. <https://doi.org/10.1007/s10706-021-01915-8>
- Ahmad M, Katman HY, Al-Mansob RA, Ahmad F, Safdar M, Alguno AC (2022) Prediction of Rockburst Intensity Grade in Deep Underground Excavation Using Adaptive Boosting Classifier. *Complexity* 2022:1–10. <https://doi.org/10.1155/2022/6156210>
- Barton N, Lien R, Lunde J (1974) Engineering classification of rock masses for the design of tunnel support. *Rock Mechanics* 6(4):189–236. <https://doi.org/10.1007/BF01239496>
- Bieniawski ZT (1973) Engineering Classification of Jointed Rock Masses. *Transaction of the South African Institution of Civil Engineers* (15):335–344
- Birba DE (2020) A Comparative study of data splitting algorithms for machine learning model selection, *Computer and Information Sciences*. KTH, School of Electrical Engineering and Computer Science (EECS)
- Breiman L (2001) Random Forests. *Machine Learning* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
- Chandrasah NS, Choudhary BS, Teja MV, Venkataramayya MS, Prasad NSRK (2022) XG Boost Algorithm to Simultaneous Prediction of Rock Fragmentation and Induced Ground Vibration Using Unique Blast Data. *Applied Sciences* 12(10):5269. <https://doi.org/10.3390/app12105269>
- Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. <https://doi.org/10.48550/ARXIV.1603.02754>
- Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297. <https://doi.org/10.1007/BF00994018>
- Crone SF, Lessmann S, Stahlbock R (2006) The impact of preprocessing on data mining: An evaluation of classifier sensitivity in direct marketing. *European Journal of Operational Research* 173(3):781–800. <https://doi.org/10.1016/j.ejor.2005.07.023>
- Erdogan Erten G, Bozkurt KS, Yavuz M (2021) Grid Search Optimised Artificial Neural Network for Open Stope Stability Prediction. *International Journal of Mining, Reclamation and Environment* 35(8):600–617. <https://doi.org/10.1080/17480930.2021.1899404>
- Evgeniou T, Pontil M (2001) Support Vector Machines: Theory and Applications. In: Paliouras G, Karkaletsis V, Spyropoulos CD (eds) *Machine Learning and Its Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 249–257

- 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. *Front Energy Res* 9:652801. <https://doi.org/10.3389/fenrg.2021.652801>
- Feng X-T, Zhao H, Li S (2004) Modeling non-linear displacement time series of geo-materials using evolutionary support vector machines. *International Journal of Rock Mechanics and Mining Sciences* 41(7):1087–1107. <https://doi.org/10.1016/j.ijrmms.2004.04.003>
- Freund Y, Schapire RE (1996) Experiments with a New Boosting Algorithm. *International Conference on Machine Learning*
- Gu Z, Cao M, Wang C, Yu N, Qing H (2022) Research on Mining Maximum Subsidence Prediction Based on Genetic Algorithm Combined with XGBoost Model. *Sustainability* 14(16):10421. <https://doi.org/10.3390/su141610421>
- Hadjigeorgiou J, Leclair J, Potvin Y (1995) An update of the stability graph method for open stope design. *CIM Rock Mechanics and Strata Control session*, Halifax, Nova Scotia :14–18
- Hastie T, Tibshirani R, Friedman JH (2017) *The elements of statistical learning: data mining, inference, and prediction*, Second edition, corrected at 12th printing 2017. Springer, New York, NY
- Henning JG, Mitri HS (2007) Numerical modelling of ore dilution in blasthole stoping. *International Journal of Rock Mechanics and Mining Sciences* 44(5):692–703. <https://doi.org/10.1016/j.ijrmms.2006.11.002>
- Hu H, Cao Y (2009) Numerical Simulation Modeling and Calculation Analysis on Stope Roof Stability under the Complex Geological Conditions in Deep Mining. In: 2009 International Conference on Engineering Computation. IEEE, Hong Kong, China, pp 175–177
- Jain E, Neeraja J, Banerjee B, Ghosh P (2022) A Diagnostic Approach to Assess the Quality of Data Splitting in Machine Learning. <https://doi.org/10.48550/ARXIV.2206.11721>
- Jorquera M, Korzeniowski W, Skrzypkowski K (2023) Prediction of dilution in sublevel stoping through machine learning algorithms. *IOP Conf Ser: Earth Environ Sci* 1189(1):012008. <https://doi.org/10.1088/1755-1315/1189/1/012008>
- Kaplan UE, Dagasan Y, Topal E (2021) Mineral grade estimation using gradient boosting regression trees. *International Journal of Mining, Reclamation and Environment* 35(10):728–742. <https://doi.org/10.1080/17480930.2021.1949863>
- Ke G, Meng Q, Finely T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y (2017) LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: *Advances in Neural Information Processing Systems* 30 (NIP 2017)
- Krzywinski M, Altman N (2014) Visualizing samples with box plots. *Nat Methods* 11(2):119–120. <https://doi.org/10.1038/nmeth.2813>

- Li C, Zhou J, Du K, Dias D (2023) Stability prediction of hard rock pillar using support vector machine optimized by three metaheuristic algorithms. *International Journal of Mining Science and Technology* 33(8):1019–1036. <https://doi.org/10.1016/j.ijmst.2023.06.001>
- Li D, Liu Z, Armaghani DJ, Xiao P, Zhou J (2022) Novel ensemble intelligence methodologies for rockburst assessment in complex and variable environments. *Sci Rep* 12(1):1844. <https://doi.org/10.1038/s41598-022-05594-0>
- Li P, Tan Z, Yan L, Deng K (2011) Time series prediction of mining subsidence based on a SVM. *Mining Science and Technology (China)* 21(4):557–562. <https://doi.org/10.1016/j.mstc.2011.02.025>
- Liang W, Luo S, Zhao G, Wu H (2020) Predicting Hard Rock Pillar Stability Using GBDT, XGBoost, and LightGBM Algorithms. *Mathematics* 8(5):765. <https://doi.org/10.3390/math8050765>
- Liu R, Liu E, Yang J, Li M, Wang F (2006) Optimizing the Hyper-parameters for SVM by Combining Evolution Strategies with a Grid Search. In: Huang D-S, Li K, Irwin GW (eds) *Intelligent Control and Automation*. Springer Berlin Heidelberg, pp 712–721
- Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems* :Pages 4768–4777
- Mathews KE, Hoek E, Wyllie DC, Stewart, S (1980) *Prediction Of Stable Excavation Spans for Mining at Depths below 1000 Metres in Hard Rock*. Ottawa, ON
- Nawi NM, Atomi WH, Rehman MZ (2013) The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks. *Procedia Technology* 11:32–39. <https://doi.org/10.1016/j.protcy.2013.12.159>
- Nickson SD (1992) Cable support guidelines for underground hard rock mine operations. <https://doi.org/10.14288/1.0081080>
- Potvin Y (1988) Empirical open stope design in Canada. <https://doi.org/10.14288/1.0081130>
- Purwanto, Shimada H, Sasaoka T, Wattimena RK, Matsui K (2013) Influence of Stope Design on Stability of Hanging Wall Decline in Cibaliung Underground Gold Mine. *IJG* 04(10):1–8. <https://doi.org/10.4236/ijg.2013.410A001>
- Qi C, Fourie A, Du X, Tang X (2018a) Prediction of open stope hangingwall stability using random forests. *Nat Hazards* 92(2):1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>
- Qi C, Fourie A, Ma G, Tang X, Du X (2018b) Comparative Study of Hybrid Artificial Intelligence Approaches for Predicting Hangingwall Stability. *J Comput Civ Eng* 32(2):04017086. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000737](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000737)

Raju VNG, Lakshmi KP, Jain VM, Kalidindi A, Padma V (2020) Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification. In: 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, Tirunelveli, India, pp 729–735

Saadaari FS, Mireku-Gyimah D., Olaleye BM (2020) Development of a Slope Stability Prediction Model Using Ensemble Learning Techniques - A Case Study. *GM* 20(2):18–26. <https://doi.org/10.4314/gm.v20i2.3>

Samui P (2008) Slope stability analysis: a support vector machine approach. *Environ Geol* 56(2):255–267. <https://doi.org/10.1007/s00254-007-1161-4>

Santos AEM, Amaral TKM, Mendonça GA, Silva DDFSD (2020) Open stope stability assessment through artificial intelligence. *REM, Int Eng J* 73(3):395–401. <https://doi.org/10.1590/0370-44672020730012>

Srivastava M (2023) Addressing Spurious Correlations in Machine Learning Models: A Comprehensive Review. *Open Science Framework*

Suorineni FT, Kaiser PK, Tannant DD (2001) Likelihood statistic for interpretation of the stability graph for open stope design. *International Journal of Rock Mechanics and Mining Sciences* 38(5):735–744. [https://doi.org/10.1016/S1365-1609\(01\)00033-8](https://doi.org/10.1016/S1365-1609(01)00033-8)

Szmigiel A, Apel DB (2022) Predicting the stability of open stopes using Machine Learning. *Journal of Sustainable Mining* 21(3):241–248. <https://doi.org/10.46873/2300-3960.1369>

Tishkov M (2018) Evaluation of caving as a mining method for the Udachnaya underground diamond mine project. In: *Proceedings of the Fourth International Symposium on Block and Sublevel Caving*. Australian Centre for Geomechanics, Perth, pp 835–846

Vinay LS, Bhattacharjee RM, Ghosh N, Kumar S (2023) Machine learning approach for the prediction of mining-induced stress in underground mines to mitigate ground control disasters and accidents. *Geomech Geophys Geo-energ Geo-resour* 9(1):159. <https://doi.org/10.1007/s40948-023-00701-5>

Wang J, Milne D, Pakalnis R (2002) Application of a neural network in the empirical design of underground excavation spans. *Mining Technology* 111(1):73–81. <https://doi.org/10.1179/mnt.2002.111.1.73>

Wang R (2012) AdaBoost for Feature Selection, Classification and Its Relation with SVM, A Review. *Physics Procedia* 25:800–807. <https://doi.org/10.1016/j.phpro.2012.03.160>

Wang R, Chen S, Li X, Tian G, Zhao T (2023) AdaBoost-driven multi-parameter real-time warning of rock burst risk in coal mines. *Engineering Applications of Artificial Intelligence* 125:106591. <https://doi.org/10.1016/j.engappai.2023.106591>

Xu Y, Goodacre R (2018) On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *J Anal Test* 2(3):249–262. <https://doi.org/10.1007/s41664-018-0068-2>

Zhou J, Huang S, Tao M, Khandelwal M, Dai Y, Zhao M (2023) Stability prediction of underground entry-type excavations based on particle swarm optimization and gradient boosting decision tree. *Underground Space* 9:234–249. <https://doi.org/10.1016/j.undsp.2022.08.002>

Zhou J, Li X, Mitri HS (2015) Comparative performance of six supervised learning methods for the development of models of hard rock pillar stability prediction. *Nat Hazards* 79(1):291–316. <https://doi.org/10.1007/s11069-015-1842-3>

Zhou J, Shi X, Du K, Qiu X, Li X, Mitri HS (2017) Feasibility of Random-Forest Approach for Prediction of Ground Settlements Induced by the Construction of a Shield-Driven Tunnel. *Int J Geomech* 17(6):04016129. [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000817](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000817)

CHAPTER 7: SUMMARY, CONCLUSION AND RECOMMENDATIONS

This chapter, drawing from previous chapters, provides a summary of the whole thesis as well as the conclusion. In addition, this chapter also tells the research prospects about using machine learning methods in predicting the stability of open stopes in underground mining operations.

7.1 Summary of the research

This thesis addresses the critical need to assess the stability of open stopes in mining operations, highlighting the importance of ensuring safety and longevity in underground excavations. Despite the existence of classical stability assessment methods, there remains a demand for improvement in predicting stope stability.

Evaluating the possibility and applicability of various machine learning approaches was a primary objective of this research, driven by the recognition of the potential for these techniques to enhance stope stability assessment in mining operations. By systematically comparing and evaluating different machine learning models, I aimed to identify the most effective approaches for predicting stability conditions. This process involved exploring a diverse range of algorithms, considering factors such as model performance, interpretability, and computational efficiency. Through rigorous evaluation, I aimed to uncover insights into the strengths and limitations of each approach, enabling us to make informed decisions about their suitability for practical implementation. By conducting this comprehensive analysis, I aimed to advance the field of stope stability assessment and contribute to the development of effective and reliable predictive models that can support decision-making in mining engineering.

Chapter 2 delves into the engineering background of open stope mining operations, presenting the terminology and classification methods essential for understanding mining techniques. It offers insights into rock mass classification, laying the groundwork for subsequent stability assessments. In Chapter 3, a comprehensive literature review examines contemporary methods for stope stability assessment, including empirical approaches, statistical analyses, and machine learning techniques. This review critically evaluates existing methodologies, highlighting their strengths, limitations, and potential applications in mining engineering.

Chapters 4, 5, and 6 focus on developing and evaluating machine learning models to predict stope stability. Chapter 4 utilizes the Potvin database to compare Random Forest and Logistic Regression models, while Chapter 5 conducts an in-depth analysis to develop an effective Artificial Neural Network (ANN) model. Both chapters explore feature importance to identify parameters influencing stope stability. Finally, Chapter 6 expands on these findings by analyzing a larger database from the literature and comparing multiple machine learning models. This

chapter provides insights into the effectiveness of various models in predicting stability conditions, emphasizing the importance of feature selection and model evaluation.

7.2 Conclusions of this research

Throughout this research, a comprehensive methodology has been developed to employ machine learning techniques for predicting the stability of open stopes in underground mining operations. This methodology included the development and application of machine learning models using database obtained from prior publications and publicly available datasets. The process involved data collection, preprocessing, feature engineering, and model evaluation to ensure effectiveness and reliability. As a result, several key conclusions have emerged from this study. These conclusions will be presented chapter by chapter, presenting the insights gained from each of them.

The introductory Chapter 1 presents the foundation for understanding the critical importance of assessing the stability of underground openings in mining operations. Several tasks that were to be accomplished in this research are presented, including an extensive literature review, a thorough data collection, development of machine learning models and feature importance analysis. All listed research objectives were achieved in this study.

Chapter 2 provides a comprehensive overview of open stope mining methods employed, where the operations are conducted by extracting substantial ore blocks using drilling and blasting techniques within underground ore deposits. The chapter presents the classification of open stope mining methods, explaining various factors influencing their categorization. Factors such as mining direction (longitudinal or transverse), the use of pillars and backfill, and drillhole diameter (longhole or blasthole) are discussed, highlighting the different approaches required for different mining scenarios. Moreover, the chapter explores popular rock mass classification methods developed over the years, emphasizing their significance in assessing the stability of openings. These classification systems offer systematic approaches to evaluate rock mass conditions, incorporating various parameters to quantify its behavior. By presenting these methods, Chapter 2 sets the stage for subsequent discussions on stability assessments and predictive modeling.

Chapter 3 delves into various methodologies developed for evaluating the stability of open stope mining. The chapter begins by discussing analytical methods used to assess stress-induced and gravity-induced failures, including kinematic, beam failure, and plate buckling analyses. Empirical methods, such as stability graphs, are also explored, emphasizing their importance, followed by the review of statistical and numerical modeling methods employed and presented in literature over the years.

One of the highlights of Chapter 3 is the review of machine learning methods for predicting stope stability. Several studies are discussed, showcasing the effectiveness of machine learning algorithms in analyzing data to identify patterns and correlations related to stope stability. Erdogan Erten et al. (2021) introduced a hybrid artificial neural network approach, while Saadaari et al. (2020) investigated the viability of ensemble learning methods. Qi et al. (2018b) explored various artificial intelligence strategies for predicting stope stability, highlighting the remarkable predictive capabilities of these machine learning models. Overall, the chapter emphasizes the importance of ongoing research in stope stability assessment, particularly in the context of advancing mining engineering practices.

Chapter 4 explores the application of Machine Learning (ML) models to predict open stope stability in underground mining, addressing the limitations of traditional stability assessment methods. Using the Potvin database, Logistic Regression and Random Forest algorithms were employed, achieving promising results. Logistic Regression demonstrated an average accuracy of 0.68, while Random Forest performed slightly better with an average accuracy of 0.71, especially in predicting unstable zones. The Potvin database, comprising data from 34 mines, provided essential parameters related to rock mass characteristics, stress conditions, and other factors influencing stope stability. Machine learning models were trained and evaluated using various methods such as k-fold Cross-Validation, Confusion Matrix, and ROC-AUC score. Logistic Regression, based on stability number (N) and shape factor (HR), struggled with correctly predicting unstable zones due to similar values to other classes. Despite this, its AUC score was 0.81 for the training set and 0.78 for the test set, which are considered excellent and acceptable, respectively. Cross-Validation yielded an average accuracy of 0.68.

Random Forest, after hyper-parameter tuning, achieved an accuracy of 0.84 for the training set and 0.75 for the test set. The AUC score was 0.96 for training and 0.83 for testing, indicating excellent and outstanding performance. Cross-Validation resulted in an average accuracy of 0.71. Both algorithms showed satisfactory performance, with Random Forest outperforming Logistic Regression, particularly in predicting unstable zones.

Chapter 5 of the study focuses on the development and optimization of an artificial neural network (ANN) model. The study again utilized Potvin's database, consisting of cases of open stopes, each characterized by various features describing the surrounding rock mass properties, geological conditions, and opening dimensions. In this study, more parameters were used as input parameters for the model. In previous study presented in Chapter 4, stability number N and hydraulic radius (HR) were considered, however in this research, parameters such as Q value, rock stress factor (A), joint orientation adjustment factor (B), effect of the gravity (C) and hydraulic radius (HR) were passed as separate inputs for the model.

Before constructing the ANN model, the extensive preprocessing of the data was conducted. This involved techniques such as feature scaling (using both `MinMaxScaler` and `StandardScaler` methods), encoding categorical data, and splitting the dataset into training, validation, and testing sets. The ANN model's architecture was carefully designed to suit the characteristics of the dataset. The input layer consisted of five neurons, corresponding to the five input features in Potvin's database. Subsequent hidden layers were developed based on experimentation, literature review, and validation on a separate validation set. During model development, various configurations were tested and evaluated using metrics such as accuracy, loss, confusion matrix, and classification report. The objective was to find a balance between accuracy and generalization, ensuring that the model could make accurate predictions on unseen data without overfitting to the training set. Through careful adjustments and fine-tuning of hyperparameters, the optimal configuration with the best results was identified: using `StandardScaler` for data scaling and `Swish` activation function for all hidden layers. The performance of the final ANN model was assessed on an independent testing set, consisting of 18 examples not used for training or validation. The model achieved an overall accuracy of 83%, indicating its ability to correctly predict the stability class for the majority of instances. Furthermore, the precision, recall, and F1 score metrics provided insights into the model's performance across different stability classes, highlighting its ability to minimize

misclassifications. In addition to evaluating the model's predictive capabilities, the analysis of the importance of features using the SHAP (Shapley Additive exPlanations) tool was performed. This analysis revealed that the Hydraulic Radius (HR) and Q value had the most significant influence on stability predictions.

Chapter 6 of my study focuses on the development, optimization, and comparison of various machine learning model, followed by thorough feature importance analysis. In this study I considered even more factors as input parameters, particularly hydraulic radius (HR), which in this case was separated into Height (H), Span (S) and Length (L). This allowed us to evaluate each dimension of the stopes separately, rather than one shape parameter (HR) representing a ratio of those dimensions.

I started by analyzing the Adoko (2022) dataset, which contains stope dimensions and geomechanical properties from three mines in Ghana. The objective in this study was to explore and compare various machine learning algorithms, including Random Forest, Support Vector Machine (SVM), AdaBoost, XGBoost, LightGBM, and Artificial Neural Network (ANN), to determine their effectiveness in predicting stope stability.

After preprocessing the data, which includes feature scaling, encoding categorical variables, and splitting the dataset into training, validation, and testing sets, I trained each model and evaluated their performance using metrics such as accuracy, precision, recall, and F1 score. I find that ANN consistently outperforms the other models, achieving an accuracy score of 91% on the testing set. To gain insights into the factors influencing stability predictions, I conduct SHapley Additive exPlanations (SHAP) analysis. This analysis reveals that the Q value, representing rock mass quality, is the most influential factor in models like Random Forest and SVM. In contrast, models like AdaBoost, XGBoost, and LightGBM prioritize shape parameters, particularly span. However, despite their focus on shape parameters, these models exhibit less accurate overall performances. The ANN model identifies span (S) as the most influential factor, but due to its inconsistency, I recommend relying on the second most influential factor, Q, for more accurate stability predictions.

This research presents a novel approach in assessing the stability of open stopes. Unlike traditional graph-based methods, which typically allow for stability assessment using only two

parameters, the ML techniques employed in this study enabled the incorporation of a wider range of parameters into the stability evaluation process. This comprehensive inclusion of multiple factors facilitated the discovery of more complex patterns and relationships that were previously undetectable with conventional methods.

In summary, this study establishes a framework for integrating machine learning techniques into the prediction of slope stability. The findings demonstrate the viability of data-driven approaches in traditional mining engineering practices, indicating their potential to complement existing experience-based and mechanism-based methods while offering a new perspective on challenges like stability prediction. Although the adoption of data-driven methods in mining has been limited so far, the success of these applications suggests a growing confidence in applying machine learning within mining engineering disciplines.

7.3 The prospects of future research

Looking ahead, the future prospects in the field of employing machine learning techniques for slope stability predictions appear promising, with several prospects for advancement and exploration. As technology continues to evolve and computational power increases, there is vast potential for the refinement and optimization of machine learning models to enhance their predictive accuracy and reliability. Additionally, the integration of advanced sensing technologies and real-time monitoring systems within mining operations can provide a continuous stream of data, enabling more dynamic and responsive stability assessments. Furthermore, the development of hybrid models that combine machine learning with physics-based modeling approaches have a great potential for improving predictive capabilities while incorporating domain-specific knowledge and principles.

However, it is essential to acknowledge the limitations of current research efforts, such as the reliance on relatively small databases, which may restrict the generalizability and applicability of machine learning models. Addressing this challenge will require concerted efforts to collect larger and more diverse datasets, potentially through collaboration among industry. Future studies could explore the potential benefits of combining the two databases that were investigated in this research. This integrated approach could improve the predictive accuracy and reliability of the models. Additionally, the collection of databases from different world regions could also enhance

the applicability of the models. By incorporating data from a diverse range of geological conditions, the models can be trained on more varied and representative data. This would allow for the development of machine learning models with improved generalizability, capable of providing accurate stability predictions across different mining environments globally. Moreover, ongoing research should focus on refining feature selection methods, improving model interpretability, and addressing issues related to data quality. By addressing these challenges and focusing on improving technologies and methodologies, the field of machine learning-based stope stability predictions is capable to make significant advancements, ultimately contributing to safer and more efficient mining practices in the future.

BIBLIOGRAPHY

Adoko AC, Saadaari F, Mireku-Gyimah D, Imashev A (2022) A Feasibility Study on The Implementation of Neural Network Classifiers for Open Stope Design. *Geotech Geol Eng* 40(2):677–696. <https://doi.org/10.1007/s10706-021-01915-8>

Barton N, Lien R, Lunde J (1974b) Engineering classification of rock masses for the design of tunnel support. *Rock Mech Felsmech Mec Roches* 6(4):189–236. <https://doi.org/10.1007/BF01239496>

Bienawski ZT (1976) Rock Mass Classification in Rock Engineering. In *Proceedings of the Symposium on Exploration for Rock Engineering*

Bienawski ZT (1989) *Engineering rock mass classification.*, John Wiley&Sons. New York

Bieniawski ZT (1978) Determining rock mass deformability: experience from case histories. *Int J Rock Mech Min Sci Geomech Abstr* 15(5):237–247. [https://doi.org/10.1016/0148-9062\(78\)90956-7](https://doi.org/10.1016/0148-9062(78)90956-7)

Brady BHG, Brown ET, Brown ET (2006) *Rock mechanics for underground mining*, 3. ed., repr. with corr. Springer, Dordrecht

Capes GW (2009a) Open stope hangingwall design based on general and detailed data collection in unfavourable hangingwall conditions, NR62618 Ph.D. The University of Saskatchewan (Canada)

Carbonell JG, Michalski RS, Mitchell TM (1983) Machine Learning: A Historical and Methodological Analysis. *AI Mag* 4(3):69–69. <https://doi.org/10.1609/aimag.v4i3.406>

Erdogan Erten G, Bozkurt Keser S, Yavuz M (2021) Grid Search Optimised Artificial Neural Network for Open Stope Stability Prediction. *Int J Min Reclam Environ* 35(8):600–617. <https://doi.org/10.1080/17480930.2021.1899404>

Goodman RE (1989) *Introduction to rock mechanics*, 2nd ed. Wiley, New York

Hoek E, Brown ET (1980b) Underground excavations in rock, Rev. Institution of Mining and Metallurgy, London

Hoek E, Brown ET (1996) Underground excavations in rock, Rev. 1. ed., repr. The Institution of Mining and Metallurgy, London

Hoek E, Kaiser PK, Bawden WF (2000) Support of Underground Excavations in Hard Rock, 0 edn. CRC Press

Mathews, K. E. Hoek, D. C. Wyllie and S. B. V. Stewart (1981) Prediction Of Stable Excavation Spans for Mining at Depths below 1000 Metres in Hard Rock. Ottawa, ON

Pakalnis RT (1986) Empirical stope design at the Ruttan Mine, Sherritt Gordon Mines Ltd. <https://doi.org/10.14288/1.0081095>

Potvin Y (1988a) Empirical open stope design in Canada. <https://doi.org/10.14288/1.0081130>

Potvin Y, Hudyma M (2000) Open Stope Mining in Canada

Potvin Y, Hudyma M (1989) Open Stope Mining Practices in Canada. Quebec City

Potvin Y, Nedin P (2003) Management of rockfall risks in underground metalliferous mines : a reference manual. Minerals Council of Australia

Qi C, Fourie A, Du X, Tang X (2018a) Prediction of open stope hangingwall stability using random forests. Nat Hazards 92(2):1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>

Qi C, Fourie A, Du X, Tang X (2018b) Prediction of open stope hangingwall stability using random forests. Nat Hazards 92(2):1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>

Qi C, Fourie A, Ma G, Tang X, Du X (2018c) Comparative Study of Hybrid Artificial Intelligence Approaches for Predicting Hangingwall Stability. J Comput Civ Eng 32(2):04017086. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000737](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000737)

Saadaari FS, Mireku-Gyimah D., Olaleye BM (2020) Development of a Stope Stability Prediction Model Using Ensemble Learning Techniques - A Case Study. Ghana Min J 20(2):18–26. <https://doi.org/10.4314/gm.v20i2.3>

Santos AEM, Amaral TKM, Mendonça GA, Silva D de FS da (2020a) Open stope stability assessment through artificial intelligence. REM - Int Eng J 73(3):395–401. <https://doi.org/10.1590/0370-44672020730012>