Prediction of Mechanical Properties of Rocks using ML-Applications

Students: Chapati Anusha, Chitturi Laasya Priya
Faculty Advisor: Dr.Ankesh Kumar

Abstract

This project explores the application of machine learning (ML) techniques to predict the mechanical properties of rocks, specifically the uniaxial compressive strength (UCS), Elasticity and Brazilian Tensile Strength. The study employs various ML algorithms, including Support Vector Machine (SVM),Back propogation Artificial Neural Network (BPNN),Random Forest (RF), K-Nearest Neighbors (KNN), Linear Regression.The research utilizes a dataset comprising indirect parameters such as density ,porosity, p-wave velocity, and point load strength index as inputs to train and evaluate the performance of the selected machine learning models. The models undergo rigorous comparison and analysis to identify the most effective approach for predicting UCS, Elasticity and Brazilian Tensile Strength.Based on the findings from parallel test results, it can be concluded that SVM is particularly advantageous, especially when dealing with large-scale problems. This preference stems from the high efficiency of its solution procedure, making SVM a favorable choice for addressing challenges associated with extensive datasets or complex computational tasks.

1. Introduction

In Geotechnical engineering, the physico-mechanical parameters of rocks are crucial for designing structures such as caverns, tunnels, dams, slope stability, underground space, and excavation projects. Uniaxial Compressive Strength (UCS), Elasticity and Brazilian Tensile Strength are pivotal metrics in this context. Traditional methods include subjecting rock samples to uniaxial compression tests on cylindrical or cubical rock samples until failure occurs for UCS and employing stress-strain tests for Elasticity. However, these methods are time-consuming and labor-intensive and may lead to rock breakage, compromising sample integrity. Additionally, the repetitive nature of these tests falls short of capturing the intricate and non-linear interactions influencing rock behavior. Recognizing these limitations, the integration of Machine Learning (ML) applications becomes transformative, offering an efficient and non-destructive means of predicting UCS , Elasticity and Brazilian Tensile Strength. This paradigm shift mitigates risks associated with repeated laboratory tests and enhances the feasibility and reliability of rock property predictions, contributing significantly to the advancement of geotechnical engineering practices in diverse construction endeavors, including those related to dams, underground spaces, and excavation projects.

2. Methodology

2.1. Overview

1. Data Collection
2. Data Preprocessing
    - Fixing input and output Parameters
    - Finding missing Data
    - Splitting dataset into training and test set
    - Feature scaling
3. Machine Learning Algorithms
    - SVM Regression
    - BPNN
    - RF
    - KNN Regression
    - Linear Regression
4. Evaluation measures
    - R2 score
    - MAE
    - MAPE
    - MSE
    - RMSE

2.2. Data Collection

To train and validate the machine learning models in predicting the mechanical properties of rocks, we first collected a few datasets containing density, porosity, p-wave velocity, point load strength index, Brazilian tensile strength, UCS, Elasticity, Cohesion and friction values, Poisson ratio of different rocks from various research papers. Typical rock types in rock engineering were included, such as granite, basalt, sandstone, limestone, phyllite, and quartzite.

2.3. Data Preprocessing

2.3.1. Fixing input and output Parameters

In our comprehensive data collection process for various properties, we meticulously filtered out values lacking complete property information. Subsequently, we identified correlations among key physical and mechanical properties. Specifically, we have opted to utilize density, porosity,p-wave velocity, and point load strength index as input parameters for our analysis. In doing so, our focus is on understanding the relationships that exist between these input parameters and the uniaxial compressive strength (UCS) ,Elasticity and Brazilian Tensile Strength, which have been selected as the output parameters for our study.

2.3.2. Finding Missing Data

In addressing missing values in our dataset, we implemented a two-pronged approach: dropping missing values and imputing them using methods like KNN, mean, and median. While dropping values offered simplicity, the potential data loss prompted a thorough exploration of imputation techniques. Among them, the KNN imputer stood out for its superior performance in preserving overall dataset structure and relationships. Consequently, we opted to rectify missing values specifically with the KNN imputation method, ensuring a comprehensive and reliable dataset for subsequent analyses.

2.3.3. Splitting dataset

In the division of our dataset for model training and testing, we allocated 80 percent for training and 20 percent for testing. This strategic split ensures a robust evaluation of our models, balancing training efficiency with effective validation of unseen data.

2.3.4. Feature Scaling

In our machine learning approach, feature scaling has been employed to standardize input variables, ensuring a consistent scale for effective model training. This normalization enhances model convergence and performance, mitigating the impact of disparate scales across features.

2.4. Machine Learning Algorithms

In our robust exploration of machine learning algorithms, we employed a diverse set to cater to the intricate patterns within our dataset. Support Vector Machine (SVM) excels in handling complex relationships by transforming data into high-dimensional space for effective separation. Random Forest (RF), a powerful ensemble learning algorithm, aggregates predictions from multiple decision trees, providing robust and accurate results. Backpropagation Neural Network (BPNN) utilizes artificial neural networks to learn intricate patterns and relationships within the data, offering flexibility in capturing non-linear dependencies. K-Nearest Neighbors (KNN) classifies data points based on proximity, making it effective for both classification and regression tasks. Linear Regression establishes linear relationships between variables, providing simplicity and interpretability in modeling.

2.5. Evaluation Measures

To assess forecasting model accuracy, we consider statistical evaluation indices such as R2, MSE, RMSE, MAE, and MAPE. Here yi is the actual value, y'i is the predicted value and n is the number of samples.

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{sum of squares total (SST)}}$$

$$\text{MAE} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \left| y_i - y_i' \right|$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - y_i'}{y_i} \right| \times 100\%$$

$$\text{MSE} = \frac{1}{n} \sum_{1}^{n} \left(y_i - y_i' \right)^2$$

$$\text{RMSE} = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^{n} \left(y_i - y_i' \right)^2}$$

3. Model Training and Results

SVM: In the model training phase of Support Vector Machine (SVM) training with various kernels, including linear, poly, and rbf After thorough trial and error, the optimal choice for the kernel was determined to be rbf. Additionally, the regularization parameter "C" in SVM, crucial for balancing the trade-off between maximizing the margin and minimizing training errors, was set to a value of 25, ensuring an effective and well-tuned model.

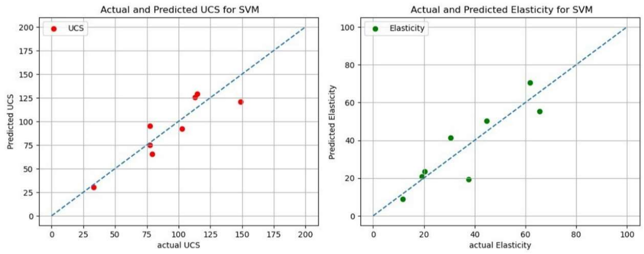
BPNN: We trained BPNN with first and second hidden layer neurons as 100 and 50 respectively, utilizing the "identity" activation function, and a random state of 42. Alternative configurations worth exploring include varying hidden layer sizes, trying different activation functions such as "logistic" or "tanh," adjusting random state values, and experimenting with additional hyperparameters like learning rate and momentum.

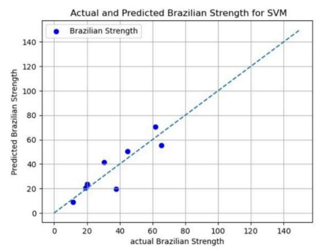
RF:Here in Random Forest we chose 200 estimators and a random state of 22, determined through iterative testing for optimal performance. Other tuning possibilities include adjusting tree depth, minimum samples for a split, minimum samples per leaf, and exploring criteria like Gini impurity or entropy for node splitting. These variations offer flexibility to enhance the model based on dataset characteristics and specific objectives.

Linear Regression: Here Linear Regression was employed, utilizing its simplicity and interpretability to establish linear relationships between input variables and facilitate predictions with transparency

KNN: Here K-Nearest Neighbors (KNN) algorithm was employed with a specified parameter of 5 neighbors. The selection of this configuration followed a process of trial and error, culminating in the determination that this particular setting yielded the most optimal performance.

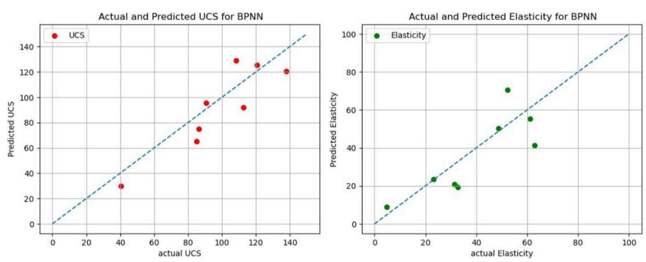
Figure 1: Results with SVM

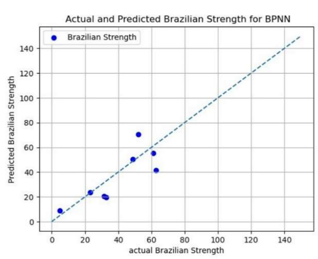




Evaluation	R2	MAE	MAPE	MSE	RMSE
UCS	0.77	12.99	13.85	226.4	15.04
Elasticity	0.78	7.57	26.28	84.41	9.19
Brezilian Tensile strength	0.70	2.86	21.80	15.80	3.97

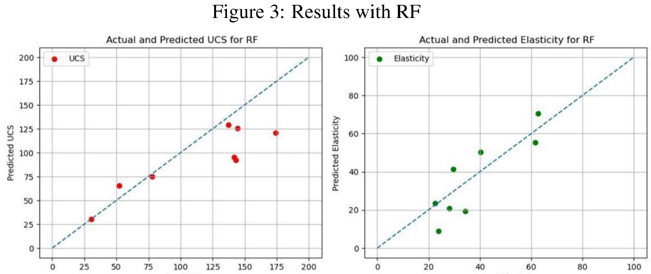
Figure 2: Results with BPNN

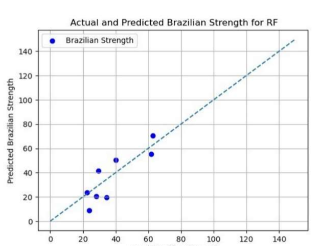




Evaluation	R2	MAE	MAPE	MSE	RMSE
UCS	0.76	13.76	17.78	229.21	15.13
Elasticity	0.64	9.4	32.31	143.04	11.9
Brezilian Tensile strength	0.38	4.26	35.92	33.49	5.78

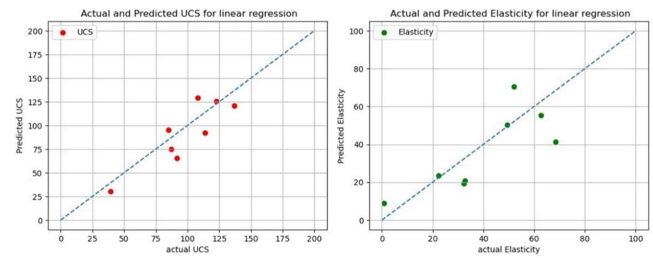
Figure 3: Results with RF

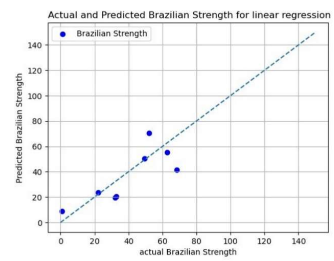




Evaluation	R2	MAE	MAPE	MSE	RMSE
UCS	0.01	24.23	24.27	1025.9	32.02
Elasticity	0.73	9.25	43.96	104.90	10.24
Brezilian Tensile strength	0.39	4.64	40.61	32.83	5.72

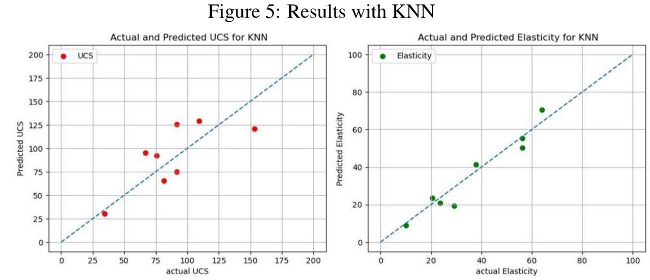
Figure 4: Results with Linear Regression

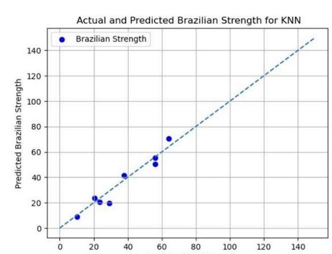




Evaluation	R2	MAE	MAPE	MSE	RMSE
UCS	0.72	15.09	19.36	280.57	16.75
Elasticity	0.52	11.01	40.79	188.89	13.74
Brezilian Tensile strength	0.39	4.02	32.06	32.91	5.73

Figure 5: Results with KNN





Evaluation	R2	MAE	MAPE	MSE	RMSE
UCS	0.48	21.07	22.40	530.41	23.03
Elasticity	0.93	4.19	15.03	25.73	5.07
Brezilian Tensile strength	0.44	4.34	33.41	30.07	5.48

4. Summary

Our results demonstrated that the SVM-based model shows a better performance than all the other models. When considering R2-Score for the overall model(i.e. taking UCS, Elasticity, and Brazilian Tensile Strength) we have got the following result. The overall R2 Score is 0.758 for SVM, 0.601 for BPNN, 0.376 for RF, 0.548 for Linear Regression, and 0.629 for KNN Regression.

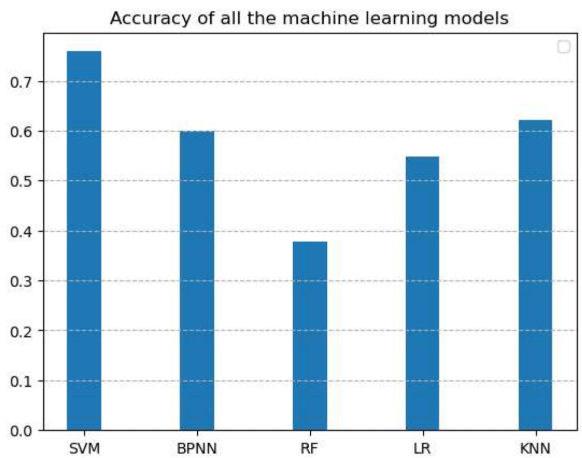


Figure 6: Overall R2 score for all the 3 Mechanical properties

Notably, the SVM algorithm exhibited a slight advantage over other algorithms. Precisely, SVM showed more relatedness to UCS, KNN to Elasticity, and BPNN to Brazilian Strength.

5. Conclusions

This project explored the viability of applying machine learning models to predict various mechanical properties of rocks using physical properties. Subsequently, five machine learning models were developed utilizing Support Vector Machines (SVM), Backpropagation Neural Networks (BPNN), Random Forest (RF), Linear Regression and K-Nearest Neighbours(KNN) Algorithms. The output of these models was the compressive strength, Elasticity, and Brazilian Tensile Strength with 4 physical properties of inputs.

Following the training and validation of the models, an analysis was conducted. The findings of this study demonstrated the efficiency of machine learning models in reliably predicting the mechanical properties of various types of rocks.

6. Future Scope

The future scope of our project involves the exploration of advanced machine learning algorithms and ensemble methods to further enhance the accuracy and robustness of our prediction models for rock mechanical properties. Consideration will be given to incorporating additional geological features or advanced testing parameters, broadening the scope of our models. Furthermore, we aim to investigate the feasibility of real-time monitoring systems, allowing for continuous data collection and the development of adaptive models that can dynamically adjust to changing geological conditions. This evolution will ensure that our ML applications for predicting the mechanical properties of rocks remain at the forefront of innovation and contribute to a deeper understanding of rock behavior in diverse environmental contexts.

7. References

1.Tarun Singh, Ashwani Jain, K.S. Rao, booktitle=Physico-mechanical Behaviour of Metamorphic Rocks in Rohtang Tunnel, Himachal Pradesh, India, year=2017,
2.Tarun SinghTarun SinghSeshagiri K. RaoSeshagiri K. RaoMainak Ghosh RoyMainak Ghosh Roy, booktitle=Engineering Behaviour and Classification of Weathered Rocks , article Geomechanics and Geodynamics of Rock Masses. Volume IIEdition: 1st EditionPublisher: CRC Press, London year=1996,
3.M. H. NasseriSeshagiri K. RaoSeshagiri K. RaoT. Ramamurthy, booktitle=Physico-mechanical Behaviour of Metamorphic Rocks in Rohtang Tunnel, Himachal Pradesh, India, year=2018,
4.Anand S. GuptaAnand S. GuptaSeshagiri K. RaoSeshagiri K. Rao, booktitle=Strength behaviour of Middle Siwalik sandstone under uniaxial and triaxial conditions, year=1996,
5.Guan Rong, Song Sha, Bowen Li, Zhiheng Chen Ziyang Zhang, booktitle=Experimental Investigation on Physical and Mechanical Properties of Granite Subjected to Cyclic Heating and Liquid Nitrogen Cooling, year=2021,
6.panelXiaoguang Wu, Zhongwei Huang, Zhen Cheng, Shikun Zhang, Hengyu Song, Xin Zhao, booktitle=Effects of cyclic heating and LN2-cooling on the physical and mechanical properties of granite, year=2019,
7.Hussein Aldeeky , Omar Al Hattamleh , booktitle=Prediction of Engineering Properties of Basalt Rock in Jordan Using Ultrasonic Pulse Velocity Test, year=2018,
8.Wen Li , Zhuoying Tan(2016) , booktitle=Research on Rock Strength Prediction Based on Least Squares Support Vector Machine

8. Signature of Faculty Advisor:

Kumar A.

Dr. Ankesh Kumar