MWD Data-Based Marble Quality Class Prediction Models Using ML Algorithms

Ozge Akyildiz[1] · Hakan Basarir[1] · Veena Sajith Vezhapparambu[2] · Steinar Ellefmo[1]

Abstract

Brønnøy Kalk AS operates an open pit mine in Norway producing marble, mainly used by the paper industry. The final product is used as filler and pigment for paper production. Therefore, the quality of the product has utmost importance. In the mine, the primary quality indicator, called TAPPI, is quantified through a laborious sampling process and laboratory experiments. As a part of digital transformation, measurement while drilling (MWD) data have been collected in the mine. The purpose of this paper is to use the recorded MWD data for the prediction of marble quality to facilitate quality blending in the pit. For this purpose, two supervised classification modelling algorithms such as conventional logistic regression and random forest have been employed. The results show that the random forest classification model presents significantly higher statistical performance, and it can be used as a tool for fast and efficient marble quality assessment.

Keywords Measurement while drilling (MWD) · Machine learning · Random forest · Logistic regression · Marble quality

1 Introduction

Brønnøy Kalk AS is an open pit mine in Northern Norway extracting marble that is processed into a whitening agent and used by the paper industry. The fragmented marble is produced by a conventional drill and blast operation. The mine provides raw materials to one of the largest suppliers of ground calcium carbonate (GCC) in the

✉ Ozge Akyildiz
    ozge.akyildiz@ntnu.no

1   Department of Geoscience and Petroleum, Norwegian University of Science and Technology, Trondheim, Norway

2   Amrita School for Sustainable Development and Amrita Wireless Network and Application, Amrita Vishwa Vidyapeetham, Kollam, Kerala, India

world. GCC is important for paper production as it is used not only as filler to reduce cellulose and energy consumption but also as pigment to improve brightness, opacity, and ink absorption. In the mine, the carbonate deposit constitutes different types of marble strata, which are characterized by impurities and visual characteristics (Watne 2001). In addition to marble, dykes and veins of dolerite, aplite, and pegmatite are abundant in the deposit. However, only the so-called speckled marble is regarded to have the required quality to be mined as a product (Vezhapparambu et al. 2018; Watne 2001). Each blast containing approximately 100,000 metric tons of raw material is divided into smaller blocks or selective mining units (SMUs) based on geometric and visual characteristics. Generally, each SMU is approximately 5,000 tonnes of rock and contains 10 to 15 blast holes. By sampling the rock chips from the blastholes in the unit, a collective sample characterizing the whole SMU is prepared and sent to the laboratory for the assessment of quality, by means of the Technical Association of the Pulp and Paper Industry (TAPPI) index, the primary quality indicator. Produced SMUs are blended carefully considering the quality specifications, and the blend is fed to the primary crusher. Lower-quality SMUs may be stockpiled for future blending with high-quality blocks (Vezhapparambu et al. 2018).

Mining and manufacturing industries will produce prominent amounts of waste which are sometimes toxic. To reduce such waste and related environmental impacts, the industries should adapt the best available techniques for production. One of the United Nation's (UN's) 2030 Sustainable Development Goals (SDG 12) is to ensure the sustainable consumption and production patterns by industries (United Nations 2022). Maximum utilization of raw material is important to minimize waste and related environmental impact. In the particular case at Brønnøy Kalk AS, due to geometry-based definition of the SMU boundaries, some high-quality marble may not be included in high-quality blocks, or low-quality marble may be left in the high-quality block boundaries. Therefore, a high-resolution and improved quality assessment method can increase the production efficiency and provide better blending plans.

Following the digitization policy, measurement while drilling (MWD) data have been used to define the geological features and rock properties since it provides high-resolution data at a low cost (Vezhapparambu and Ellefmo 2020). As part of the policy, MWD data have been regularly collected. MWD is defined as acquisition of data related to drilling parameters gathered from the drilling rig. The main parameters recorded by the MWD system, varying depending on drill rig characteristics, are dampening pressure, feeding pressure, flush air pressure, rotation pressure, percussion pressure and penetration rate.

For the development of rock property predictive models based on MWD data or other information, different machine learning algorithms such as boosting, neural networks, and fuzzy logic have been used (Kadkhodaie-Ilkhchi et al. 2010). Supervised learning algorithms aiming to find the link between inputs and outputs and to make predictions with minimum error are widely used for mining engineering implementations. Depending on the purpose, the output of the algorithm can be either a continuous quantity, as in the case of regression modelling, or labels or classes as in the case of classification modelling. Both regression and classification supervised learning algorithms are widely used in mining engineering-related modelling studies. Leighton (1982) improved blast design considering the peaks in an MWD parameter,

penetration rate, as an indicator of weak zones. Segui and Higgins (2002) also stated that MWD data can be used to improve blast results in mines. They proposed that different types and amount of explosive can be used for each blast hole, according to the information derived from MWD parameters. Basarir et al. (2017) predicted rock quality designation (RQD) by a soft computing method called adaptive neuro fuzzy inference system (ANFIS) using MWD data such as bit rotation, bit load and penetration rate. Recently, Vezhapparambu et al. (2018) established a method to classify rock type utilizing MWD data. For classification supervised learning algorithms, Chang et al. (2007) used logistic regression to predict probabilities of landslide occurrence by analysing the relationships between instability factors and the past distribution of landslides. Hou et al. (2020) established a method to classify rock mass based on a random forest algorithm using cutterhead power, cutterhead rotational speed, total thrust, penetration, advance rate and cutterhead torque collected during a tunnel boring machine (TBM) operation. Zhou et al. (2016) investigated 11 different supervised learning algorithms including naïve Bayes (NB), k-nearest neighbour (KNN), multilayer perceptron neural network (MLPNN) and support vector machine (SVM) for rock burst prediction. Caté et al. (2017) employed machine learning as a tool for the prediction of gold mineralization. They compared six classification methods in terms of detection of gold-bearing intervals using geophysical logs as a predictor.

This study investigates the relation between MWD data and the quality of the marble by means of classification modelling to assess the quality of marble, in terms of the TAPPI index. Thereby, the mine can more accurately define SMU boundaries and thus improve production efficiency. Initially a reliable database was constructed containing MWD operational parameters and TAPPI values. For model development, first, a basic stand-alone classification algorithm, logistic regression (LR) modelling, and then a more advanced ensemble machine learning method, random forest modelling (RF), were used. The prediction performances of the models were thoroughly examined using commonly used performance measures. As a next step, the model's output may be utilized to create a systematic blending strategy to prevent ore loss.

2 Field Studies

The database used in this study was gathered in a recently completed project entitled Increased Recovery in the Norwegian Mining Industry by implementing the geometallurgical concept (InRec). The project was coordinated by the Mineral Production Group at the Norwegian University of Science and Technology and funded by the Norwegian Research Council and three industry partners (Ellefmo and Aasly 2019).

The collected MWD data can be grouped into two subgroups depending on whether they are controlled by the operator or the medium that is drilled. Penetration rate (PR), flush air pressure (FAP), rotation pressure (RP) and dampening pressure (DP) were considered as dependent variables as they change with the conditions of rock characteristics. However, percussion pressure (PP) and feed pressure (FP) were labelled as operator-controlled parameters. PR automatically increases if the rig's rotation, feed and percussion pressures are increased without any change in rock type. From this, it can be concluded that penetration rate alone may not be a suitable indicator of rock

properties and should be combined with other variables. Flush air pressure is defined as the pressure required to carry drill cuttings out of the borehole, so it is highly dependent on borehole depth. Percussion pressure is controlled by the operator and defined as the pressure released by the piston while hammering. The pressure needed to push the entire hammer downward is known as feed pressure. The feed and percussion pressures are adjusted by the rig's control system to regulate the lower and higher extreme values of rotation pressure and dampening pressure (Ghosh et al. 2017; Schunnesson 1998). Rotation pressure is the amount of pressure used to rotate the drilling bit. The rotation or rotation speed is fixed by the operator. Lastly, dampening pressure is the pressure that helps control percussion pressure and is adjusted by the rig (Atlas-Copco 2019).

MWD data collected during field work are comprised of log time (YYYY-MM-DDThh:mm:ss), depth (metres), penetration rate (metre/min), flush air pressure (bar) percussion pressure, (bar), feed pressure (bar), rotation pressure (bar) and dampening pressure (bar) collected from the MWD system of the Atlas Copco (now named Epiroc) T-45 SmartROC drilling rig for 10-cm intervals. In addition, every 3.5 m, a new rod is attached to the drilling string, leading to anomalies in the MWD data. The anomalies are obvious especially in percussion pressure and feed pressure because they are adjusted by the drilling rig system or by the operator to ensure a smooth drilling process (Ghosh et al. 2017). Since such anomalies can mislead the interpretation of the data, filtering algorithms are implemented. The data beyond two standard deviations from the mean of percussion and feed pressure are filtered out. Moreover, in order to eliminate rod change effect, the data coinciding with the depth at which rod was changed was also filtered out. Figure 1 represents the distribution of filtered MWD data, and summary statistics are shown in Table 1.

Pearson correlation coefficients between MWD parameters are shown in Fig. 2. Due to the strong correlation (0.84) between dampening pressure and feed pressure, feed pressure is not included as an independent variable during the modelling process.

The quality or TAPPI data were obtained from laboratory tests conducted on the drill cuttings of five boreholes at different locations, containing pure and impure marble due to the intrusions of non-carbonate lithologies and fracture zones. Holes were selected based on geological variability in collaboration with the on-site mining geologist. Each hole was divided into several sections with different lengths ranging from 1.3 to 3.4 m. Drill cuttings were collected from 29 sections in total, and the quality of each section was determined. To be able to ensure exact georeferencing of the data, "To" and "From" depths were recorded at the beginning and the end of each section (Vezhapparambu 2019). After lab tests, the TAPPI values and corresponding quality class of the section were determined and assigned to the corresponding MWD data. It should be noted that range of TAPPI values and corresponding quality classes were defined considering the blending process necessary for achieving the desired product quality. For example, class 1 corresponds the highest quality, while class 6 refers to waste material with the lowest quality. Due to a data confidentiality agreement, TAPPI values cannot be presented. However, the normalized values used to assess the quality of marble are shown together with assigned classes in Table 2 and Fig. 3.
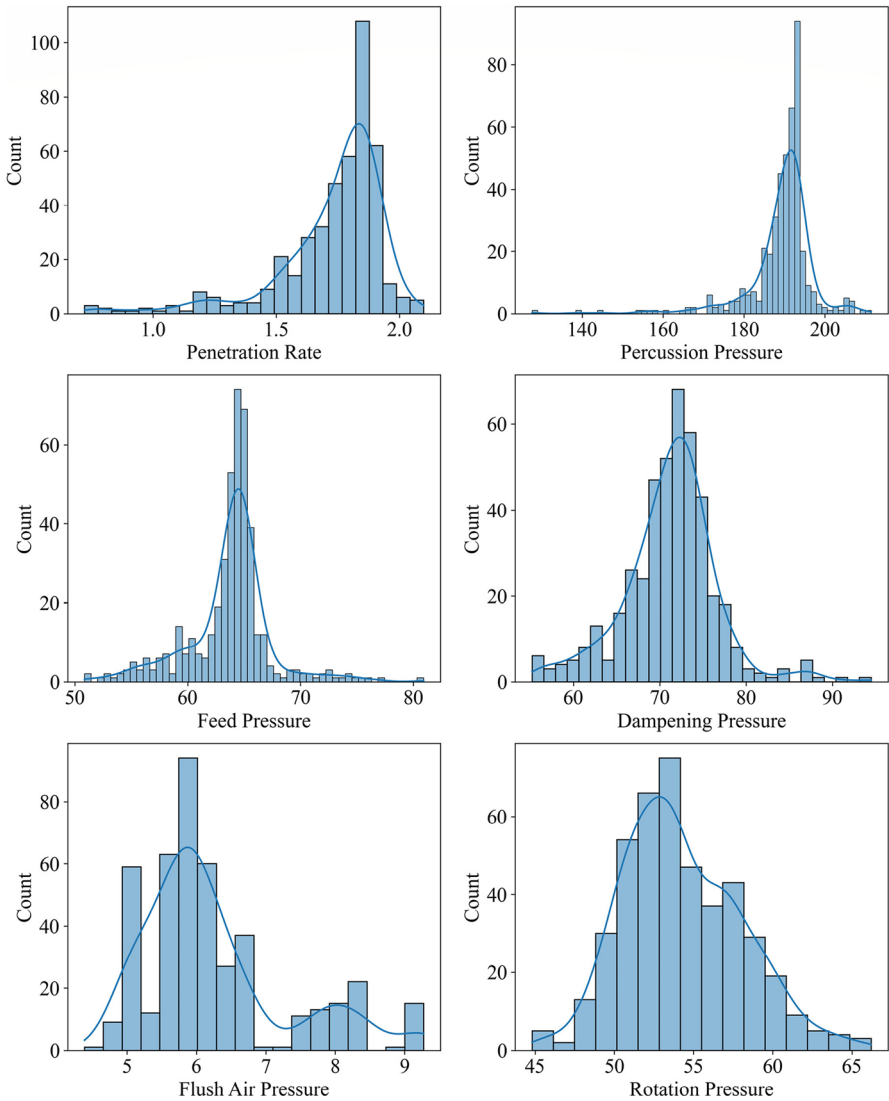
Fig. 1 MWD data distribution

The dataset comprises 441 MWD data used as independent variables and corresponding quality classes used as dependent variables. Figure 4 shows the normalized TAPPI values and corresponding MWD operational parameters together for one of the holes (number 204).

Table 1 Summary statistics of filtered MWD data

MWD	PR	PP	FP	FAP	RP	DP
Mean	1.73	189.26	63.70	6.25	54.28	71.23
Median	0.22	8.33	3.70	1.06	3.74	5.38
Min	0.72	127.63	50.87	4.39	44.81	55.22
Max	2.10	211.42	80.91	9.27	66.21	94.49
Range	1.38	83.78	30.04	4.88	21.40	39.26
SD	1.80	191.01	64.27	5.97	53.76	71.70
# of data	441	441	441	441	441	441
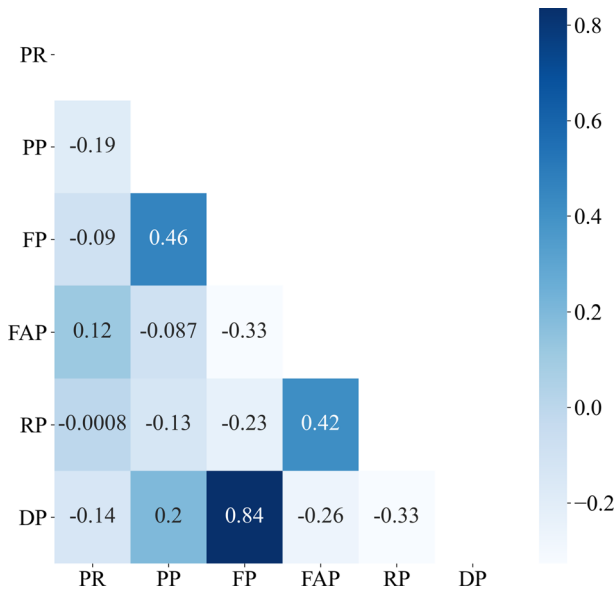
Fig. 2 Pearson correlation coefficients between recorded filtered MWD parameters

Table 2 Normalized TAPPI values and assigned quality classes

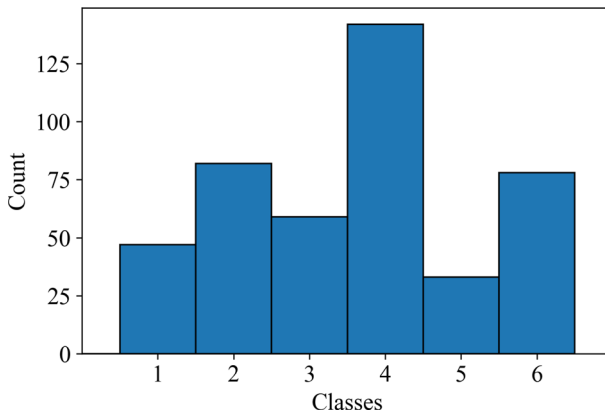| Class	Range of normalized TAPPI	
1	0.83	1.00
2	0.67	0.83
3	0.50	0.67
4	0.33	0.50
5	0.17	0.33
6	0.00	0.17

Fig. 3 Quality class distribution

3 Modelling

For the construction of marble quality classification models, a conventional stand-alone and an ensemble method are used for the prediction of the quality class. Logistic regression (LR) is employed as a conventional model. Afterwards, the ensemble method-based random forest (RF) algorithm combining several decision trees is used. When the dataset has noise and outliers, an advanced method such as RF is expected to yield more precise results compared to conventional methods (Breiman 2001).

Typically, supervised models work quite similarly in principle. The models need a p-dimensional input vector, $X = (X_1, X_2, ..., X_p)^T$ and an output vector, Y. The aim is to find the best function $f(X)$ to predict Y and minimize the loss determined by a loss function or error ($L(Y, f(X))$) measuring the difference between the predicted and measured values (Cutler et al. 2012).

To improve the model's ability to generalize, models need to be tuned by carefully choosing their hyperparameters depending on the characteristics of the data and learning capacity of the algorithm. In this study, the optimum hyperparameters were investigated by the GridSearchCV function provided by the scikit-learn library (Pedregosa et al. 2011). Although it is computationally demanding, grid search is one of the most powerful techniques among the methods used to find optimum hyperparameters. In this method, a list of values for different hyperparameters is provided, and different combinations are evaluated to find the best set in terms of model performance (Raschka 2016). The list is defined by the user and depends on the dataset characteristics such as size of the dataset. Besides grid search, a k-fold cross-validation algorithm was employed in the training stage. In this method, the data are split into two parts as training and testing before the model is trained. While the testing data are kept as unseen data, the training data are split into k folds where k–1 folds are used for the model training, and the remaining fold is used for testing (Raschka 2016). The k-fold validation method enables updating the parameters of the model k times on different k folds of data which results in reducing overfitting (Sadrossadat et al.
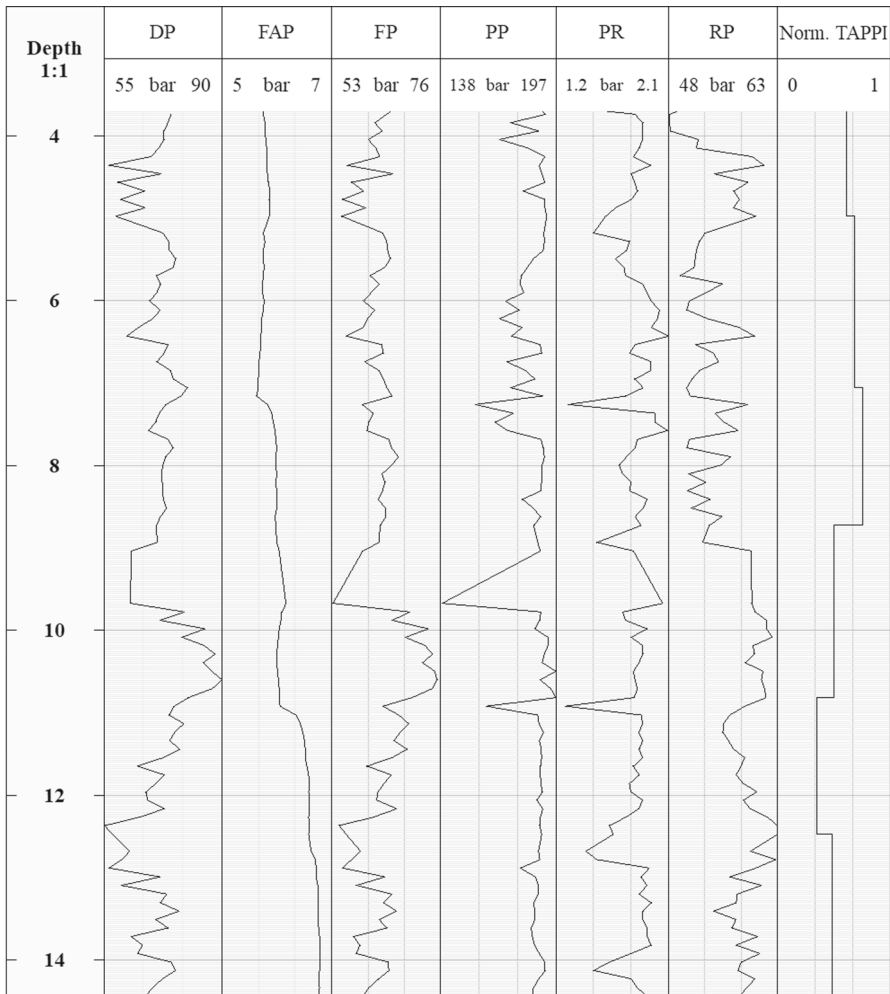
Fig. 4 Recorded MWD data and normalized TAPPI values used in modelling for one of the holes (# 204)

2022). As seen from Fig. 3, the database comprises unbalanced class distribution in the output. To prevent the drawbacks of unbalanced class distribution, a stratified k-fold cross-validation method was implemented. In the training stage, the number of folds was chosen as 10 as suggested in the literature, and an external testing set was used to assess the models' predictive performance (Kohavi 1995; Qi et al. 2020; Rodriguez et al. 2010).

The dataset was randomly divided into two groups as training and testing, adopting a stratified shuffle split method regarding the elimination of an unbalanced class split in the training and testing sets. The training set used in the construction step of the models comprised 352 datapoints (80%). The test set used for validation of the performance and predictive power of the models comprised 89 datapoints (20%). The

models were trained and tuned through Python programming (Guido Van and Fred 2009), JupyterLab web-based user interface (Kluyver et al. 2016) and the scikit-learn library (Pedregosa et al. 2011).

3.1 Logistic Regression (LR) Model

LR is one of the most widely used conventional classification algorithms as it is easy to implement, and it performs quite well for linearly separable cases. LR can be thought of as linear regression used for classification problems. The least square function of linear regression is used in LR, which establishes multivariate correlations between one or more independent and dependent variables, such as

$$Y_n = \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \cdots + \beta_m X_{mn} + \varepsilon_n, \tag{1}$$

where Y_n is the dependent variable; $X = X_{1n}, X_{2n}, \ldots, X_{mn}$ are the independent variables; β_0 is a constant; $\beta_1, \beta_2, \ldots, \beta_m$ are the coefficients of regression; and ε_n is the error. The main difference between linear regression and LR is that LR's output varies between 0 and 1, seen in Fig. 5. It uses the logistic function defined below (Belyadi and Haghigtat 2021)

$$f(x) = \frac{1}{1 + e^{-X}} \; t \in \mathbb{R}, \; f(x) \in (0, 1). \tag{2}$$

The LR model employs a variety of algorithm-specific hyperparameters, mainly penalty, solver and C value. They should be carefully chosen to avoid overfitting as well as underestimation. Penalty indicates the regularization method, for example, l_1, l_2 and Elastic-Net to reduce the loss function. Solver is the algorithm used in
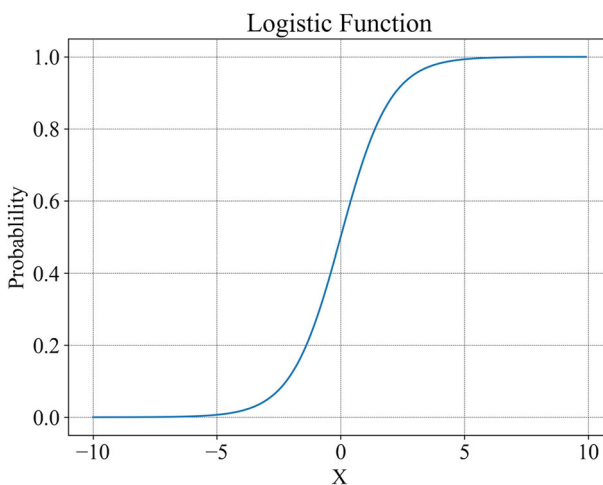

Fig. 5 Graphic representation of logistic function

the optimization problem. C value defines the strength of the regularization. The lower the C-value, the higher the regularization strength. After the hyperparameter optimization by grid search method, the accuracy of the model reached its maximum with the following hyperparameters [solver: newton-cg, penalty: l_1 and C: 10].

3.2 Random Forest (RF) Model

The RF algorithm is based on the idea of ensemble learning, defined as the aggregation of many decision trees. It was introduced by Breiman (2001) and is used for both classification and regression problems. Similarly, input features (independent variables) can be both categorical and continuous. The RF algorithm is preferred because of varying reasons such as ease of use, strong generalization ability and rapidity. Although, the mathematical idea behind a decision tree is easy to grasp, the RF algorithm is regarded as a so-called black box model since the model consists of a large number of trees (Couronné et al. 2018).

The RF algorithm has two sources of randomization. The first source is due to the random selection of the splitting input feature in the "root node" or "parent node". Bootstrapping is the second source of randomness and a very popular method in the construction step of the RF. It is defined as dividing the training data into small sub-samples randomly and growing several trees with them (Hastie et al. 2009). However, in this study, no bootstrapping over samples was needed since the model accuracy is already high, and it is computationally demanding.

The RF algorithm builds f from a set of "base learners", $h_1(x)$, …,$h_j(x)$, which are then merged to form the "ensemble predictor" $f(x)$. For the regression problems, the final output is obtained by averaging the predictions from each tree in the model, as given in Eq. (3).

$$f(x) = \frac{1}{J} \sum_{j=1}^{J} \left(h_j(x)\right).$$

(3)

The error of the model result can be calculated by different criterions, but the mean squared error is mostly used. The model is trained until the mean squared error gets its minimum value. For the classification problems, the prediction is determined by a majority vote. The final output is calculated as seen in Eq. (4).

$$f(x) = \frac{1}{J} \arg\max \sum_{j=1}^{J} I\left(y = h_j(x)\right).$$

(4)

The RF model is trained until the splitting criterion, for example, misclassification error, Gini index or entropy, reaches its minimum. In the current model, the Gini index is implemented as splitting criterion, which is defined in Eq. (5). The input feature for splitting in the "root node" is selected randomly. Gini index measures the impurity in the "child node" corresponding to each input feature and selects the one with the smallest Gini index as the next splitting feature. Each node is split into two nodes until

the Gini index is zero. The nodes having a zero Gini index are called "terminal nodes".

$$\text{Gini index} = \sum_{k \neq k'}^{k} p_k \, p_{k'}; \, p_k = \frac{1}{n} \sum_{i=1}^{n} I(y_i = k), \qquad (5)$$

where y_i is the predicted value, k is the number of classes, p_k is the proportion of class k observations in the node (Cutler et al. 2012).

There are several hyperparameters that can be adjusted to build a successful RF model, such as depth of the trees (max_depth), number of trees (n_estimators) and number of variables (max_features). One of the most important hyperparameters is the depth of the trees in the RF model. The number of trees in the forest is another hyperparameter that needs to be tuned. The model accuracy becomes better while the number of trees is higher in the forest. However, the model training may take a much longer time with more trees (Agrawal 2021). The other important hyperparameters is the number of variables employed when looking for the best split. The highest accuracy is achieved with the following hyperparameters: [max_depth: 6, n_estimators: 118 and max_features: 2].

3.3 Performance Indicators

There are several error metrics to investigate the performance of a classification model. In this study, the predictive performance of the models is assessed by considering accuracy score, precision, recall and F-score, which are defined in Eqs. (6), (7), (8) and (9). These metrics are derived from the confusion matrix, visualizing the predictions with respect to the real classes. An example confusion matrix with two classes is shown in Table 3 and used for explanation. The positive class is represented by 1, while the negative class is represented by 0. Values in the diagonal are the number of correct predictions, while the rest are the number of incorrect predictions. In the metrics, the number of correctly predicted positive class examples (true positive), correctly predicted examples that do not belong to the class (true negative) and incorrectly predicted classes (false negative or false positive) are considered. Accuracy is defined as the overall success of the classification model. It is calculated by taking the ratio of correct predictions and total number of predictions. Precision is the percentage of correct positive predictions; that is, the proportion of positive identifications the have really been correct. Recall or sensitivity is the percentage of true positives that are successfully predicted. In other words, it explains what proportion of real positives

Table 3 Example confusion matrix with two classes

True class	Predicted class	
	Positive (1)	Negative (0)
Positive (1)	True positive (TP)	False negative (FN)
Negative (0)	False positive (FP)	True negative (TN)

have been correctly predicted. F-score is the weighted average of precision and recall. It can be used for the interpretation of precision and recall together (Sokolova and Lapalme 2009).

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FN} + \text{TP} + \text{TN}}, \tag{6}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{8}$$

$$F\text{-score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{9}$$

4 Results and Discussion

To measure the performance of LR and RF models, confusion matrices including six quality classes were used. In Figs. 6 and 7, confusion matrices for the training, testing and overall dataset for both models are shown. It can, as expected, be clearly seen that the RF model outperforms the LR model. For all data, LR model predicts 41 out of 47
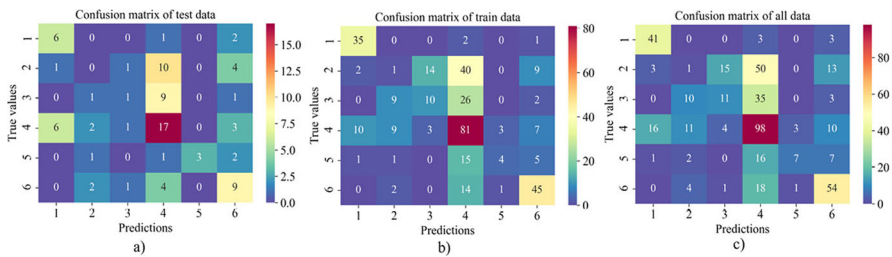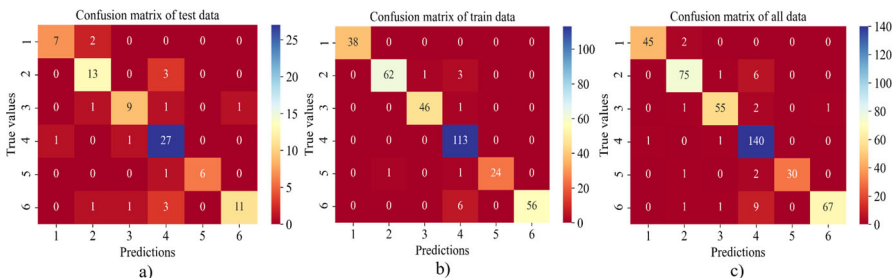

Fig. 6 Confusion matrices for the LR model

Fig. 7 Confusion matrices for the RF model

Table 4 Performance metrics of the models

Model		Accuracy	Precision	Recall	F-score
LR					
	Train	0.50	0.43	0.50	0.45
	Check	0.40	0.37	0.40	0.36
	All	0.48	0.42	0.48	0.43
RF					
	Train	0.96	0.98	0.96	0.96
	Check	0.82	0.83	0.82	0.82
	All	0.93	0.94	0.93	0.93

data correctly (true positive) for class 1. However, 20 data are labelled as class 1 (false positive), while three of them belong to class 2, 16 belong to class 4 and one belongs to class 5. Similarly, predictions for the rest of the classes are highly inaccurate, as seen in Fig. 6. On the other hand, the RF model labelled 45 out of 47 data correctly (true positive), and there is only one false positive prediction for class 1. The accuracy is quite high for the rest of the classes as well.

The performance of models was also examined during the training and testing stages to ensure whether they were overfitting or not. Table 4 represents the accuracy, precision, recall and F-score of the multiclass models utilized for the prediction of quality classes. For each model, overall precision, recall and F-score values were calculated by taking weighted average of each class as shown in Table 4. The tree-based ensemble classification (RF) model produced better predictions than the LR model.

The marble quality classes predicted by the RF model can suggestively be utilized as inputs to geostatistical methods for building a quality block model. As a result, SMU boundaries may be drawn based on the predicted marble quality rather than based on the procedure the mine uses today that are based on tonnage and geometrical constraints. In such a case, taking into account the predicted marble quality, the SMU boundaries will be chosen so that the quality variability within each SMU is lower (homogeneous quality) and higher to ease the blending process between the SMU variability.

Furthermore, the SMU boundaries defined by the suggested algorithm can also improve production efficiency by reducing the probability of high-quality marble being discarded or low-quality marble being mixed in the product.

5 Conclusions

MWD data-based marble quality class prediction models were developed via the use of different supervised learning type classification modelling methods, LR and RF. MWD parameters were selected based on statistical evaluations and used as independent

variables. The RF model provides higher accuracy, precision, recall and F-score than the LR model. The constructed RF model can be used as reliable tool for the preliminary prediction of quality class of marble.

The suggested RF model allows quality assessment and provides high-resolution quality data that can be used by geostatistical methods as valuable input to create a quality block model for efficient production and blending plans. The predicted classes can potentially be used in an indicator kriging approach, where each predicted class is transformed into a binary class to evaluate the portion of the blast or the SMU above some quality. Such block model may help the mine for a more accurate definition of SMU boundaries to reduce production and ore loss and to improve and strengthen the blending plans. This way, the productivity will enhance and contribute to SDG 12, the responsible consumption and production pattern.

As proved by the statistical performance indicators, the suggested models are reliable and accurate; however, it is always possible to improve model performance by adding new data representing a broader range of the geology in the studied mine. For other mines in a similar situation, the presented methodology can be followed to improve production efficiency and reduce production loss.

Data Availability Data sharing is not applicable. However, the code utilized in the training and testing stage of the models can be found at https://github.com/ozgeakyldz/Marble-quality-assessment-using-measurement-while-drilling-data.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Agrawal T (2021) Hyperparameter optimization in machine learning: make your machine learning and deep learning models more efficient. Apress, Berkeley
Atlas-Copco SMARTRIG™: Handbook (2019). http://exeldrilling.com.au/wp-content/uploads/SmartRig%20Handbook%20Low%20Res.pdf. Accessed 11 Jan 2019

Basarir H, Wesseloo J, Karrech A, Pasternak E, Dyskin A (2017) The use of soft computing methods for the prediction of rock properties based on measurement while drilling data. In: Deep mining 2017: eighth international conference on deep and high stress mining, Perth, WA, pp 537–551

Belyadi H, Haghigtat A (2021) Machine learning guide for oil and gas using python: a step-by-step breakdown with data, algorithms, codes, and applications. Gulf Professional Publishing, Cambridge

Breiman L (2001) Random forests. Mach Learn 45:5–32

Caté A, Perozzi L, Gloaguen E, Blouin M (2017) Machine learning as a tool for geologists. Lead Edge 36:215–219. https://doi.org/10.1190/tle36030215.1

Chang K-T, Chiang S-H, Hsu M-L (2007) Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression. Geomorphology 89:335–347. https://doi.org/10.1016/j.geomorph.2006.12.011

Couronné R, Probst P, Boulesteix A-L (2018) Random forest versus logistic regression: a large-scale benchmark experiment. BMC Bioinform 19:1–14

Cutler A, Cutler DR, Stevens JR (2012) Random forest. In: Zhang C, Ma Y (eds) Ensemble machine learning, methods and applications. Springer, New York, pp 157–176

Ellefmo SL, Aasly K (2019) InRec—increased recovery in the norwegian mining industry by implementing the geometallurgical concept. Research Council of Norway

Ghosh R, Schunnesson H, Gustafson A (2017) Monitoring of drill system behavior for water-powered in-the-hole (ITH) drilling. Minerals. https://doi.org/10.3390/min7070121

Van Guido R, Fred LD (2009) Python 3 reference manual. CreateSpace, Scotts Valley

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: Data mining, inference, and prediction. Springer, New York. https://doi.org/10.1007/978-0-387-84858-7

Hou SK, Liu YR, Li CY, Qin PX (2020) Dynamic prediction of rock mass classification in the tunnel construction process based on random forest algorithm and TBM in situ operation parameters. In: IOP conference series: earth and environmental science, p 570. https://doi.org/10.1088/1755-1315/570/5/052056

Kadkhodaie-Ilkhchi A, Monteiro ST, Ramos F, Hatherly P (2010) Rock recognition from MWD data: a comparative study of boosting, neural networks, and fuzzy logic. IEEE Geosci Remote Sens Lett 7:680–684. https://doi.org/10.1109/LGRS.2010.2046312

Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Kelley K, Hamrick J, Grout J, Corlay S, Ivanov P, Avila D, Abdalla S, Willing C, Jupyter Development Team (2016) Jupyter notebooks—a publishing format for reproducible computational workflows. In: Loizides F, Schmidt B (eds) Positioning and power in academic publishing: players. Agents and Agendas, IOS press, Amsterdam, pp 87–90

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, Montreal, Quebec, Canada, pp 1137–1145

Leighton J (1982) Development of a correlation between rotary drill performance and controlled blasting powder factors. M.Sc. Thesis,University of British Columbia

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830

Qi C, Chen Q, Sonny Kim S (2020) Integrated and intelligent design framework for cemented paste backfill: a combination of robust machine learning modelling and multi-objective optimization. Miner Eng. https://doi.org/10.1016/j.mineng.2020.106422

Raschka S (2016) Python machine learning: unlock deeper insights into machine learning with this vital guide to cutting-edge predictive analytics. Packt Publishing, Birmingham

Rodriguez JD, Perez A, Lozano JA (2010) Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans Pattern Anal Mach Intell 32:569–575. https://doi.org/10.1109/TPAMI.2009.187

Sadrossadat E, Basarir H, Karrech A, Elchalakani M (2022) An engineered ML model for prediction of the compressive strength of Eco-SCC based on type and proportions of materials. Clean Mater. https://doi.org/10.1016/j.clema.2022.100072

Schunnesson H (1998) Rock characterization using percussive drilling. Int J Rock Mech Min 35:711–725

Segui JB, Higgins M (2002) Blast design using measurement while drilling parameters. Fragblast Int J Blast Fragm 6:287–299

Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manag 45:427–437. https://doi.org/10.1016/j.ipm.2009.03.002

United Nations, Department of Economic and Social Affairs (2022). https://sdgs.un.org/goals/goal12 Accessed 15 Aug 2022

Vezhapparambu VS (2019) Statistical analysis of MWD data in a geometallurgical perspective: a case study on an industrial mineral deposit. Dissertation, Norwegian University of Science and Technology

Vezhapparambu VS, Ellefmo SL (2020) Estimating the blast sill thickness using changepoint analysis of MWD data. Int J Rock Mech Min. https://doi.org/10.1016/j.ijrmms.2020.104443

Vezhapparambu V, Eidsvik J, Ellefmo S (2018) Rock classification using multivariate analysis of measurement while drilling data: towards a better sampling strategy. Minerals 8:384. https://doi.org/10.3390/min8090384

Watne T (2001) Geological variation in marble deposits: Implication for the mining of raw material for ground calcium carbonate slurry products. Dissertation, Norwegian University of Science and Technology

Zhou J, Li X, Mitri HS (2016) Classification of rockburst in underground projects: comparison of ten supervised learning methods. J Comput Civ Eng. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000553