

Classificação de maciços rochosos em minas a céu aberto no Brasil: uma abordagem baseada em inteligência artificial e análise fatorial

Allan Erlikhman Medeiros Santos

Professor, Universidade Federal de Ouro Preto, Ouro Preto, Brasil, allan.santos@ufop.edu.br

Milene Sabino Lana

Engenheira de Minas, Doutora em Tecnologia Mineral, Belo Horizonte, Brasil, milene@ufop.edu.br

Denise de Fátima Santos da Silva

Engenheira Geóloga, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, denisefss@ufmg.br

Tatiana Barreto dos Santos

Professora, Universidade Federal de Ouro Preto, Ouro Preto, Brasil, tatiana.santos@ufop.edu.br

RESUMO: O propósito deste estudo é propor novos modelos para a classificação de maciços rochosos de minas a céu aberto localizadas no Brasil, fundamentados no sistema RMR. Os modelos propostos são baseados em técnicas de inteligência artificial. A utilização destes métodos quantitativos reduz a subjetividade dos sistemas propostos. As variáveis utilizadas foram a resistência e a alteração da rocha, condição e presença de água no maciço, e o grau de fraturamento do maciço rochoso. Assim, um número menor de variáveis foi utilizado, o que difere do sistema de classificação RMR. Os modelos foram treinados de modo supervisionado, de tal forma que as classes reais do RMR são utilizadas no treinamento, assim as técnicas identificam padrões para identificar estas classes; na fase de teste as classes são previstas e comparadas às reais, obtendo assim as acurácias dos modelos. Os valores médios de acurácia dos modelos treinados foram de 0,81; 0,89; 0,87 e 0,89 para naïve Bayes, random forest, redes neurais artificiais e máquinas de vetores de suporte, respectivamente. Estes resultados mostraram a possibilidade de definir as classes do RMR com um número menor de variáveis. O trabalho apresenta uma nova abordagem para lidar com sistemas de classificação de maciços rochosos, diminuindo a subjetividade, aumentando a seletividade dos parâmetros, reduzindo a dimensionalidade de bancos de dados geotécnicos e otimizando a aplicação de sistemas de classificação.

PALAVRAS-CHAVE: maciços rochosos, classificação geomecânica, inteligência artificial.

ABSTRACT: The purpose of this study is to propose new models for the classification of open-pit mine rock masses located in Brazil, based on the RMR system. The proposed models are based on artificial intelligence techniques. The use of these quantitative methods reduces the subjectivity of the proposed systems. The variables used were rock strength and alteration, rock mass condition and water presence, and rock mass fracturing degree. Thus, a smaller number of variables was used, which differs from the RMR classification system. The models were trained in a supervised manner, such that the actual classes of RMR are used in training, allowing the techniques to identify patterns to recognize these classes; during the testing phase, the classes are predicted and compared to the actual ones, thus obtaining the accuracies of the models. The average accuracy values of the trained models were 0.81, 0.89, 0.87, and 0.89 for naïve Bayes, random forest, artificial neural networks, and support vector machines, respectively. These results showed the possibility of defining RMR classes with a smaller number of variables. The work presents a new approach to deal with classification systems of rock masses, reducing subjectivity, increasing parameter selectivity, reducing the dimensionality of geotechnical databases, and optimizing the application of classification systems.

KEYWORDS: rock masses, geomechanical classification, artificial intelligence.

1 INTRODUÇÃO

Avaliar a condição dos maciços rochosos por meio da quantificação de sua qualidade constitui uma fase preliminar nos estudos de análise de estabilidade, permitindo estimativas da resistência do maciço rochoso e das descontinuidades, parâmetros de entrada usados em análises de estabilidade. No entanto, a maioria dos sistemas de classificação foi proposta para ambientes geológicos muito diferentes da realidade brasileira, tornando muitos parâmetros inadequados para os maciços rochosos brasileiros.

As limitações dos sistemas de classificação são destacadas por diversos autores, como Pells & Bertuzzi (2007), Potvin *et al.* (2012) e Ulusay *et al.* (1992). Pells & Bertuzzi (2007), Potvin *et al.* (2012) que apontam como fatores limitadores a incerteza na estimativa de parâmetros e julgamentos subjetivos. Ulusay *et al.* (1992) destacam limitações no contexto de rochas brandas e rochas altamente alteradas, relatando que o RMR (Rock Mass Rating), proposto por Bieniawski (1989), pode gerar resultados errôneos quando aplicado em rochas brandas, estratificadas e argilosas. Neste contexto, parâmetros como alteração dos maciços rochosos, são subestimados, com um peso pequeno, quando comparado aos demais parâmetros. Posto isso, muitos autores vêm propondo correções no sistema de pesos do RMR, como Santos *et al.* (2012), com foco na alteração dos maciços rochosos.

Assim, a presente pesquisa tem como objetivo a proposição de diferentes modelos de classificação para maciços rochosos brasileiros. A proposição do sistema foi feita com base na seleção das variáveis que mais influenciam a qualidade do maciço rochoso, a partir da determinação de fatores comuns utilizando a análise fatorial (Santos *et al.* 2021). Os sistemas de classificação foram desenvolvidos a partir das variáveis relativas aos fatores extraídos do banco de dados. As técnicas utilizadas para os modelos por aprendizagem supervisionada foram o Naïve Bayes, Random Forest, Redes Neurais Artificiais e as Máquinas de Vetores de Suporte.

A ciência de dados experimenta um crescimento significativo nas últimas décadas, especialmente com a popularização das técnicas de aprendizado de máquina, refletindo o poder dessas abordagens na resolução de uma variedade de problemas. No contexto dos sistemas de classificação de maciços rochosos, as classificações desenvolvidas, como o RMR (Bieniawski, 1989), estão intrinsecamente ligadas às condições tecnológicas de sua época, resultando em limitações. No entanto, é evidente que esses sistemas podem e devem ser aprimorados ao longo do tempo, especialmente à luz do desenvolvimento contínuo da ciência de dados, que permite uma melhor aquisição e processamento de dados. Esta pesquisa oferece uma nova abordagem para o desenvolvimento de classificações geomecânicas, utilizando métodos quantitativos, estatísticos multivariados e inteligência artificial, com o objetivo de reduzir as limitações inerentes aos métodos tradicionais, como subjetividade de parâmetros e incertezas nas estimativas, e abrangendo uma variedade de tipos de maciços rochosos.

2 MATERIAIS E MÉTODOS

2.1 Banco de dados

O banco de dados utilizado nesta pesquisa é resultado de uma compilação de diferentes bancos de dados de minas brasileiras. O banco é constituído de pontos amostrais com variáveis levantadas ao longo das faces de taludes, contendo informações sobre parâmetros relacionados à classificação RMR, são 3216 amostras. Assim, para a construção dos modelos de classificação por aprendizagem supervisionada, as classes do RMR para cada ponto foram utilizadas.

As variáveis utilizadas foram selecionadas em Santos *et al.* (2021), no qual foi aplicada a análise fatorial no mesmo banco de dados, com as variáveis relativas ao RMR. Os resultados obtidos pelas cargas fatoriais indicaram o Fator 1 representando as variáveis resistência e alteração da rocha, o Fator 3 representando as variáveis espaçamento e persistência das descontinuidades, e o Fator 2 representando as variáveis água e abertura das descontinuidades. Todas as variáveis estão diretamente relacionadas a qualidade dos maciços rochosos.

Santos *et al.* (2021) aplicando o princípio da interpretabilidade, aponta no Fator 1, um fator relacionado com a competência da rocha, uma vez que nele existe a resistência da rocha intacta e a alteração.

Estes parâmetros estão relacionados uma vez que a alteração no maciço impacta diretamente a resistência deste. O Fator 3 apresenta relação direta com o grau de fraturamento do maciço rochoso, uma vez que apresenta as variáveis espaçamento e persistência das descontinuidades. Por último, o Fator 2 está relacionado com as condições de percolação de água no maciço rochoso, uma vez que o fator relaciona a presença de água com a abertura das descontinuidades.

2.2 Partição do banco de dados

Para o treinamento dos modelos, o conjunto de dados foi particionado em amostra de treinamento e teste de forma aleatória. A amostra de treinamento consistiu em 2/3 do banco de dados e a amostra de teste com o 1/3 restante, conforme recomendação de Hair *et al.* (2009). Apenas a amostra de treino foi utilizada para treinamento do modelo.

Além disso, buscando os melhores ajustes e cenários possíveis foi aplicada a técnica de random subsampling. A random subsampling foi feita com trinta seleções, ou seja, trinta modelos foram treinados para cada algoritmo de aprendizado de máquina. Cada modelo consistia em uma amostra de treinamento diferente. O processo de random subsampling foi aplicado com o objetivo de verificar a reprodutibilidade dos resultados, estabilidade do modelo e estatísticas básicas das métricas de avaliação, como intervalos de confiança, média, mínimo, máximo e variabilidade dos resultados.

2.3 Aplicação das técnicas

Para treinamento dos modelos de aprendizado de máquina, foram utilizados pacotes específicos do R (R Core Team, 2018). O pacote *e1071* (Meyer *et al.*, 2012) foi utilizado para otimizar o treinamento para as técnicas de *Support Vector Machines* (SVM) e *Naive Bayes* (NB). O pacote *randomForest* (Liaw & Wiener, 2002) foi utilizado para treinar o modelo para a técnica de Random Forest (RF). O pacote *neuralnet* (Günther & Fritsch, 2010) foi usado para treinar o modelo para a técnica de Redes Neurais Artificiais (ANN). Para selecionar os parâmetros de cada técnica no treinamento dos modelos, foram aplicados ajustes e recomendações da literatura. Além disso, todos os parâmetros ajustados para o treinamento dos modelos foram selecionados com base no desempenho do modelo (*tunning*), ou seja, os parâmetros foram testados e selecionados de acordo com o melhor desempenho do modelo.

2.3.1 Naive Bayes (NB)

Para o treinamento com a técnica NB, o parâmetro de ajuste é a correção de Laplace, que não foi necessária para o presente estudo. A correção de Laplace não foi necessária devido as classes do RMR já estarem balanceadas, ou seja, o mesmo número de amostras em todas as classes. Além disso, o método de partição utilizado garante que a amostra de treinamento e teste estejam equilibradas em relação ao número de exemplos de cada classe de forma proporcional tanto na amostra de treinamento quanto na amostra de teste. Assim, a proporção de classes nas amostras de treinamento e teste foi equivalente. Posto isso, o parâmetro de correção de Laplace foi 0.

2.3.2 Random Forest (RF)

Para o treinamento do RF dois parâmetros são importantes no ajuste do modelo: o número de árvores (n_{tree}); e o número de variáveis amostradas aleatoriamente como candidatas em cada nó (m_{try}). O n_{tree} não deve ser um número muito pequeno, para garantir que cada variável seja prevista pelo menos algumas vezes. O m_{try} foi selecionado de acordo com as recomendações da literatura para problemas de classificação; $m_{try} = 2$. Para ajustar o n_{tree} , vários valores foram testados procurando o valor mínimo para a precisão máxima. Dentre os valores testados, observou-se que a partir de $n_{tree} = 100$ houve uma estabilização do valor de acurácia do modelo. Valores acima de 100 não adicionaram ganho no desempenho do modelo e aumentaram o custo computacional. Então $n_{tree} = 100$ foi assumido.

2.3.3 Redes Neurais Artificiais (ANN)

Para o treinamento da ANN o método utilizado foi a rede neural *perceptron* multicamadas com aprendizado supervisionado por meio do algoritmo *resiliente backpropagation*, com uma camada intermediária e função de ativação tangente hiperbólica. Vale ressaltar que diferentes funções de ativação foram testadas, sendo que a função tangente hiperbólica apresentou os melhores resultados em termos de acurácia. A função tangente hiperbólica aproxima-se da identidade, variando os valores de 1 a -1, proporcionando os melhores resultados para o treinamento da rede; portanto, foi a escolha adequada para a função de ativação.

Com relação à estrutura da rede neural, a camada de entrada possui 24 neurônios. Em relação aos princípios sobre o número de neurônios na camada intermediária e para evitar problemas de *underfitting* (modelo de aprendizado não consegue capturar a estrutura subjacente dos dados, resultando em um desempenho ruim tanto nos dados de treinamento quanto em dados de teste) e *overfitting* (modelo se ajusta excessivamente aos dados de treinamento, ou seja, o modelo captura não apenas os padrões gerais dos dados, mas também o ruído aleatório) 25 neurônios foram usados na camada oculta. A camada de saída da rede possui 5 neurônios, equivalentes às classes RMR existentes no banco de dados, classe I, classe II, classe III, classe IV e classe V.

2.3.4 Support Vector Machines (SVM)

Para a construção do modelo de classificação por SVM o *tunning* também foi aplicado, buscando otimizar a função do kernel e seus parâmetros considerando o melhor desempenho do modelo. As funções linear, polinomial, radial e sigmoide foram testadas para o kernel, com melhores resultados de acurácia para a função radial. Assim, para o treinamento do SVM foi utilizado o kernel radial, com $(1/2\sigma^2) = 1$.

A função kernel do SVM é uma função matemática que calcula o produto interno entre dois vetores de características, mas de forma implícita, sem precisar realmente calcular a transformação explícita dos dados para um espaço de maior dimensionalidade. Isso é útil porque permite que o SVM opere em espaços de características de alta dimensionalidade sem a necessidade de armazenar ou calcular explicitamente os vetores transformados.

Essencialmente, a função do kernel avalia a similaridade entre dois pontos no espaço original das características. Isso é crucial para a operação do SVM, pois permite que ele encontre um hiperplano de separação ótimo no espaço de características transformado, mesmo que os dados não sejam linearmente separáveis no espaço original. Alguns exemplos de funções kernel comuns incluem o kernel linear, o kernel polinomial e o kernel radial. Cada tipo de kernel tem suas próprias características e é escolhido com base na natureza dos dados e no problema específico a ser resolvido.

2.4 Validação dos modelos propostos

Considerando os modelos por aprendizagem supervisionada, trinta modelos foram treinados com diferentes amostras de treinamento e validados com suas respectivas amostras de teste. Assim, a partir dos 30 modelos treinados foram obtidos os valores médios para as métricas de avaliação do modelo. Posto isso, as métricas de acurácia com seus respectivos intervalos de confiança; índice Kappa; eficiência; precisão; e índice AUC foram aplicados e avaliados.

3 RESULTADOS E DISCUSSÕES

3.1 Treinamento dos modelos

A primeira verificação dos modelos treinados foi o desempenho nas amostras de treinamento e teste, para observar a existência de problemas de *overfitting* e *underfitting*. A métrica de desempenho do modelo para verificação de *overfitting* e *underfitting* foi a acurácia, ou seja, a probabilidade geral de sucesso, em cada treinamento realizado. Na Figura 1 estão apresentados os desempenhos nas amostras de treino e teste para cada técnica, ANN, RF, NB e SVM, respectivamente.

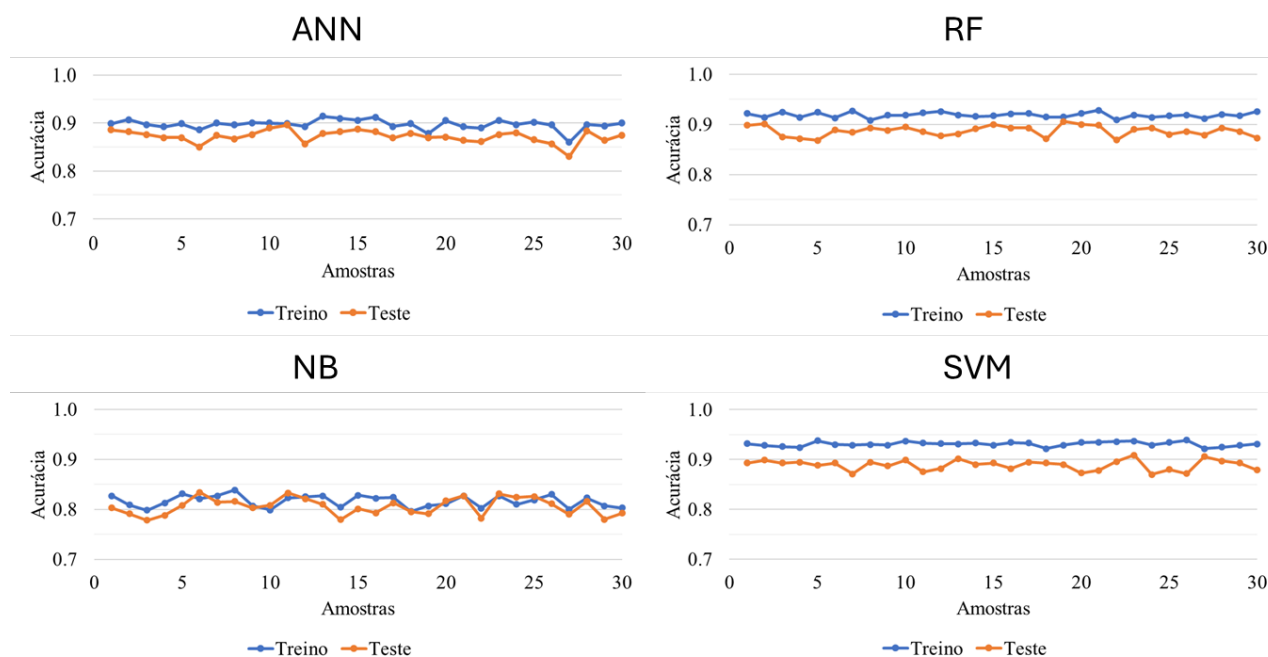


Figura 1. Desempenho nas amostras de treino e teste para os modelos.

O *underfitting* ocorre quando o desempenho do modelo não é satisfatório, considerando as amostras de treinamento e teste. Valores baixos significam que os modelos não são recomendados. Na Figura 1, todos os modelos apresentaram alta acurácia, sem *underfitting*, o que mostra a recomendação de seu uso. O *overfitting* ocorre quando o aprendizado do modelo é memorizado e funciona bem apenas na amostra de treinamento, com valores ineficientes na amostra de teste. O *overfitting* pode ser diagnosticado a partir de diferenças significativas entre o desempenho do modelo nas amostras de treinamento e teste. A partir da Figura 1, é possível observar uma semelhança entre os desempenhos no treinamento e teste, o que descarta problemas de *overfitting* nos modelos.

Os resultados apresentados na Figura 1 são importantes, pois além de evidenciar a ausência de problemas de *overfitting* e *underfitting*, é possível visualizar a adequação dos modelos solução de classificação dos maciços rochosos. Os algoritmos utilizados apresentam bons resultados para a classificação do maciço rochoso, principalmente considerando o uso de um número menor de variáveis quando comparado ao RMR tradicional. Isso mostra a sincronia do presente estudo, pois a redução do tamanho do banco de dados por meio da seleção das variáveis realmente relacionadas à qualidade do maciço rochoso é validada, pois é possível chegar às classificações por meio de diferentes algoritmos de aprendizado de máquina.

3.2 Validação dos modelos

Uma visualização das médias da acurácia dos modelos, que representa a probabilidade geral de sucesso, com seus respectivos intervalos de confiança superior e inferior está apresentada na Figura 2. Os valores médios de acurácia dos modelos foram 0.81, 0.89, 0.87 e 0.89 para NB, RF, ANN e SVM, respectivamente. A acurácia representa o número de respostas corretas em relação ao total de dados na amostra de teste. Dito isso, as taxas de erro são 0.19 (NB), 0.13 (ANN) e 0.11 (RF e SVM). Assim, é possível verificar que todos os modelos apresentam desempenho satisfatório, sendo o NB o menor valor.

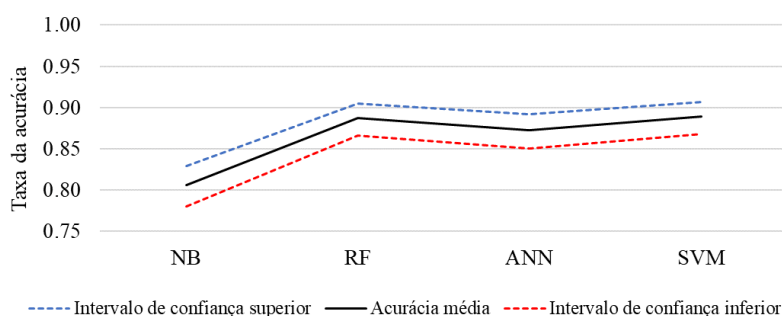


Figura 2. Acurácia média dos modelos com intervalos de confiança.

Dentre os algoritmos aplicados o que apresenta menor grau de complexidade, ou seja, o mais simples é o NB. Acredita-se que a simplicidade do algoritmo possa ter impactado no menor valor de acurácia. Embora, o valor de acurácia do NB seja o menor, 81%, isso não significa que modelo deva ser descartado. Este estudo apenas mostra que existem algoritmos com melhor taxa de desempenho. Outro ponto a ser citado é que a literatura indica que a premissa da interdependência entre os preditores pode impactar no resultado do NB, o que pode ser reafirmado aqui neste estudo. Ainda no contexto do NB, a influência da interdependência entre os preditores pode variar dependendo do contexto do problema e das características dos dados. Em alguns casos, a suposição de independência condicional feita pelo Naive Bayes pode ser violada se os preditores estiverem correlacionados entre si, levando a uma performance “pior” do modelo.

Apesar disso, ao se considerar a complexidade da classificação dos maciços rochosos, por se tratar de um problema de várias classes, com 5 classes distintas, o que eleva o grau de dificuldade do problema, o resultado de precisão de 0.81 para o NB é considerado satisfatório e pode ser aplicado.

Em relação aos maiores valores de acurácia, sejam eles RF, ANN e SVM, alguns pontos devem ser comentados. O modelo da ANN com o algoritmo Rprop e a aplicação de variáveis *dummies* potencializam os resultados, possibilitando melhor relação (interação) entre os neurônios nas camadas. A sofisticação do método SVM, principalmente devido ao aumento da dimensionalidade para ajuste do hiperplano durante o treinamento do algoritmo, implica em bons resultados. O RF funciona como um método de *ensemble* para árvores de decisão, ou seja, várias árvores são treinadas, o que contribui para a melhoria dos resultados. Nenhum dos modelos apresentou resultados que inviabilizassem seu uso, porém o RF, a ANN e a SVM foram destacados.

Para avaliar a reprodutibilidade dos resultados obtidos, foram analisados o índice Kappa e a área sob a curva ROC (AUC). As curvas ROC não foram plotadas por haver muita sobreposição entre as curvas do modelo e não apresentar a melhor visualização. Assim, optou-se pelos valores abaixo da curva em forma de tabela. Posto isso, a Tabela 1 apresenta os resultados do índice Kappa e AUC.

Tabela 1. Valores do Índice Kappa e AUC para os modelos.

Métrica/Modelo	NB	RF	ANN	SVM
Índice Kappa	0.76	0.86	0.84	0.86
AUC	0.95	0.96	0.96	0.96

O valor AUC aponta para o comportamento não aleatório dos modelos treinados. A avaliação do índice Kappa permite interpretar o NB como concordância substancial e outros como concordância quase perfeita. A partir dos resultados dos índices, verifica-se a reprodutibilidade dos modelos.

Os resultados da análise fatorial reduziram o número de variáveis, o que torna o banco de dados diferente das variáveis RMR originais, e mesmo assim, os algoritmos de aprendizado de máquina conseguiram obter as classes RMR com precisão. Assim, acredita-se na consistência e racionalidade da metodologia proposta. A metodologia proposta é uma abordagem que funciona para o problema de classificação de maciços rochosos.

Para o estudo em questão, se tratando de 5 classes de classificação a avaliação do desempenho dos modelos em relação a cada classe foi realizada utilizando as métricas de eficiência e precisão. Essas métricas são interessantes para observar o comportamento dos classificadores em relação a cada classe, identificando

os pontos fortes e fracos. A eficiência se reflete em uma acurácia balanceada, sendo uma média que entre sensibilidade e especificidade. A sensibilidade refere-se à capacidade do teste em identificar corretamente os casos positivos. Em outras palavras, é a proporção de verdadeiros positivos em relação ao total de casos positivos, expressa como uma porcentagem. Uma sensibilidade alta indica que o teste é bom em detectar casos positivos, minimizando os falsos negativos. Já a especificidade refere-se à capacidade do teste em identificar corretamente os casos negativos. É a proporção de verdadeiros negativos em relação ao total de casos negativos, expressa como uma porcentagem. Uma alta especificidade indica que o teste é bom em excluir casos negativos, minimizando os falsos positivos. A precisão mostra a taxa de acerto de previsão para cada classe. A Figura 3 apresenta os resultados para eficiência e precisão de cada classe para os modelos treinados.

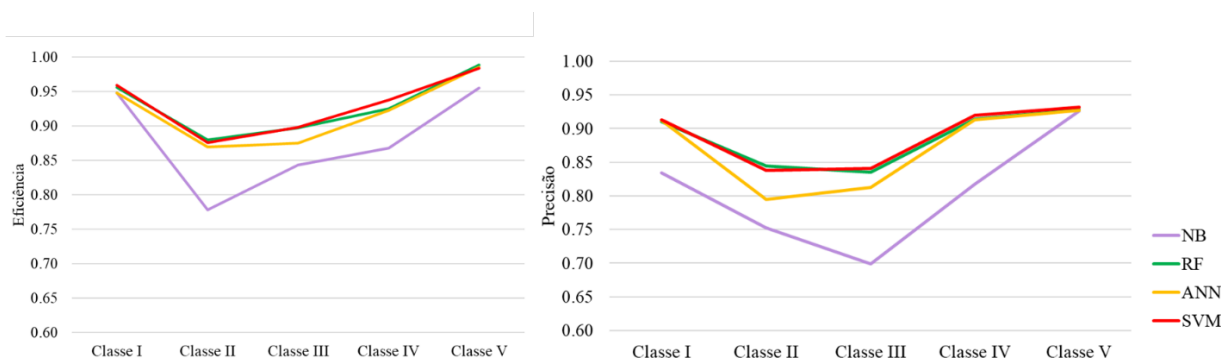


Figura 3. Eficiência das classes.

De acordo com a Figura 3, existe uma similaridade de desempenho entre os modelos RF, ANN e SVM. O NB apresentou, em média, valores menores quando comparado aos demais modelos, o que reflete no menor resultado de acurácia geral do modelo, conforme mostrado na Figura 2.

A característica geral observada nesse resultado é a predominância dos melhores valores para as métricas em classes extremas, por exemplo, o caso da Classe I e Classe V. Este resultado é esperado uma vez que essas classes são muito bem definidas, a Classe I é a rocha muito boa e a Classe V é a rocha muito pobre. Assim, os menores valores das métricas encontram-se nas Classes II, III e IV, o que pode ser interpretado como uma zona de transição entre as classes no problema em estudo. Acredita-se, que as características que definem as classes I e V são bem determinadas, particulares, para o maciço rochoso.

Em relação aos modelos com melhores taxas de desempenho, destaca-se o ANN, por ser um modelo com menor complexidade em termos de compreensão, visualização e concepção do algoritmo quando comparados aos modelos de RF e SVM. De fato, pela literatura, os modelos de ANN foram os primeiros modelos propostos para a inteligência artificial.

4 CONCLUSÕES

Os modelos propostos apresentaram bons índices de desempenho mostrando a adequabilidade das técnicas para a solução de problemas de classificação geomecânica. Além dos bons índices de desempenho destacam-se os resultados referentes a interpretação, os modelos conseguem mostrar as relações entre os parâmetros geomecânicos e a abordagem da classificação de maciços rochosos.

O banco de dados utilizado é extenso, resultante da compilação de várias minas brasileiras, com variáveis relacionadas ao sistema de classificação RMR. A seleção de variáveis na análise fatorial permitiu minimizar a subjetividade típica dos sistemas de classificação, pois seleciona as variáveis diretamente relacionadas com a qualidade do maciço rochoso. As variáveis não utilizadas, tais como preenchimento e rugosidade estão muito mais relacionadas à resistência ao cisalhamento das descontinuidades do que à qualidade do maciço propriamente dita. Além disso destacar que a alteração do maciço rochoso não é uma variável considerada no RMR (no RMR considera-se alteração das paredes da descontinuidade), no entanto a alteração do maciço rochoso é extremamente importante e determinante na qualidade dos maciços de rocha branda.

Neste ponto, consegue-se a redução da dimensionalidade do banco de dados. Portanto, todos os modelos de classificação criados foram feitos com um número menor de variáveis, sendo diferente do sistema de classificação RMR. Outro ponto vantajoso da metodologia é a abrangência do banco de dados, que inclui rochas de diferentes classes, desde maciços com baixa qualidade até alta qualidade, o que garantiu a aplicabilidade de modelos preditivos para qualquer tipo de maciço rochoso.

Os modelos treinados por aprendizagem supervisionada foram validados, provando ser precisos e eficazes. Os resultados mostram que os modelos de aprendizado de máquina forneceram bons modelos preditivos e permitiram mostrar que as variáveis selecionadas são realmente os parâmetros importantes para determinar a qualidade do maciço rochoso, visto que os modelos alcançam precisão nas validações. Os valores médios de acurácia dos modelos treinados foram de 0,81; 0,89; 0,87 e 0,89 para *Naive Bayes*, *Random Forest*, *Redes Neurais Artificiais*, *Support Vector Machine*, respectivamente. Estes resultados mostraram a possibilidade de alcançar as classes do RMR com um número menor de variáveis.

Neste contexto o trabalho apresenta novas abordagens para lidar com sistemas de classificação de maciços rochosos, diminuindo a subjetividade, aumentando a seletividade dos parâmetros, reduzindo a dimensionalidade de bancos de dados geotécnicos e otimizando a aplicação de sistemas de classificação. Assim, acredita-se que a seletividade dos parâmetros e a redução da dimensionalidade dos dados permitem a simplificação dos levantamentos geotécnicos sem influenciar na estimativa da qualidade do maciço rochoso.

AGRADECIMENTOS

Os autores do artigo agradecem. o apoio recebido do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Brasil (Projeto 402756/2023-1).

REFERÊNCIAS BIBLIOGRÁFICAS

- Bieniawski Z. T. (1989) *Engineering rock mass classifications: a complete manual for engineers and geologists in mining, civil, and petroleum engineering*. John Wiley & Sons, New York, pp. 251.
- Günther F., Fritsch S. (2010) *neuralnet: Training of neural networks*. The R Journal. 2(1). p.30-38.
- Hair, J. F. Anderson, R. E. Tatham, R. L. Black, W. C. (2009) *Análise Multivariada de Dados*. Tradução de A. S. Sant'anna e A. Cloves Neto, 6.ed. Porto Alegre: Bookman, p. 593.
- Liaw, A. Wiener, M. (2002) *Classification and regression by randomForest*. R News, 2(3), p.18-22.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F. (2012) *Package 'e1071'*.
- Pells, P. J. N., Bertuzzi, R. (2007) *Limitations of rock mass classification systems for tunnel support designs*. Tunnels Tunn Int, 1-11.
- Potvin, Y., Dight, P. M., J. Wesseloo. (2012) *Some pitfalls and misuses of rock mass classification systems for mine design*. Journal- South African Institute of Mining and Metallurgy 112, 697-702.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Santos, A. E. M. Lana, M. S. Martins, T. P. (2021) *Rock Mass Classification by Multivariate Statistical Techniques and Artificial Intelligence*. GEOTECHNICAL AND GEOLOGICAL ENGINEERING, v. 39, p. 2409-2430
- Santos, Y. L., Chaves, S., Paredes, J., Zenóbio, Â. (2012) *Proposta de índice de ajuste (Aw) utilizando o grau de intemperismo no sistema RMR (Rock Mass Rating) de classificação do maciço rochosos*. CBMINA 12.
- Ulusay, R., Ozkan, I., Unal, R. (1992) *Characterization of weak, stratified and clay bearing rock masses for engineering applications*. Proc Conference on Fractured and Jointed Rock Masses, Lake Tahoe, p. 233–240.