Submission code: ZGAVKNBJW6

FORECASTING OF DRILLING PARAMETERS THROUGH GEOPHYSICS

MAXIMILIANO SIMAO (Vale S.A.), Luciano Assis (Vale S.A.)

FORECASTING OF DRILLING PARAMETERS THROUGH GEOPHYSCS

Abstract Summary

The use of geophysical studies through electrical resistivity is widely used in mining, however, the monitoring of geotechnical structures using this method is still in rapid advancement. This work presents the approaches for data acquisition and processing, as well as the use of AI as an integrator and tool for detecting areas with geotechnical anomalies.

Introduction

Drilling data plays a fundamental role in geotechnical, geophysical and mineral studies. In engineering, this data is essential for the preparation of projects, as it enables the analysis of terrain characteristics, such as resistance, composition, stability and load capacity.

In mining, drilling is one of the main sources of direct research into the possible mineral deposit, being essential to quantify, qualify and geologically model the mineral deposit. This information is crucial for planning exploration efficiently, reducing uncertainty in the calculation of mineral resources and reserves and dimensioning of geotechnical projects.

Method

The use of Machine Learning (ML), which is a subfield of Artificial Intelligence (AI), emerges as a promising solution to these challenges (Gehring, 2017). These technologies allow the analysis of large volumes of data, identifying patterns and relationships that would not be easily perceived otherwise. By applying ML to drilling data, it is possible to create predictive models that estimate soil properties in areas where direct drilling has not been carried out.

For example, ML algorithms can be trained with historical drilling data and other parameters, including CPTU (Cone Penetration Test with Pore Pressure Measurement) and NSPT (Number of Strokes of SPT) test parameters. In the case of CPTU, the main parameters are:

• fs: lateral friction resistance

• qc: tip resistance

• u2: pore pressure

These parameters provide detailed information about the strength and properties of the soil at depth. NSPT is another important geotechnical parameter, which measures the strength of the soil by the number of blows required to penetrate a given distance. This data is crucial for predicting the bearing capacity and other characteristics of the soil.

Using ML can also help identify the most promising areas for future drilling campaigns, optimizing planning and reducing operational costs. This is especially useful in large-scale projects, where the logistics and resources required to conduct drilling can be significant.

Geophysical data is information obtained through non-invasive methods that allow the analysis of the physical properties of the subsoil. This data is essential for a preliminary understanding of the geological characteristics of an area, without the need for extensive drilling. Among the most common geophysical methods are resistivity surveys, which measure the electrical resistance of soil and rocks, providing information about the composition and structure of the subsoil. Resistivity can be measured aerially, using sensors mounted on aircraft, or by electrical survey, where electrodes are inserted into the soil to record resistivity variations along a profile.

The relationship between geophysical data and geotechnical data is close, as both provide complementary information about the subsoil. While geotechnical data are obtained directly through drilling and field tests, geophysical data offer a broader and more continuous view of the subsoil properties. This combination allows for a more complete and accurate analysis of the geological and geotechnical conditions of an area, facilitating the planning and execution of engineering and mining projects.

To consolidate this information into an analytical base, we use a grid technique for the x- and y-axis coordinates. This grid technique is an efficient way to deal with the heterogeneity of measurement points and allows us to create a consistent analytical base, containing only those regions of the grid that have all the necessary information.

The purpose of this work is to investigate the feasibility and effectiveness of using geophysical data, especially resistivity, to predict geotechnical drilling parameters, with a focus on improving the modeling and prediction process in geotechnology. The study seeks to evaluate the applicability of different artificial intelligence approaches, comparing simpler models with more complex models, and to verify the generalization capacity of the models created in different geographic areas. Based on these objectives, we seek to provide a deeper understanding of how advanced modeling techniques can be applied to optimize the analysis of geotechnical and geophysical data, contributing to more accurate and efficient decision-making in engineering and mining projects.

During the development of the project, a key step consisted of testing different modeling configurations and architectures to ensure the accuracy and robustness of the CPTU and NSPT parameter predictions. These tests were performed comprehensively, with applications and evaluations across all case studies, ensuring a consistent basis for comparative analysis across different methods and conditions.

Complementary features were developed to enrich the data and enhance model performance. These features were carefully designed to address different aspects of the data, such as local context, spatial and sequential variations, which are fundamental for robust and accurate modeling. The main features developed include: Resistivity Context, Depth, Sequential Difference.

Tests were performed with traditional models, initially using traditional machine learning methods, such as linear regression, decision trees, random forests and support vector machines (SVM). To speed up the selection of the most appropriate model for the problem, we used PyCaret, an open-source library in Python that facilitates the comparison and evaluation of different machine learning algorithms. This approach made it possible to efficiently identify which type of model presented the best performance for the proposed problem.

We also explored an autoregressive approach, since our goal is to predict the probing parameters at depth, where the prediction made for a point is used as a feature in the next prediction, as we can see in Figure 1, where we have this representation of the predictions feeding back into the model throughout the sequence, SOS (start of sequence) to EOS (end of sequence). This type of model was tested to evaluate its ability to capture temporal patterns and sequential dependencies in the data.

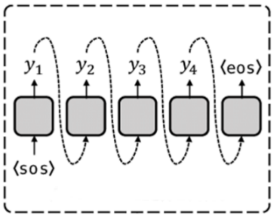

Figure 1: Schematic of how an autoregressive model works. The figure shows a sequence of blocks connected in a feedback process, where the first receives the SOS (Start of Sequence) to represent the beginning of the sequence, and the last block outputs.

We invested in seq2seq (sequence-to-sequence) architectures, a type of model, as shown in Figure 2, composed of two main modules, encoder and decoder, and designed to transform an input sequence (such as a sentence, a set of temporal measurements or sequential data). These models are widely used in problems where the order and context of the data are important.

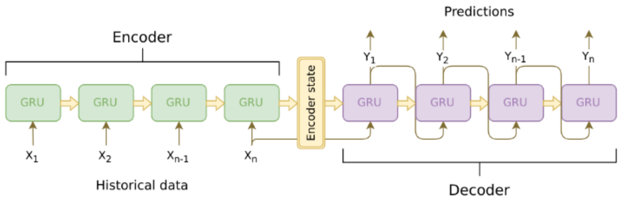

Figure 2: Illustrative diagram of how a seq2seq (sequence-to-sequence) model works. The figure shows an input sequence (Historical data) processed by the encoder layers, which transform the data into a compact intermediate representation (Encoder state). The decoder then uses this representation to generate the corresponding output sequence (Predictions), highlighting the model's ability to transform sequential data in a contextual and ordered way.

These architectures were implemented using recurrent neural networks (RNNs), a type of neural network that processes sequential data, such as time series or text, one step at a time, allowing previous information to influence subsequent steps. More advanced variants of RNNs, such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit), were used for their ability to handle longer dependencies in the data, meaning they can remember important information from earlier parts of the sequence for longer.

Additional tests included: Variation in the number of layers and neurons, Attention mechanisms, Inclusion of dropout, Convolutional layers.

Results

These choices presented previously allowed us to explore the best ways to model and analyze the data, leveraging the strength of different techniques to deal with complex sequential relationships and local patterns relevant to the problem.

As can be seen in figure 3, the model is able to separate rocky material and less resistant soil based on the Fs parameter of a region of interest, in which the model was trained in a neighboring area.

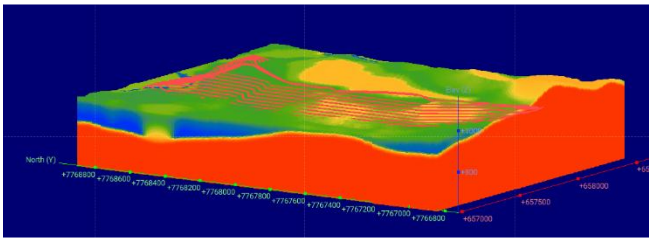

Figure 3: Representation of the Fs (lateral friction) of the area of interest.

Conclusions

In summary, the application of ML and AI in drilling data prediction represents an innovative and efficient approach to overcome the traditional challenges associated with obtaining and analyzing such data. By providing a powerful tool for predictive analysis, these technologies can significantly contribute to the success of geotechnical, geophysical and mineral projects, ensuring more accurate and cost-effective exploration.

These studies have shown that geophysical variables have great predictive power for geotechnical drilling parameters, regardless of the method used to collect the parameter. This ability makes geophysical variables valuable tools in planning and optimizing drilling campaigns.

Furthermore, it was found that, although more complex models, such as Seq2Seq with RNN, require greater computational cost, they offer considerably superior results compared to simpler models. This difference is justified by their ability to capture sequential and non-linear patterns more efficiently, demonstrating their suitability in scenarios that demand greater accuracy in predictions.

Finally, the analyses carried out reinforce the importance of hybrid and integrated approaches, using geophysical data as support in geotechnical studies, to increase the reliability of predictions and assist in decision-making in projects.

The next steps involve consolidating data from various geotechnical and geophysical areas into an integrated analytical base, with the aim of creating a more generalist and robust model. This approach will allow the use of a wider range of information to improve the accuracy of predictions and increase the applicability of the model in different contexts. In addition, the study of other predictor variables, such as seismic variables, which have great potential to complement and enrich existing models, will be explored. The inclusion of these new variables can bring significant benefits, expanding the prediction capacity and offering a more detailed and accurate understanding of subsurface conditions.

References

Gehring J., Auli M, Grangier D, Yarats D., and Dauphin Y. N. (2017). Convolutional sequence to sequence learning.

Simandoux, P. 1963. Mesures Dielectrique em Milloux Poreux, Application a Mesure dês Saturations em Eaux, Etude du Comportement dês Massifs Argileux: Ver. De lÍnstitut Français du Petrole, Supplementary Issue.