Application of Machine Learning Methods to Improve of the Roller Press Performance in the Pelletizing Process

Thiago Nicoli de Abreu[1,2][a], Andrea G. Campos Bianchi[3][b] and Saul Emanuel Delabrida Silva[3][c]
[1]*Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração (PROFICAM), Universidade Federal de Ouro Preto (UFOP) and Instituto Tecnológico Vale (ITV), Ouro Preto, Brazil*
[2]*Vale S.A., Vitória, Brazil*
[3]*Departamento de Computação (DECOM), Universidade Federal de Ouro Preto (UFOP), Ouro Preto, Brazil*

Keywords: Modeling, Machine Learning, Data Mining, Database, Roller Press, Pelletizing.

Abstract: In recent years, the technology of the roller press has become very useful in the pelletizing processes to comminute the pellet feed and increase the specific surface of the iron ore. It is known that the surface gain is directly related to the productivity and quality gains in the pelletizing process. In view of its importance, the increase in efficiency of the press becomes increasingly necessary, mainly due to its direct impact on the production chain. The large number of variables involved in its operation demonstrate that conventional methods and the knowledge of this process can be improved. For this, the work identifies the variables with the highest production in the specific surface gain, develops a classification model to determine rules of optimal operation settings and presents a model for the prediction of the specific surface variable, seeking gains in determining performance of this asset.

1 INTRODUCTION

At the beginning of the 1990s, a series of pelletizing industrial plants began to implement the Roll Press in industrial circuits for pressing iron ore and pellet feed. This was an important advance in the area of comminution and mineral processing (Barrios, 2014). One of the greatest benefits of the Roller Press in the pelletizing process is the increase in the specific surface gain, a property that directly contributes to the improvement of the physical and mechanical properties of the pellets and the quality of the finished product. When a material has a high specific surface, the greater the capillarity of the vessels, resulting in a more compact, more finished pellet with better mechanical resistance (Silva, 2008).

The inefficiency in the steps prior to the pressing process, possible problems with the adjustment of the operating parameters and the operational unavailability of this equipment (due to failures) are some of the factors that impact the performance of this equipment in terms of increasing the specific surface of the ore. The low efficiency of the press results in disturbances for the next stages of the process, mainly in the pelletizing stage (formation of pellets), causing losses in production and even plant shutdowns. A mining company may incur losses corresponding to millions of dollars in a few hours due to its production stoppage. Besides, an inappropriate manufactured product can lead to rework or non-compliance with the customer's quality requirements.

The presses operation is complex because it involves the combination of dozen variables. This leads to an exponential combination of setups for human decisions without computer support. In addition, the variations in the characteristics of the production process, in the previous and subsequent stages, physical and chemical characteristics of the ore and process restrictions introduce new production scenarios.

This work proposes an analysis of the roller press in a pelletizing plant, seeking to answer the following questions: what are the main factors that interfere directly or indirectly in the efficiency of this equipment? What are the possible setups (setting of

[a] https://orcid.org/0000-0001-5940-1632
[b] https://orcid.org/0000-0001-7949-1188
[c] https://orcid.org/0000-0002-8961-5313

Nicoli de Abreu, T., Bianchi, A. and Silva, S.
Application of Machine Learning Methods to Improve of the Roller Press Performance in the Pelletizing Process.
DOI: 10.5220/0010443906770684
In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 1*, pages 677-684
ISBN: 978-989-758-509-8

operating values) for the greatest gain in the process? The study was performed in a database installed in a mining company. The responses lead to improvement and increased productivity (production volume) and the specific surface of the ore.

Data mining techniques were used to identify the variables with the greatest influence on the gain of the press and a classification model to determine optimal operating setups of the roller press was developed. Our main findings and contribution are:

- Identify the most influential variables in the pressing process.
- Identify the rules for decision making using different values of optimal setups of operation of the roller press.
- Develop a classification model to predict the specific press surface.

The rest of this paper is organized as follows: Section 2 describes the main challenges for operating the roller press in the pelletizing process. Section 3 describes the work related to the use of computational intelligence techniques in this same process. Section 4 describes the materials and methods that were used to develop this work. Section 5 describes the results obtained and discussions carried out with the use of machine learning techniques to: determine the most relevant variables, obtain the rules of optimal setups and the prediction model of the specific press surface. Section 6 describes the conclusions after the analysis carried out.

2 PROBLEM STATEMENT

The Roller Press consists of two rollers rotating in opposite directions, one roller called a fixed roller, rotating on a fixed axis, and the other as a mobile roller, which rotates on a mobile axis that performs a translation movement towards the fixed roller through a hydraulic system (Figure 1). This system allows a variation of the specific compression force exerted on the bed of particles between the rollers.

The material feed is introduced in the opening between rollers, where the comminution of the particles occurs through the interparticle force breaking mechanism.

Due to the complexity of this equipment, variables such as working pressure of the rollers, clearance between the rollers, feed rate, level of the press rail, humidity of the pressed ore, speed of the rollers, torque and current of the motors, among others, need be parameterized (manipulated by the equipment operator) and monitored during its operation.

This scenario of multiple variables to be controlled emphasizes the need for a study of intelligent computational methods that can cooperate in the adjustments and in the best decision making in their operation. Therefore, the knowledge to operate the roller press is a major factor in fulfilling its objective, which is to increase the specific surface of iron ore. Dependence on human action just to control and maintain an operational standard in the face of all the complexity for analysis is practically impossible.

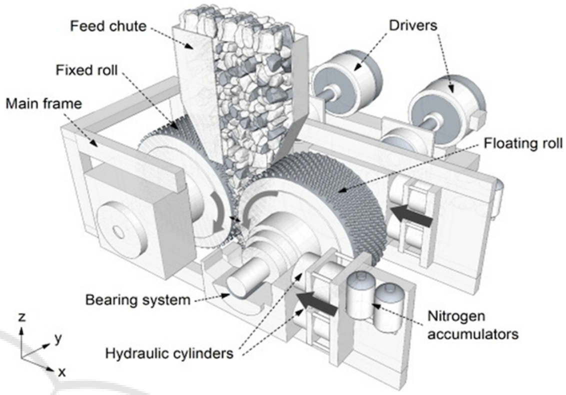

Figure 1: Diagram of the roller press. Source: Barrios (2014).

For the period of data evaluated for this work, the plant under analysis had 11.22 hours of stoppage and 273.37 hours of lost or reduced production. Considering the average production value of the plant at 700 t/h and the commercial price of a ton of iron ore as US$ 100.00/t, the estimated value of economic loss was approximately US$ 785,000.00. The main cause related to loss of production is related to the low values of specific surface after the pressing process.

Another factor of great attention is related to the measurement of the specific surface of the pressed pulp. Currently, this measurement is performed in the laboratory, in the interval of 4 hours, by means of manual collection performed in the post-circuit phase of pressing. This aspect causes a delay in the perception of the performance of this equipment, delaying corrections of possible failures and anomalies during the operation. Decision-making becomes late due to the lag in this measurement.

Thus, predicting the classification of this measurement online would be a more agile way of evaluating the performance of the equipment and the process in real time.

This work proposes an analysis of the roller press through artificial intelligence techniques to identify the variables of greatest influence to determine the gain of the press. Finally, it develops a classification model to determine optimal operating

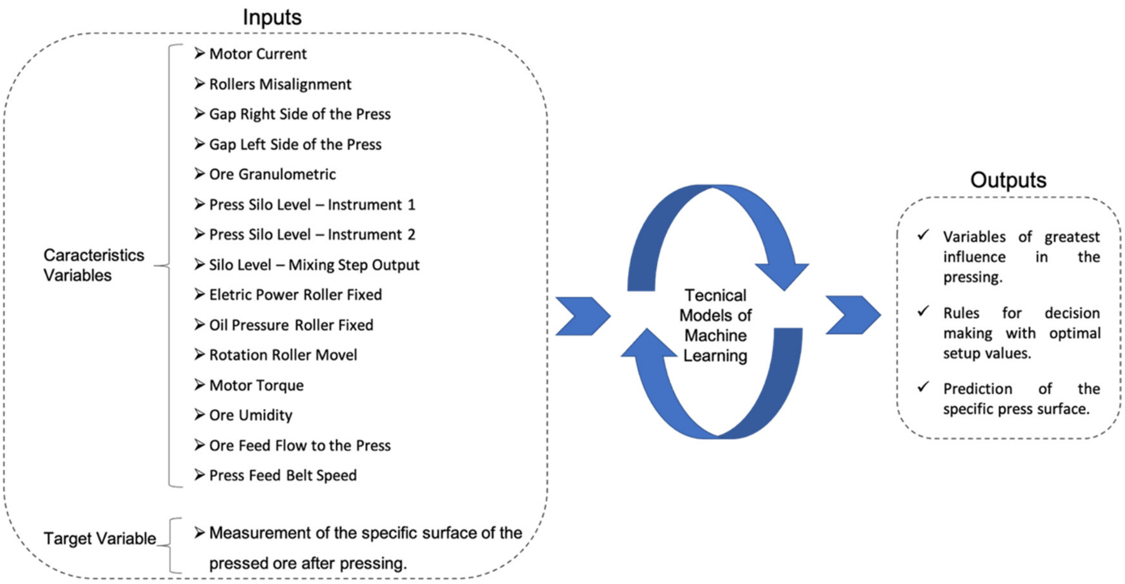

Figure 2: Overview of the concept applied in this project. Detailing of input variables, model learning and expected output results.

settings for the roller press. This model will be used to predict the specific surface variable of the pressed pulp, seeking gains in determining the performance of this asset.

3 RELATED WORK

Studies carried out in pelletizing processes, pressing area and applications of mathematical models in mining processes are essential as a reference and development of this work. The models developed until then did not consider the dynamics of the key parameters during the operation of the press to perform the forecast of the equipment performance variables.

Campos (Campos, 2018), for example, discusses the phenomenological mathematical model developed by Torres and Cassali (Torres & Cassali, 2009) capable of predicting the capacity, power and granulometric distribution of the product generated in the equipment and address a series of tests at different scales for pellet feed and iron ore pressing. Vyhmeister (Vyhmeister et al., 2019), on the other hand, presents a modeling study for roller press based on predictive control model (MPC), showing that the growing need for analysis in multivariable controls for complex processes requires increasingly robust and advanced strategies. Hasanzadeh and Farzanegan (Hasanzadeh & Farzanegan, 2011) apply a method based on genetic algorithms to estimate the parameters of mathematical models developed for roller presses, based on the model explained by Torres and Cassali (Torres & Cassali, 2009).

Tohry (Tohry, Chelgani, Matin, & Noormohammadi, 2020) presents a predictive model for power draw prediction based on operating parameters for an industrial ball mill. Still Tohry (Tohry, Yazdani, Hadavandi, Mahmudzadeh, & Chelgani, 2020) demonstrates in another work a robust artificial intelligence technique to fill gaps related to the modeling of energy consumption in high pressure grinding rolls on an industrial scale.

In this context, is evident that there is a need to develop accurate forecasting models to improve the operations of roller press. In addition, most representations of this equipment are based on steady state models for offline design and optimization, making them unsuitable for controlling online process and optimization.

Therefore, this work shows that there is a great potential yet to be explored in the use of the artificial intelligence area with the iron ore pressing processes in the industry.

4 METHODS AND MATERIALS

The data related to the pressing were extracted from the private base of the pelletizing system of a plant of a mining company collected at intervals of 10 minutes, over a period of 8 months. The resulting database is a matrix with 32,998 records (rows) of 16 variables (columns).

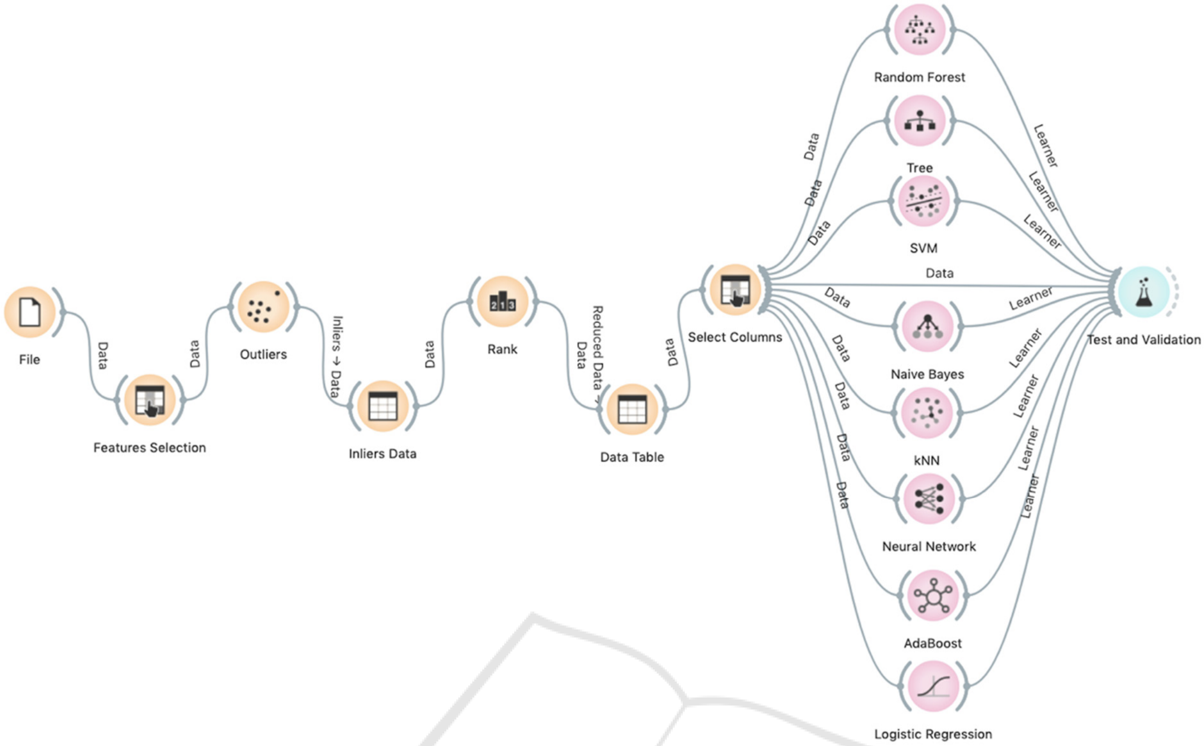

Figure 3: Modeling KDD processes in Orange software, showing all Orange processes and components used for modeling and project development.

The columns represent the press plant variables. The data was evaluated using the Knowledge Discovery in Databases (KDD) process. The Orange Data Mining[1] software was used for the development of this work, exploring the components of machine learning, data mining and libraries for classification and regression (Viterbo et al., 2016) and (Naik, 2016). Figure 3 shows the flow performed on this work.

The sequencing of the KDD processes was used to determine the ranking of the variables of greatest influence on the specific surface, to obtain the best classification model and the rules for optimal setup of the process variables.

The stage of data preparation is essential for this process, where tools were used for selection (pressing process rules and filters), outliers detection, discretization and data normalization. Still according to Shedroff (Shedroff, 1999), data "are the product of discovery, research, collection and creation. It is the raw material we find or create that we use to build our communications "and information" makes data meaningful to the public, because it requires the creation of relationships and standards between the data.

The most significant variables for the gain of the pressing process were identified using the Rank component of Orange with evaluation of the selection methods Information Gain, Information Gain Ratio, Gini, Anova, chi-square (X2), ReliefF and Fast Based Correlation Filter (FCBF).

The machine learning models tested were Decision Tree (Tree), Support Vector Machine (SVM), Naive Bayes, K-Nearest Neighbors (kNN), Neural Network, AdaBoost, Logistic Regression and Random Forest. The evaluation was performed according to the performance measures CA - classification accuracy and AUC - area under the ROC curve.

The cross-validation sampling method for learning, training and testing of the database was used during the validation of the methods, with 10 folds as a parameter.

5 RESULTS

The data were processed using the Outlier component of Orange to exclude outliers. For each of the 16

[1] More information at: https://orange.biolab.si

variables, the selection of the Inliers and Outliers data sets was obtained. Only Inliers data were used.

Comparing the variables “motor current” and “press trough level”, for example, it is observed that the Inliers data resulting from the execution of the One-Class SVM algorithm with non-linear Outliers detection kernel presents a considerable improvement for the data set when compared to the original data from the database (Figure 4). The same happens when comparing the variables of engine torque, roll misalignment and oil pressure, which confirms the reduction of the standard deviation of the values after the execution of the algorithm and the increase of the average value due to the concentration of the resulting data.

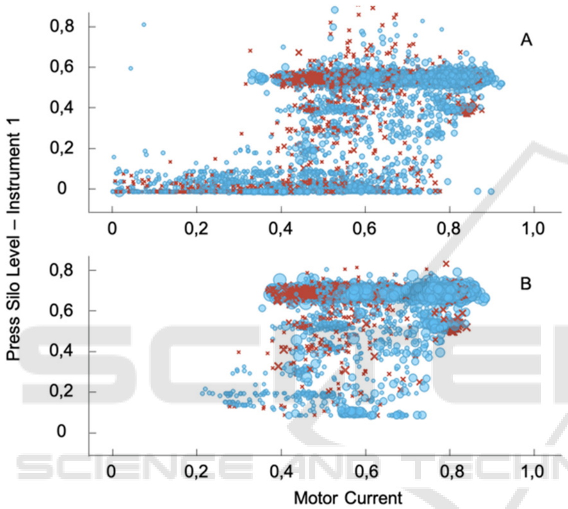

Figure 4: Exclusion of outliers of the roller press database. Values and density of the database of the variable "Press Silo Level – Instrument 1" in relation to the variable "Motor Current". Figure A: without data treatment to exclude the Outliers. Figure B: with data treatment to exclude the Outliers.

5.1 Identification of the Most Influential Variables in the Pressing Process

The ranking of the 8 most significant variables was performed using the Gain Ratio method. Figure 5 shows the order of significance of the variables.

These results show that the ore moisture is the variable of greatest interference in this process. According to Saramak and Kleiv (Saramak & Kleiv, 2013), there is an optimum humidity range for each granulometric distribution of the feed together, under a certain operational condition, which directly interferes with the ore comminution in the roller press.

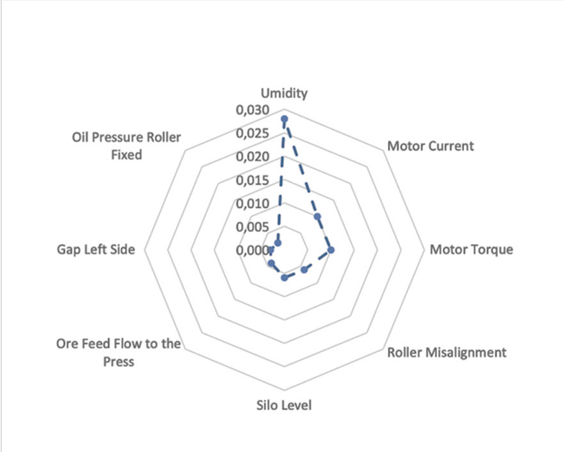

Figure 5: Ranking of the 8 most significant variables for the specific result of the roller press.

The importance of measuring motor current, ranked second by this method, opens a horizon for discussion since there is currently no direct reference to it, that is, it is not attributed the criticality of impact on the process. This may allow a more careful study of its impacts on the performance of the press, mainly because it is an electrical quantity with high dynamism for control and diagnosis. The engine torque is the main control variable in the pressing process used by the specialist system for optimizing the plant under analysis. Therefore, this result supports the validation of the control used in the production process.

The variable of misalignment of the rollers is related to the granulometric dispersion during the comminution process, which can occur in the application of a disproportionate force along the rollers due to misalignment. This fact directly interferes in the process of breaking the grains and consequently in the gain of specific surfaces of the press. Currently, this variable act only to protect the equipment. From this result, the influence of this measure can be discussed in more detail, considering the degree of significance found.

Another important variable is the silo level, resulting from the step prior to the pressing process, the filtration step. This level has a direct impact on the feed level of the press silo. According to Oliveira (Oliveira, 2015) the press feed chute cannot limit the flow of material to the compression zone, which directly impacts the capacity and performance of this equipment. Therefore, maintaining a constant level of supply shows a strong indication and relevance in the process.

Still according to Oliveira (Oliveira, 2015), specific capacity is one of the main aspects of press performance, with granulometric distribution as one of its factors of great importance. The operational opening (directly proportional to the gap between the rollers) has a high degree of influence on the particle size distribution, being identified in the gap distance variables on the left and right sides. The fact that they are ranked in different positions opens up a discussion about the possibility that the press is working in a possible state of misalignment.

The oil pressure that is applied to the mobile roller is another signaled variable with a high degree of importance, which is responsible for adjusting the pressure roller increase / decrease over the ore.

The rotation speed is currently used as the manipulated variable (output variable) of the existing PID controller (proportional-integral-derivative) to control the level of the press feed silo (due to the complexity of measuring this level, two instruments were installed to validate these values), which aims to maintain its stability for the ore pressing process. Both variables were not ranked as the most important for the process, demonstrating that possibly, due to the stability of the PID control, it was not possible to verify high interference of these variables from the direct gain of the press. The fact that both level measurements are ranked with the same weight shows that the model is consistent in the classification and ranking analysis. The result of classification of electrical power confirms the low importance of this quantity, as verified in the process today.

Finally, the result of classification of the ore granulometry variable does not seem to be in accordance with the expected, as this is of great importance in the performance of the process. According to Campos (T. M. Campos, Barrios, Bueno, & Tavares, 2017) a series of challenges involves the process of pressing the pellet feed from the point of view of the granulometry of the product and, in some cases, its high humidity. These facts make it difficult to increase the surface area of the material and cause greater difficulty in controlling the process. Therefore, it is concluded that the database for this variable must be verified and due to the fact that they are values obtained by measurements external to the process (information acquired at a very high frequency), a study in instrumentation is necessary to enable a measurement that allows data to be obtained in shorter and more assertive intervals of time.

Therefore, the results indicate consistency with the implications of the pressing process and the knowledge of the criticality classification of the variables can allow a more assertive and optimized decision making to improve this process. In addition, variables that, at first, were not pointed out as criticisms, they can be objects of study and analysis for the best performance of the press. This indicates that the application of ranking models in machine learning can cooperate to optimize the gain of the roller press.

5.2 Prediction of Classification of Specific Press Surface

Comparing the eight machine learning models evaluated, the Random Forest classifier model initially presented the best results in terms of precision, measured by the metrics CA = 0.875 and AUC = 0.913. Several tests to increase the number of trees in the Random Forest were performed and there is stability in the values of CA = 0.901 and AUC = 0.957 after reaching the number of fifty trees. These values show an excellent discrimination power, indicating great assertiveness in the prediction of the specific surface goal (> = 2.100g/cm³).

Table 1: Result of the evaluation of machine learning methods.

Model	AUC	CA
Random Forest	0.957	0.901
KNN	0.896	0.845
AdaBoost	0.826	0.859
Tree	0.807	0.860
Neural Network	0.804	0.791
Naive Bayes	0.652	0.727
Logistic Regression	0.647	0.720
SVM	0.443	0.626

The determination of the most significant variables and the subsequent model obtained by the Random Forest method in the pressing process demonstrate its effectiveness in the application of artificial intelligence methods in the industry. This result is in line with the result presented by Aldrich (Aldrich, 2020), showing this technique is constantly growing for use in analyzes of variable importance in a wide range of technical disciplines, including the mineral processing industries, such as in comminution processes, object of this study. Still according to Tohry (Tohry, Chelgani, et al., 2020), the use of the Random Forest model can greatly improve the prediction of the ball mill`s energy consumption in mineral processing plants, reinforcing the wide applicability of the techniques used in this work.

5.3 Validation of the Classification Model of the Specific Press Surface

To validate the Random Forest model for classification of the specific press surface, a new database, with 1015 records, was used. The overlap between the real and predicted values by the model results in 93.69% of assertiveness.

Therefore, the model achieves the utilization objective for a prediction of the press performance in the process with high assertiveness, enabling its application for measurement and adjustments in a more agile way and, consequently, the improvement for greater operational efficiency (Figure 6).

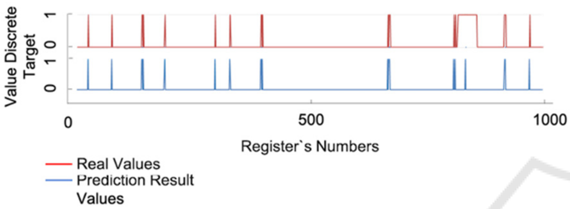

Figure 6: Comparison of real data (red) with the result obtained by the Random Forest prediction model (blue).

5.4 Identification of Optimal Setup Rules for Press Operation

The results that establish several rules to reach the target of the specific surface value (values> = 2.100 g/cm³) were obtained from the Orange Pythagorean Forest and Tree Viewer. Figure 7 shows 8 of the 50 trees generated by the model.

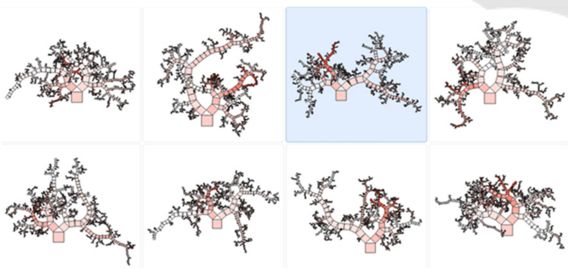

Figure 7: Pythagorean trees from the Random Forest model. The tree selected for extracting the information in the Tree Viewer component is highlighted, with the darker colors being the branches that guide to reach the goal of the prediction model.

A part of tree with the variables used for decision making, containing the varied and conclusive results to achieve the model's objective, is shown in Figure 8.

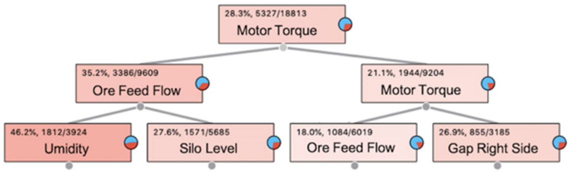

Figure 8: Decision tree levels from the Random Forest model. Detail of part of the decision tree, three depth levels, obtained from the result of the Random Forest model. Highlighted is one of the final nodes to achieve the "target" for the specific surface value. The variables belonging to each square in the figure demonstrate the variables that are used for decision making to achieve the goal of the model.

The 50 trees generated produce numerous decision rules, varying the number of rules for each of the trees. The Random Forest model creates several smaller decision trees by selecting random subsets of the characteristics in order to form a forest of trees that make up the global solution. One of the rules obtained from the example of one of the small decision trees comprises 7 of the 8 characteristics mapped in the model (Figure 9).

This result shows the values to be evaluated during the operation of the press, allowing to know some of the best setups for each verified condition. A priori, the rule extracted from the model may allow a control of the oil pressure variable when the roll misalignments greater than 0.13 mm and less than 0.70 mm occur or when the gap on the left side of the press exceeds the 5.44 mm value. So, for both conditions above, the oil pressure increase must be carried out until reaching a value greater than 94 kgf/cm² conditioned to the increase in torque (greater than 82.01% and less than 82.45%) and motor current (greater than 179.17 A), this when the ore feed flow for the pressing process is greater than 604 t/h and less than 633.85 t/h.

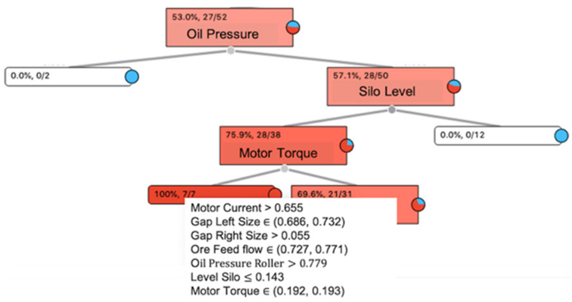

Figure 9: Result of decision tree obtained by the Random Forest model. Detail of a small tree with highlighted window shows one of the rules for reaching the goal of the model, showing the values (normalized) of each variable to be controlled to obtain the expected result for the process.

Therefore, from the discussion of this rule, one realizes the power of analysis that is provided by the machine learning model. The use of these models demonstrates the ability to make decisions in the face of varied process conditions and the correlation between the most significant variables, allowing gains with adjustments that drive the optimization of the expected result of the press.

6 CONCLUSIONS

The results of this work make it possible to speed up the predictive analysis of the performance of the roller press, automating the correlation of information from the various available systems and enabling the diagnosis of the press performance in real time, meaning a great advance since currently this performance needs an analysis laboratory with results available only in an interval of 4 hours.

In addition, it shows effective results of a multivariate analysis, contrasting the human limitation for the evaluation of numerous parameters. Thus, this work allows the decision making of the technical and operational team to be strengthened in order to support the challenge of reducing costs and increasing revenue and quality of the production process.

The applicability in the industry as well as its scalability are highly possible, since the possibility of implantation can be applied and customized for other existing roller presses, for the other different equipment in the pelletizing process (such as ball mill, filters, pelletizing discs and others) and even different processes, as long as they are evaluated for each need and peculiarity.

Besides that, the prediction of the process performance can open a wide discussion and possibility of study for the prediction of the useful life of this equipment adopting the various machine learning techniques.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001, the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), the Fundação De Amparo a Pesquisa Do Estado De Minas Gerais - FAPEMIG grant code APQ-01331-18, the Instituto Tecnológico Vale (ITV), the Universidade Federal de Ouro Preto (UFOP) and Vale S.A.

REFERENCES

Aldrich, C. (2020). Process Variable Importance Analysis by Use of Random Forests in a Shapley Regression Framework. *Minerals 10, 5*. doi:https://doi.org/10.3390/min10050420

Barrios, G., Tavares, M., Pérez-Prim, J. (2014). *High pressure grinding rolls simulation using the discrete element method dynamic coupling interface*. Paper presented at the XXVII International Mineral Processing Congress, Santiago, Chile.

Campos, T. M., Barrios, G. K. P., Bueno, G., & Tavares, L. M. (2017). Desafios na modelagem da capacidade e potência consumida da prensa de rolos. XXVII Encontro Nacional de Tratamento de Minérios e Metalurgia Extrativa, 1743-1753.

Campos, T. M. (2018). *Modelagem matemática da prensa de rolos aplicada à cominuição de minério de ferro.* Universidade Federal do Rio de Janeiro,

Hasanzadeh, V., & Farzanegan, A. (2011). Robust HPGR model calibration using genetic algorithms. *Minerals Engineering, 24*, 424-432. doi:10.1016/j.mineng.2010.12.004.

Naik, A., Samant, L. (2016). Correlation Review of Classification Algorithm Using Data Mining Tool: WEKA, Rapidminer, Tanagra, Orange and Knime. *Procedia Computer Science, 85*, 662-668. doi:10.1016

Oliveira, R. N. M. d. (2015). *Performance analysis of the HRCTM HPGR in pilot plant.* (Master). Universidade de São Paulo, São Paulo.

Saramak, D., & Kleiv, R. A. (2013). The effect of feed moisture on the comminution efficiency of HPGR circuits. *Minerals Engineering, 43-44*, 105-106.

Shedroff, N. (1999). Information Interaction Design: A Unified Field Theory in Design. In R. JACOBSON (Ed.), *Information Design* (pp. 267-292): The MIT Press.

Silva, L. (2008). *Influência da umidade no processo de pelotização.* Faculdade do Centro Leste – UCL.

Tohry, A., Chelgani, S. C., Matin, S. S., & Noormohammadi, M. (2020). Power-draw prediction by random forest based on operating parameters for an industrial ball mill. *Advanced Powder Technology,, 31*, 967-972.

Tohry, A., Yazdani, S., Hadavandi, E., Mahmudzadeh, E., & Chelgani, S. C. (2020). Advanced modeling of HPGR power consumption based on operational parameters by BNN: A "Conscious-Lab" development. *Powder Technology, 381*. doi:10.1016/j.powtec.2020.12.018

Torres, M., & Cassali, A. (2009). A novel approach for the modelling of high-pressure grinding rolls. *Minerals Engineering, 22*, 1137–1146.

Viterbo, J., Boscarioli, C., Bernardini, F., & Teixeira, M. (2016). Avaliação de Ferramentas de Apoio ao Ensino de Técnicas de Mineração de Dados em Cursos de Graduação. In Anais do XXIV Workshop sobre Educação em Computação, (pp. 11-20). Porto Alegre: SBC. doi:10.5753/wei.2016.9644.

Vyhmeister, E., Reyes-Bozo, L., Rodriguez-Maecker, R., Fúnez-Guerra, C., Cepeda-Vaca, F., & Valdés-González, H. c. (2019). Modeling and energy-based model predictive control of high pressure grinding roll. *Minerals Engineering, 134*.