APPLIED MULTIVARIATE ANALYSIS FOR SINTER FEO PREDICTION*

Paulo Lopes[1]
Alei L. A. Domingues[2]
Bianca Nakandakari[3]

Abstract

This work is one of the results of a successful proof of concept, with the objective to analyze and model the iron ore sintering process on a pilot plant scale, aiming at predicting the final FeO content in the sinter, as a function of the ore blend, fuel, fluxes and other process parameters. The model was developed with real data from a sintering pilot plant, considering around 300 tests with different iron ore mixtures. Multivariate analysis and machine learning techniques were applied, and a final mathematical model with R^2 greater than 0.92 was obtained, confirming the strength of the proposed methodology.

Keywords: Machine Learning; Sintering process; Iron ore, Sinter feed.

1 Co-founder - Beyond Mining - paulo@beyondmining.tech
2 Technical Marketing - Vale - alei.domingues@vale.com
3 Co-founder - Beyond Mining - bianca@beyondmining.tech

1 INTRODUCTION

The iron ore sintering process is a critical step in the production of iron and steel. It involves the agglomeration and heating of iron ore fines, fluxes, and coke breeze into a porous mass that is then loaded into a blast furnace. The efficiency and quality of the sintering process have a significant impact on the productivity and profitability of steelmakers. As such, there is a growing interest in applying machine learning techniques to optimize various aspects of the sintering process.
FeO in the iron ore sinter is formed during the sintering firing process, through oxidation/reduction reactions of iron oxides. Several factors such as mineralogical composition, temperature, fuel, and permeability can be related to the formation of FeO during burning. The FeO content of the sinter is directly related to its quality, as it affects physical strength and metallurgical behavior during the reduction in A&F.
Given this, measuring the FeO of the sinter right after it is produced is essential to adjust the process. Currently, the main way to make this measurement is by wet chemical tests, which require about 2 to 3 days to be completed. Wet tests still require samples to be prepared (sprayed) before being sent to the laboratory. Alternative solutions able to measure FeO directly or indirectly in sinters produced in the pilot sintering machine are useful and desirable in this context.

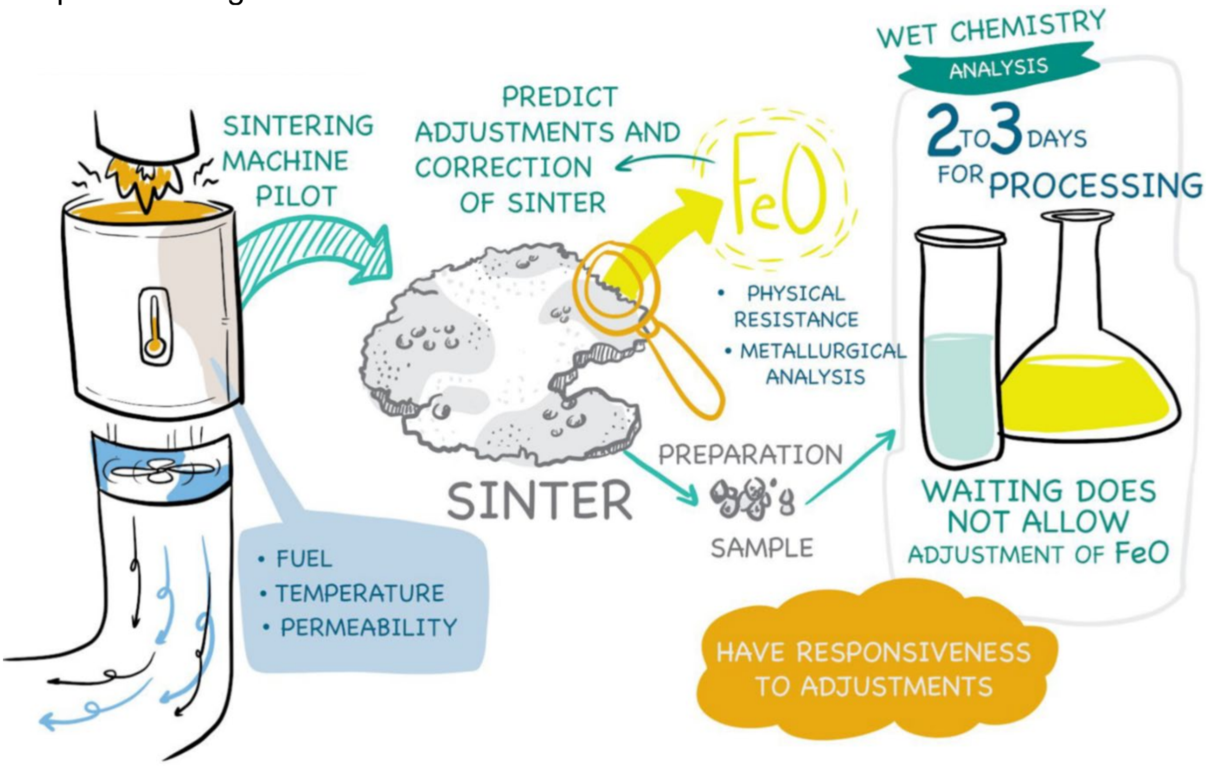

Figure 1. Infographic summarizing the challenge (Mining Hub's Challenge e-book)

The main objective of the project was the development of a predictive mathematical model for the final FeO content of the sinter after the burn process, as a function of the process parameters and the quality of the ore. A secondary objective is to achieve a model that is not too simple, but also not too complex, implying overfit and compromising its extrapolation and generalization.
There are many studies applying machine learning techniques in the sintering process, including ANNs, decision tree regression, SVMs, and fuzzy logic, that can be effectively applied to optimize and predict the sintering process's various parameters, such as sinter yield, sinter strength, and energy consumption.

These proposed approaches can provide accurate predictions and significantly improve the sintering efficiency, product quality, and energy consumption. However, there is still room for further research to optimize the sintering process's multiple objectives, such as improving sinter yield and sinter strength while reducing energy consumption. Furthermore, the development of intelligent control systems using machine learning techniques can significantly improve the process's stability and efficiency. The application of machine learning techniques in the iron ore sintering process shows promising results and can potentially provide significant benefits to the industry.

2 LITERATURE REVIEW

2.1 Machine learning on FeO prediction

There are many attempts to apply machine learning techniques in the iron ore sintering process, with a vast variety of research methods, findings, and implications of these studies to gain insights into the current state of research in this field.
Kwang et al. (2000) utilized diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) to investigate various multivariate calibration methods for accurately quantifying the FeO content in sinter ores. The calibration techniques explored include ridge regression with variable selection, principal component regression, ridge principal component regression, and partial least square regression, utilizing both linear and nonlinear mapping through neural networks. Before modeling, the spectral data underwent preprocessing steps such as signal correction and scaling. Cross-validation was employed to determine the optimal biasing parameter for ridge-related regression and the optimal number of principal components (or latent variables) for component-related modeling. We also considered the possibility of reducing the number of variables in the models while maintaining prediction accuracy to propose a final prediction model. Based on our analysis, we recommend utilizing component-related regressions on auto-scaled orthogonal signal correction for the quantitative determination of FeO content in sinter ores, as these calibration methods demonstrate appropriate performance.
Ziang et al. (2020) evaluated sinter quality relies significantly on the FeO content present in it. Due to the extreme temperatures and challenging conditions involved, real-time online detection of FeO content is not feasible. In order to address this issue, a novel approach that combines heat transfer mechanisms with data-driven models has been proposed to enable online prediction of FeO content. Initially, a model based on temperature distribution mechanisms is developed to represent the sintering process. This model categorizes the sinter into three groups based on the maximum temperature it reaches. Subsequently, three distinct long short-term memory models are created to predict the FeO content under different conditions. The effectiveness and practicality of this proposed model are confirmed through its application in a sintering plant, where the predicted results furnish dependable FeO content information for the sintering site.

2.2 Multivariate analysis

Principal Component Analysis (PCA) is a widely used technique for dimensionality reduction and data visualization. It is a linear transformation that maps the original data onto a new coordinate system such that the variance of the data in each

dimension is maximized. In recent years, there has been growing interest in developing more efficient and accurate PCA algorithms for various applications in machine learning (Jolliffe, I., Cadima, J., & Sandoval, J. 2016).
Linear Discriminant Analysis (LDA) is a widely used supervised classification technique that has been applied to a variety of machine learning applications, including image recognition, bioinformatics, and finance. In LDA, the goal is to find a linear projection of the data that maximizes the separation between classes while preserving the within-class structure of the data. In recent years, there has been a growing interest in developing more efficient and accurate LDA algorithms for various applications in machine learning (Sugiyama, M., & Borgwardt, K. M. 2015).
Multivariate regression is a popular statistical technique that models the relationship between multiple independent variables and a dependent variable. In the context of machine learning, multivariate regression has been applied to a wide range of applications, including finance, bioinformatics, and natural language processing. In recent years, there has been significant progress in developing new and more efficient algorithms for multivariate regression (Xiong, S., & Pan, W. 2018).

2.3 Data augmentation

Data augmentation is a technique that has been increasingly adopted in the field of machine learning as it allows us to generate more training data from existing datasets by applying various transformations. The use of data augmentation techniques has been shown to be effective in improving the performance of machine learning models, particularly in computer vision applications. In recent years, there has been growing interest in using stochastic simulation-based data augmentation techniques to further improve model performance (Zhang et al. 2018).

3 DATA AND METHODS

3.1 Methodology

The PCA, LDA and multiple regression techniques were applied in sequence, both for the raw dataset and for the augmented dataset.

3.2 Data collection

The data used for training the model was obtained from tests on a pilot sintering machine at the Ferrous Technology Center (Figure 2). Results from 150 firings were used, totaling 300 data about raw material properties and process parameters.

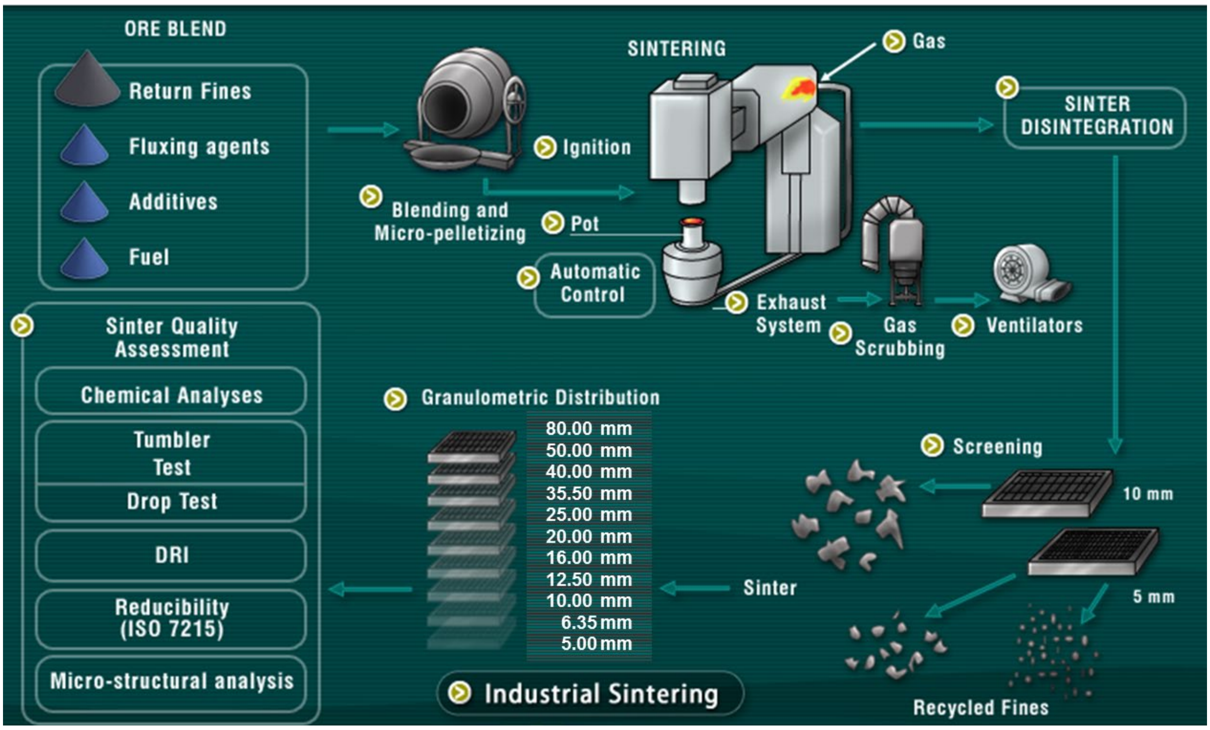

Figure 2. Sintering pot tests experimental procedures (Vale internal collection)

Among the 150 burns selected for this project, mixtures composed of ores from all over the world were selected, including mixtures with Brazilian, Australian, Canadian and Chinese ores. The binary basicities of these mixtures range from 1.7 to 2.2.

Table 1. Description and details of model variables

	Variable	**Description**
Quality parameters from ore blend	SiO2 ores mix	Iron ore blend SiO2 from chemistry analysis
	Al2O3 ores mix	Iron ore blend Al2O3 from chemistry analysis
	Fe ores mix	Iron ore blend FeT from chemistry analysis
	PF ores mix	Iron ore blend LOI from chemistry analysis
	FeO ores mix	Iron ore blend FeO from chemistry analysis
Quality parameters from the mixture	CaO sinter mix	All raw materials (iron ore+fluxes+fuel+burnt lime) CaO calculated from sinter mix
	SiO2 sinter mix	All raw materials (iron ore+fluxes+fuel+burnt lime) SiO2 calculated from sinter mix
	Al2O3 sinter mix	All raw materials (iron ore+fluxes+fuel+burnt lime) Al2O3 calculated from sinter mix
	MgO sinter mix	All raw materials (iron ore+fluxes+fuel+burnt lime) MgO calculated from sinter mix
	Fe sinter mix	All raw materials (iron ore+fluxes+fuel+burnt lime) FeT calculated from sinter mix
	b2 sinter mix	All raw materials (iron ore+fluxes+fuel+burnt lime) binary basicity (SiO2/CaO) calculated from sinter mix
Process parameters	Suction pressure mmH2O	Set up of suction pressure for each burn
	% Moisture	% moisture of granulation
	% Fuel.	% fuel charged on the mixture
	% Lime MP	% burnt lime charged on the mixture
	Tempo sint. (min)	Sintering time for each burn
	Temp máxima T1 (ºC)	Maximum temperature during ignition process
	Temp máxima T3 (ºC)	Maximum temperature at wind box

Table 2. Descriptive statistics of the dataset used

Variable	**Mean**	**StDev**	**Minimum**	**Maximum**
SiO2 ores mix	4.87	0.45	4.48	6.67
Al2O3 ores mix	1.36	0.2	1	1.85
Fe ores mix	61.53	0.96	59.25	63.08
PF ores mix	4.94	1.47	2.03	8.07
FeO ores mix	0.26	0.15	0.04	0.71
CaO sinter mix	10.81	0.61	9.64	12.8

Variable	Mean	StDev	Minimum	Maximum
SiO2 sinter mix	5.66	0.36	5.35	7.11
Al2O3 sinter mix	1.54	0.15	1.28	1.88
MgO sinter mix	1.41	0.31	1.1	1.9
Fe sinter mix	55.8	0.75	52.84	56.98
b2 sinter mix	1.91	0.09	1.73	2
Suction pressure mmH2O	1513	61.5	1500	1800
% Moisture	7.38	0.32	6.4	8.1
% Fuel.	3.9	0.22	3.45	4.4
% Lime MP	2.69	1.06	1.41	8.88
Tempo sint. (min)	28.5	4.2	19.95	38.32
Temp máxima T1 (ºC)	1162	20.2	1124	1200
Temp máxima T3 (ºC)	430.38	29.2	355	502

This number of samples increased to 95473 after applying the data augmentation methodology. Table 1 shows the categories of data that were used.

4 RESULTS AND DISCUSSION

Applying the modeling methodology to the raw data set, three different predictive models are achieved, one for each phase of the process, or even for the available data set, as shown in figure 3.

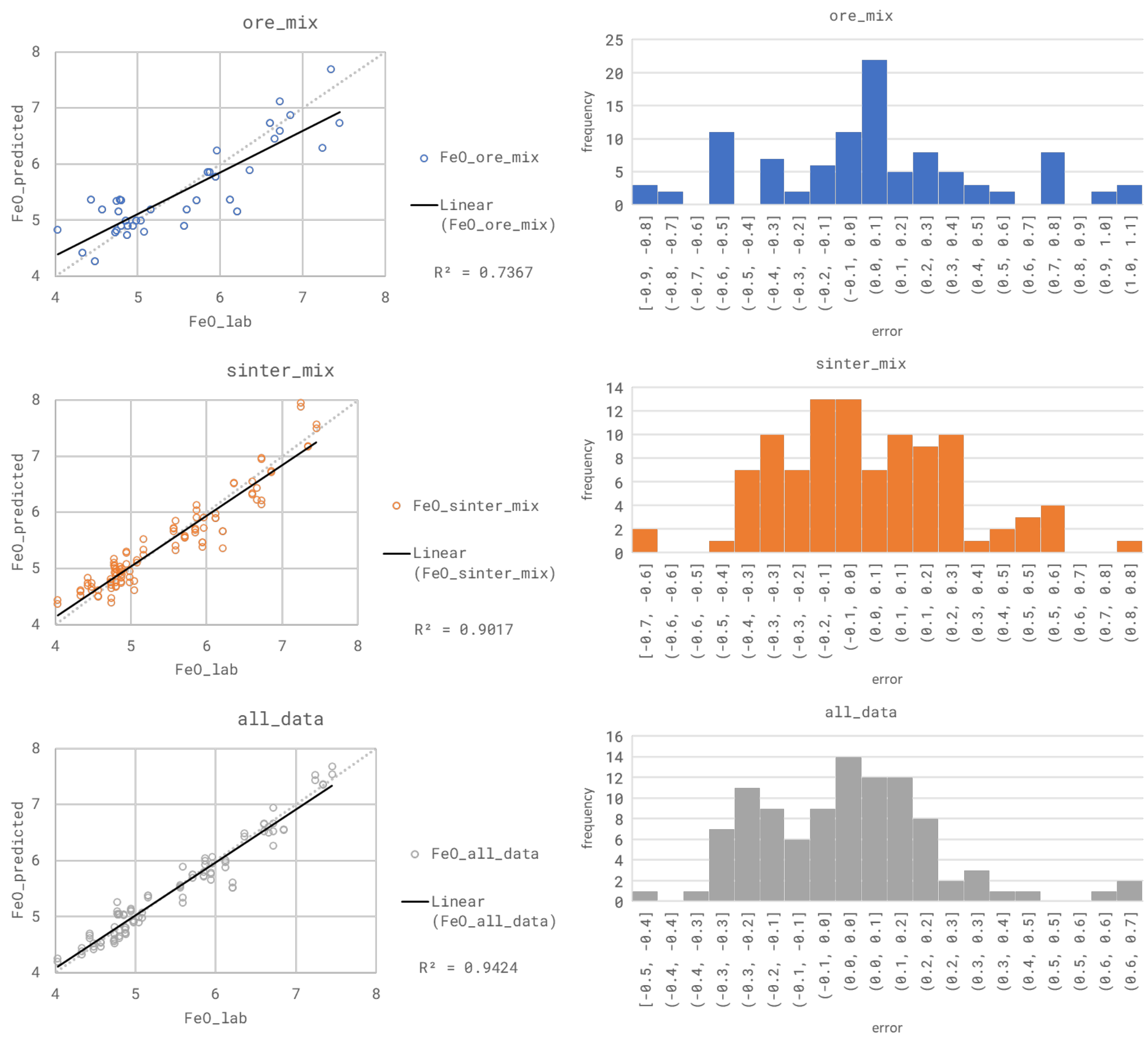

Figure 3. Summary modeling results for the raw dataset

* Technical contribution to the 9º Simpósio Brasileiro de Aglomeração de Minérios, part of the ABM Week 7th edition, August 1st-3rd, 2023, São Paulo, SP, Brazil.

In order to apply data augmentation, it is necessary to preserve the original variability and correlations intrinsic to the raw data set. In this way, the entire process of generating the augmented data set started from the correlation matrix of the original data, shown in figure 4.

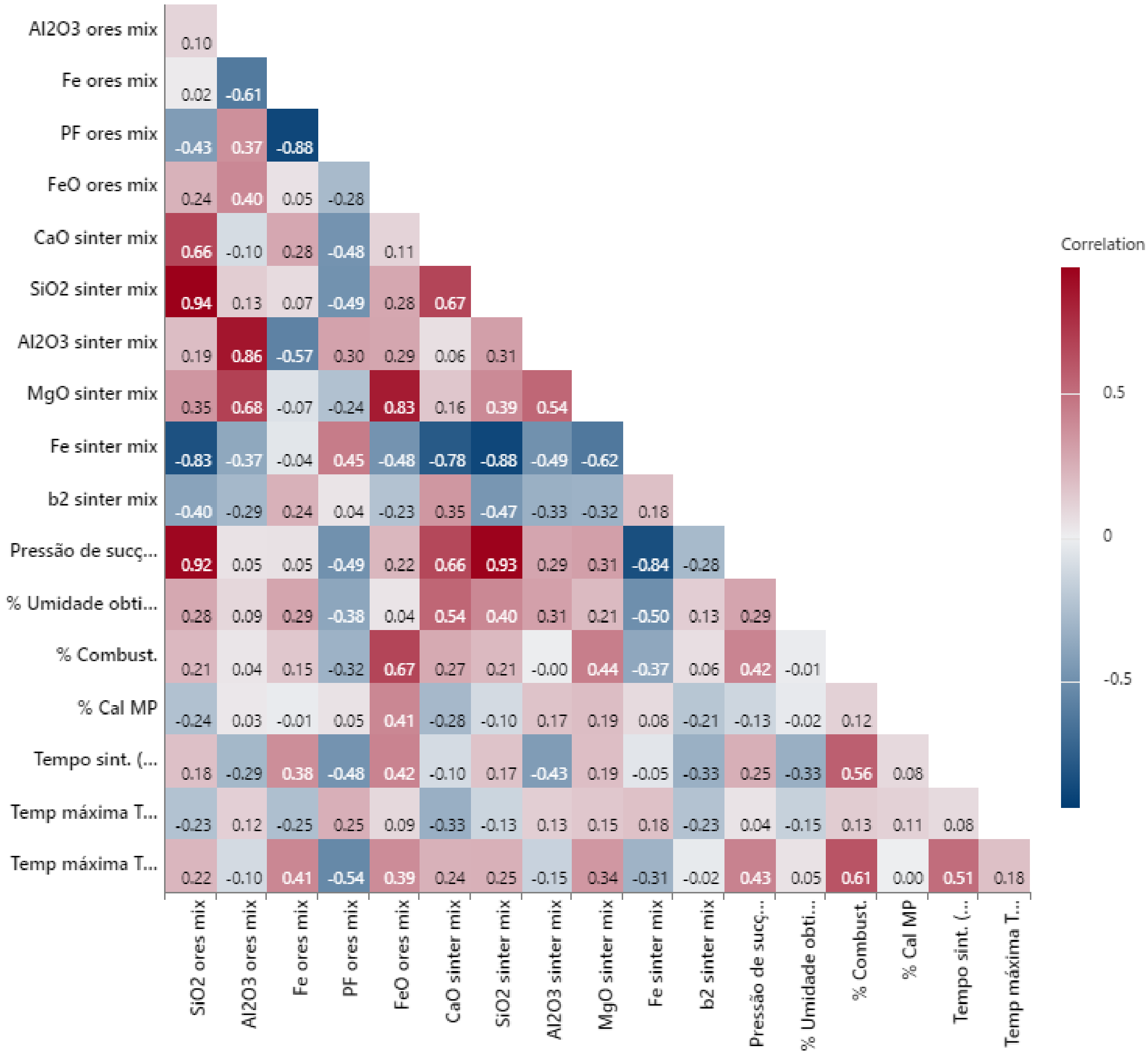

Figure 4. Correlation matrix of the raw dataset

Applying the modeling methodology to the augmented data set, three different predictive models are achieved, one for each phase of the process, or even for the available data set, as shown in figure 5

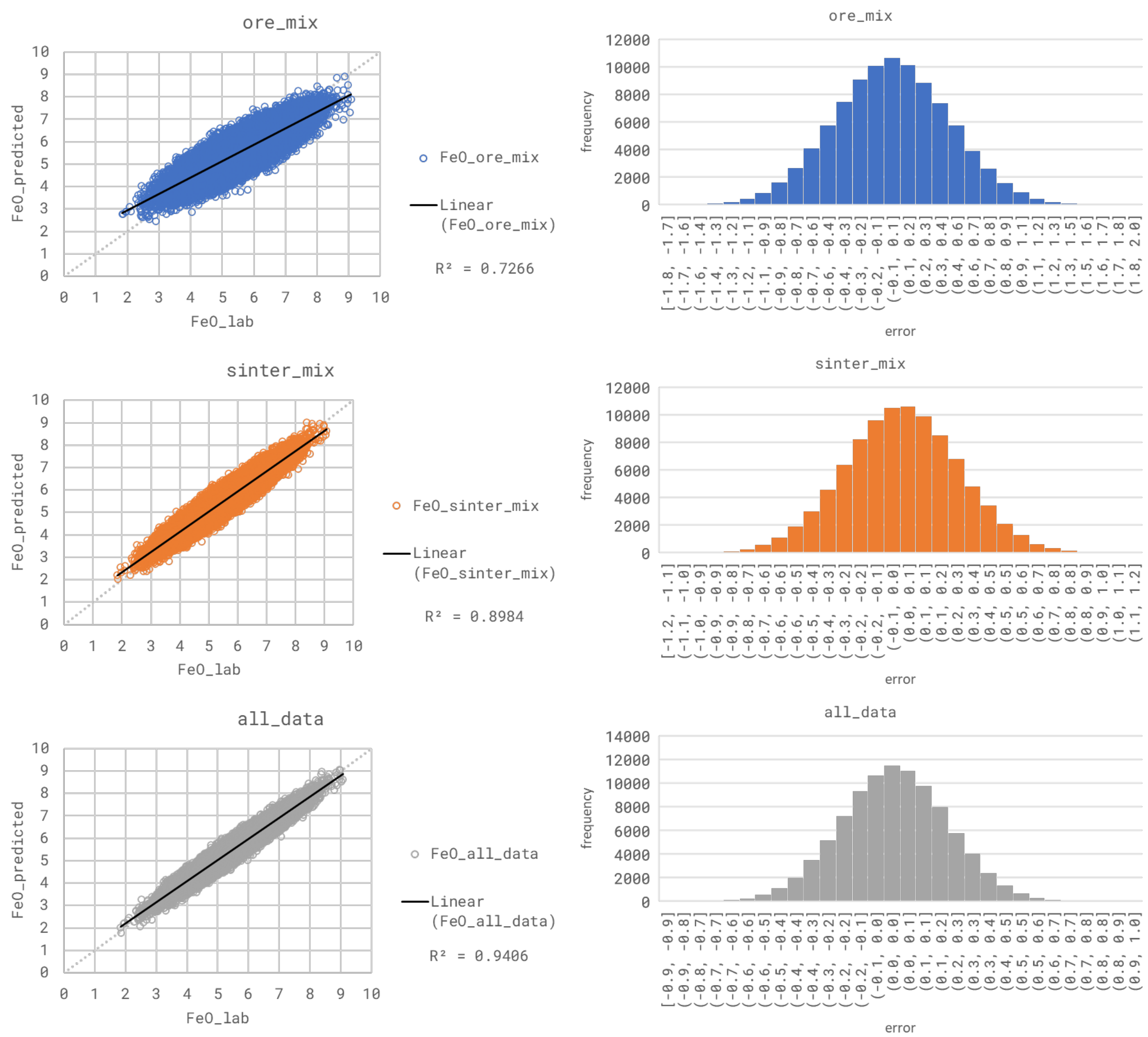

Figure 5. Summary modeling results for the augmented dataset

5 CONCLUSION

The developed model is able to predict the FeO of the sinter mixtures instantly and allows us to better plan the projects, without having to wait for the result from the laboratory tests.

Primary predictive models, using only the raw dataset, have already achieved determination coefficients high enough for model acceptance. With the data augmentation technique, it was possible to extrapolate and generalize these models, validating their broad and unrestricted application.

In summary, we can summarize the project and its results through the flowchart in figure 6.

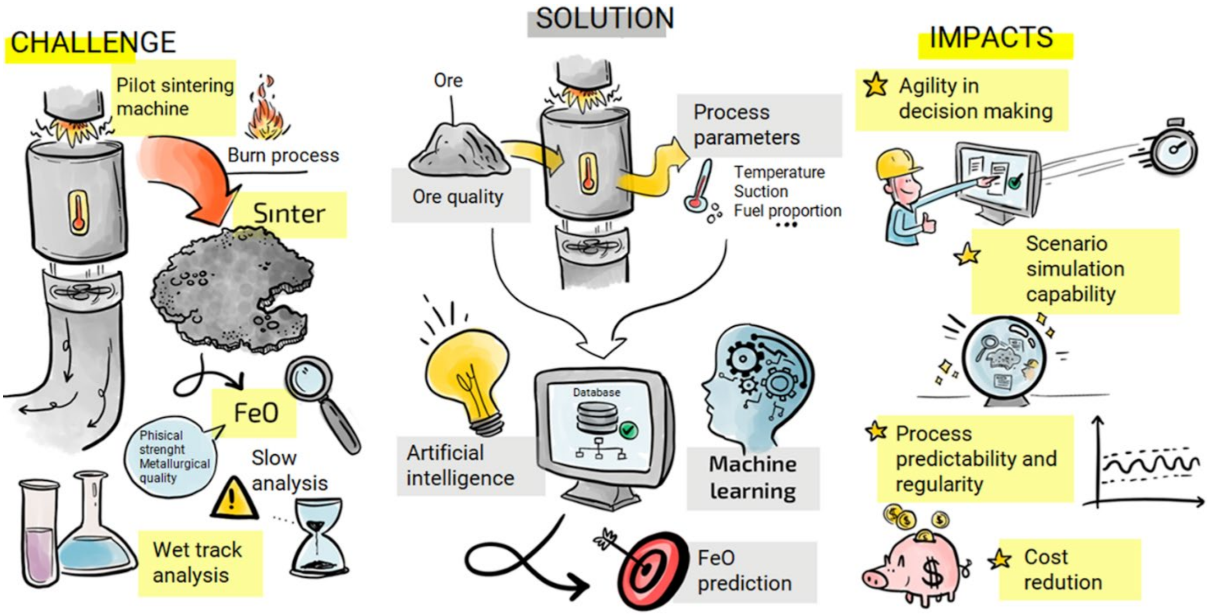

Figure 6. Didactic representation of the final achieved solution (Mining Hub's Challenge e-book)

Once implemented, this solution brings agility in decision-making, reduces costs with laboratory tests, brings predictability and regularity to the process, in addition to allowing the simulation of real or hypothetical scenarios.

Acknowledgments

Special thanks to the Vale (Ferrous Technology Center team) and the Mining Hub, involved in the realization of this project.

REFERENCES

1 Chen, J., Sun, Z., Li, X., Li, H., & Chen, J. (2019). Optimization of iron ore sintering process based on improved particle swarm optimization algorithm. ISIJ International, 59(8), 1451-1459.
2 Gao, J., Zhang, Y., Ma, X., & Wang, W. (2019). Development of a prediction model for sinter strength using machine learning techniques. Ironmaking & Steelmaking, 46(3), 238-245.
3 Jolliffe, I., Cadima, J., & Sandoval, J. (2016). Deep Principal Component Analysis. Journal of Machine Learning Research, 17(1), 6339-6365.
4 Kwang-Su Park, Hyeseon Lee, Chi-Hyuck Jun, Kwang-Hyun Park, Jae-Won Jung, Seung-Bin Kim, Rapid determination of FeO content in sinter ores using DRIFT spectra and multivariate calibrations, Chemometrics and Intelligent Laboratory Systems, Volume 51, Issue 2, 2000, Pages 163-173, ISSN 0169-7439, https://doi.org/10.1016/S0169-7439(00)00067-8.
5 Mining Hub (2021). M-Spot Cycle 1 Vale Challenge e-book.
6 Sugiyama, M., & Borgwardt, K. M. (2015). Linear Discriminant Analysis in High-Dimensional Spaces: Theoretical Issues and Applications. Journal of Machine Learning Research, 16(1), 307-349.
7 Xiong, S., & Pan, W. (2018). Multivariate Regression with Calibration. Journal of Machine Learning Research, 19(1), 5719-5756.
8 Z. Jiang, L. Huang, K. Jiang and Y. Xie, "Prediction of FeO Content in Sintering Process Based on Heat Transfer Mechanism and Data-driven Model," 2020 Chinese Automation Congress (CAC), Shanghai, China, 2020, pp. 4846-4851, doi: 10.1109/CAC51589.2020.9327289.