USO DE INTELIGÊNCIA ARTIFICIAL, AUTO APRENDIZAGEM DE MÁQUINA, PARA MODELAGEM PREDITIVA DE PROPRIEDADES METALÚRGICAS DE PRODUTOS DE AÇO LAMINADOS A QUENTE*

Alisson Paulo de Oliveira[1]
Leonardo Sene de Lourenço[2]

Resumo

Este trabalho teve como objetivo a modelagem preditiva, com Inteligência Artificial, das propriedades mecânicas de perfis de aço ARBL laminados a quente. Os modelos foram baseados em dados históricos de propriedades mecânicas, e na composição química das corridas e parâmetros do processo de laminação. Foi utilizada uma plataforma *Auto-Machine* Learning. Esta ferramenta é capaz de testar simultaneamente dezenas de algoritmos visando o menor erro. Modelos simplificados foram construídos com base em análises estatísticas da base de dados, e modelos ampliados foram desenvolvidos utilizando todos os dados disponíveis. Os modelos foram desenvolvidos para serem metalurgicamente coerentes com as tendências científicas, apesar da precisão matemática. Os resultados alinharam-se bem com as tendências esperadas na maioria dos casos. Foi possível avaliar o efeito isolado das variáveis. Os modelos expandidos foram capazes de gerar previsões com menor erro estatístico. A variabilidade dos dados é um fator importante para o sucesso dos modelos preditivos. Tais modelos permitem que o projeto de liga seja realizado com maior precisão, menores custos de produção e melhor compreensão do efeito da variável de entrada. A tomada de decisões baseada em dados nas atividades de produção e P&D é aprimorada com o uso de ferramentas preditivas baseadas em Inteligência Artificial.

Palavras-chave: Inteligência Artificial; Aprendizado de Máquina; ARBL; Modelagem.

USE OF ARTIFICIAL INTELLIGENCE, AUTO MACHINE LEARNING, FOR PREDICTIVE MODELLING OF METALLURGICAL PROPERTIES OF HOT-ROLLED STEEL PRODUCTS

Abstract

This work aimed at the predictive modeling, with Artificial Intelligence, of the mechanical properties of hot-rolled structural steel sections, produced in High Strength and Low Alloy steel, HSLA. The models were based on historical data of mechanical properties as well as the chemical composition of the heats and rolling process parameters. An Auto-Machine Learning (Auto-ML) platform was used. This tool is capable of simultaneously testing dozens of algorithms aiming for the lowest error. Simplified models were built based on database statistical analyses, and expanded models were developed using all available data. The models were developed to be metallurgically coherent with scientific trends, despite mathematical precision. The results aligned well with expected trends in most cases. It was possible to evaluate the isolated effect of the variables. The expanded models were able to generate predictions with lower statistical error. The variability of the data is an important factor for the success of predictive models. Such models allow the steel chemical composition design to be carried out with greater accuracy, lower production costs, and improved understanding of the effect of input variable. Data-Driven Decision-Making in production and R&D activities are enhanced with the use of predictive tools based on Artificial Intelligence.

Keywords: Artificial Intelligence; Machine Learning; HSLA; Modelling.

[1] *ABM Member, Metallurgist Engineer, D.Sc., M.Sc., Metallurgical, Materials and Mining Engineering. Executive Director, NSigma Consulting, LTD, Belo Horizonte, MG, Brazil.*

[2] *Data Scientist, Ph.D. and Tech Lead in AI, A10 Consulting - IN Group Now Vertical, São Paulo, SP, Brazil.*

** Contribuição técnica ao 26° Seminário de Automação e TI, parte integrante da ABM Week 8ª edição, realizada de 03 a 05 de setembro de 2024, São Paulo, SP, Brasil.*

1. INTRODUCTION

The rolling process of structural steel sections presents a major challenge in adjusting the chemical composition to meet the metallurgical properties required by different international standards. This challenge arises from understanding the effect of various variables involved in the evolution of the properties throughout the rolling process. Numerous techniques have been applied with the aim of modeling the rolling process, providing not only predictive capability but also a greater understanding of the phenomena involved. Artificial Intelligence techniques are particularly efficient in predicting the results of mechanical tests with low error and high correlation, often aligning with the metallurgical phenomena involved.

The application of machine learning techniques in materials science serves a crucial purpose in accelerating the discovery of novel materials. Over time, as materials science research has advanced, a substantial amount of data from experimental or simulation studies has been amassed and continues to grow steadily. This accumulation of data forms the foundation for the widespread utilization of machine learning techniques in materials research and development processes. In line with the progress in materials data science, the Materials Genome Initiative (MGI) was introduced by the United States **[1]**. This initiative aims to expedite the research cycle and reduce economic costs through the adoption of high-throughput computing, data-driven methods, big data technologies, and more. In terms of methodology, machine learning methods based on data mining are closely related to applied statistics, given that machine learning forms a pivotal component of data science, primarily focusing on statistical data processing. Hence, machine learning methods typically encompass data preprocessing and the selection of suitable algorithms for training, testing, and validation **[1-2]**. Machine learning techniques facilitate the establishment of associations between composition, microstructure, process, and performance, enabling the prediction of new materials with exceptional performance in unknown domain spaces, a commonly employed strategy in materials development. Despite the emergence of numerous novel materials, no metal material has yet surpassed steel's dominant role as one of the fundamental structural materials. Steel material differs from others in terms of data characteristics, with the accumulation of substantial data during its development process providing a favorable environment for exploring artificial intelligence strategies **[1-2]**.

Further advancements in materials science, has led to machine learning techniques based on data-driven approaches emerging as a prominent focus in current materials research **[2–12]**. Unlike traditional physical modeling methods with intricate mechanisms, machine learning methods require sufficient data for training to uncover underlying rules. They excel in identifying correlations among multiple data points, making them effective for addressing multivariate nonlinear problems and potentially establishing accurate prediction models based on existing data **[13]**. The utilization of multi-objective optimization methods in conjunction with neural network adjustments to simultaneously reduce prediction error and average temperature difference has resulted in the identification of a model with enhanced predictability **[14]**. Additionally, the application of genetic algorithms (GA) has rendered the alloy discovery and optimization process computationally more affordable **[15]**. Advancements in microstructural observations facilitated by AI technology have paved the way for the automatic recognition of steel microstructures. This progress signifies a significant development in the understanding and control of steel microstructures, offering an

effective tool for researchers. Moreover, a design system integrating machine learning (ML) and high-throughput optimization algorithms has been devised to derive optimal solutions for reduced activation ferritic/martensitic (RAFM) steels. This system incorporates composition and treatment process modifications aimed at improving both yield strength and impact toughness **[16]**. Automated Machine Learning, also known as AutoML, refers to tools and services that abstract the details and knowledge required to perform Machine Learning, automating tasks necessary for ML to occur. Typically, they cover steps such as data normalization and feature engineering; training models of different types and with different hyperparameters; and evaluating and comparing results. There are also other tools, libraries, and services that assist in these tasks, but with a lower level of abstraction, and thus are not called AutoML. The goal of AutoML is to democratize access to analytical tools for non-data scientists by providing tools that do not require code or require very little code **[17, 18]**.

The present work focuses on modeling the relationship between tensile test results (Yield Strength, Tensile Strength, and Elongation) of hot-rolled products, which are the output data calculated from the models, and the results of chemical composition and rolling parameters. The training process also utilized historical tensile test data as input.

2. DEVELOPMENT

2.1. Materials and Methods

2.1.1. The steel section

A steel section, produced in HSLA (High Strength Low Alloy) steel, with a nominal thickness of 11.0 mm was chosen due to the following characteristics:

- Tensile testing sampling taken from the flange, resulting in lower variability of results.
- High quantity of tests conducted according to ASTM A572-50 standard. For the present study, a total of 461 sets of results were used.

The following information was used in the development of the Mechanical Properties Prediction Models: Steel chemical composition, final rolling temperature on the flange, measured thickness of the test specimen from the flange of the steel section (The flange thickness is the same as that of the test specimen) and historical data for Tensile Strength (LR), Yield Strength (LE), and Elongation (A). For the calculation of the percentage reduction during the rolling process, the initial thickness was taken as the thickness of the cast raw material in the flange region (equation 1).

$$\%Reduction = \frac{(E_f - E_i)}{E_i} * 100 \quad \textbf{(1)}$$

In this equation, E_f represents the final thickness of the rolled profile at the sampling location, and E_i represents the initial thickness of the cast raw material.

2.1.2. Database statistical treatment

The following analyses and actions were carried out:

- Correlation analysis between the various input variables and the outputs.
- Statistical characterization of the input data.
- Data Treatment: Range within +/- 3 standard deviations.
- Elimination of outliers.

These techniques were used to eliminate the presence of data that could compromise the reliability of the database and the final models.

2.1.3. Models Development

A total of six different models were built, comprising three simplified models and three expanded models. The output variables are Yield Strength (LE), Ultimate Tensile Strength (LR), and Elongation (A). The simplified models used a smaller amount of process parameters (YS, UTS, Elongation, %C, %Mn, %Si, %S, %Cr, %Nb, %N, Final Rolling Temperature [TFL] in °C, and Reduction) while the expanded models used all available data (YS, UTS, Elongation, %C, %Mn, %Si, %P, %S, %Cu, %Ti, %Cr, %Ni, %Nb, %Mo, %V, %B, %Al, %Sn, %W, %Zr, %As, %Ca, %Co, %Sb, %N, %Te, Final Rolling Temperature [TFL], and %Reduction). The goal was to understand if simplified models, initially less complex, could provide superior results in terms of correlation and mean errors. Regressive models were used through cross-validation, choosing a k-fold (randomized subsets obtained from the training bases, thus reducing bias in model training) equal to 5, for a base partitioned into 90% training and 10% testing. Using the Average Percentage Error (MAPE) as the error measurement metric. As an advantage of using the Auto ML software the data quality verification steps, as well as the field engineering (feature engineering), are carried out automatically. The last step is disabled due to the need to explain the variables that make up the model **[18]**. With the results obtained, it was possible to visualize all the modeling stages, both in the descriptive and predictive parts, since the adjustment curves were carried out, as well as the analysis of the residuals of the model.

2.2. Results

2.2.1. Selection of model parameters

- Simplified Models: Correlation analysis was carried out between the output variables and the input variables (chemical composition and rolling variables). The choice of input variables was made based on the correlation analysis between these and the output variables, involving the elaboration of the correlation matrix, the evaluation of the correlation coefficients obtained, and the elimination of chemical elements considered residual, that is, not intentionally added in the steelmaking process. The following variables were included in the simplified models:

    - Chemical Composition: Only the elements C, Mn, Si, S, Cr, Nb and N.
    - Process Variables: Final Rolling Temperature and Total Rolling Reduction.

After this step and the elimination of discrepant data points, a database was obtained with 461 occurrences with the variables indicated in table 1 (a).

- Expanded Models: The variables shown in table 1 (b) were used together with the variables shown in table 1 (a) to build this model. The maximum and the minimum values for additional variables can be seen in the table 1:

Table 1: Input and output variables (LE, LR and A) included in the simplified models (a) and expanded models (b).

Variable	Description	Maximum	Minimum
Y1	LE, MPa: Yield Strength	461,68	378,35
Y2	LR, MPa: Tensile Strength	551,99	469,10
Y3	A, %: Elongation	39,80	20,00
X1	C, %	0,1309	0,0808
X2	Mn, %	14540	12020
X3	Si, %	0,2500	0,1560
X4	S, %	0,0160	0,0031
X5	Cr, %	0,0470	0,0130
X6	Nb, %	0,0360	0,0200
X7	N_2, %	0,0069	0,0019
X8	Final Rolling Temperature (TFL), °C	1007,05	939,84
X9	Total Rolling Reduction, %	86,49	84,73

Variable	Description	Maximum	Minimum
X10	P, %	0.0312	0.0124
X11	Cu, %	0.011	0.001
X12	Mo, %	0.017	0.001
X13	V, %	0.0035	0.001
X14	Al, %	0.0429	0.0007
X15	Sn, %	0.0022	0.0003
X16	Ti, %	0.003	0.0001
X17	Ni, %	0.021	0.008
X18	Co, %	0.0058	0.0029
X19	Sb, %	0.0118	0.0006
X20	Zr, %	0.0052	0.0003
X21	As, %	0.0024	0.0004
X22	Te, %	0.0015	0.0001
X23	B, %	0.0004	0.0001
X24	Ca, %	0.0016	0.0001
X25	W, %	0.004	0.001

2.2.2. Architecture of the AI models

The Auto ML tool features competition between models, enabling functions such as mixing models (Blend Models) in addition to modeling in parallel with different families of models, highlighting the Ensemble, Neural Networks and Traditional Regressive models with their strands. This type of modeling allows the indication of the most appropriate model and, depending on the need for use, the user will be able to choose the one that had the best performance combined with the best way of presenting the results. This architecture allowed obtaining the optimal result in a few minutes, since the adjustment of missing data had already been carried out in an automated way, with the recommendation of the widely used modeling methodology (cross-validation), it was possible to obtain models with appropriate metrics for applications found in literature and academic works, according to the master's thesis of one of the authors.

2.2.3. Training of the models: Expanded LR Model

The blue "Forecast" line, in the Lift Chart, figure 3 (a), displays the average prediction score for the rows in that bin. The orange, "Actual" line, displays the actual percentage of the lines in that bin. A strong correlation between these two lines signifies the model's predictive accuracy; a steadily increasing line is another positive indicator of satisfactory model performance. Another important key aspect: There is an alternation of real values above or below the line of predicted values. This indicates that the model is not biased. At the extremes, where there is little data, the model tends to have a good accuracy. The Matrix Plot on the figure 3 (b) shows the association between the process variables (input and output data). This matrix provides information about the strength of association between numeric and categorical feature pairs that are visually indicated by color opacity and feature clusters. The greater the opacity of the pair, the weaker the association between the

variables. The lower the opacity of the pair, the stronger the association between the variables. LR has a strong association with Als, Alt, C, Cr, Mn, N, Nb, P, S, Si, Mo, Reduction and As. LR has weak association with Pb, Sb, Te, Zr, Ca, Co, Ti, V, W, Final Rolling Temperature [TFL], B, Cu and Sn. Clusters, or families of resources indicated by colors in the matrix, are resources partitioned into groups based on their membership structure.

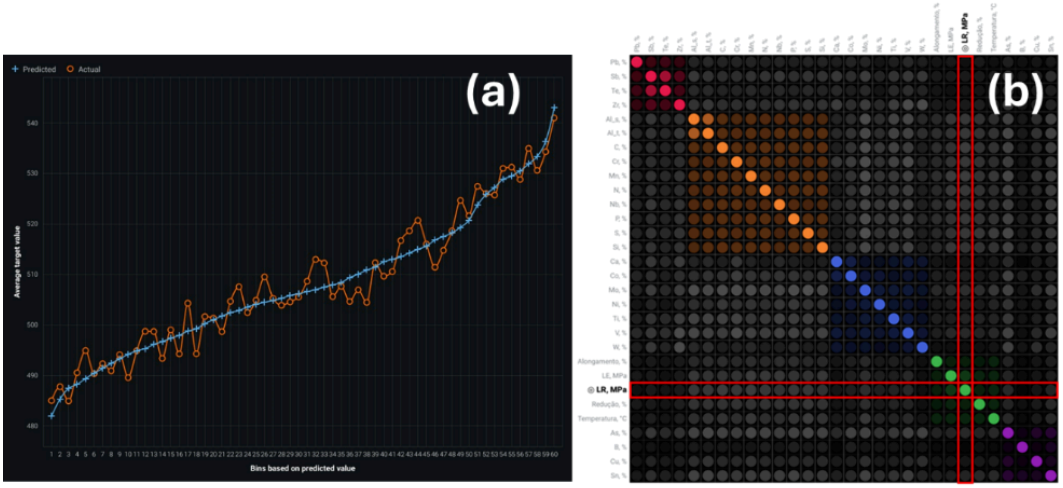

Figure 3: Training results for the expanded LR Model: (a) Lift Chart, (b) Matrix Plot.

The impact of the variables in the LR parameter can be seen in the plot 4 (a) (Figure 4). The variables that most impact LR are: %C, %Mn, %Al, %Nb, %Ti and %Reduction. Mn has approximately 60% of the influence of C. The variables that have the least impact on LR are %S, %Cu, %N, %Si and %P. These results are in line with what was expected, metallurgically.

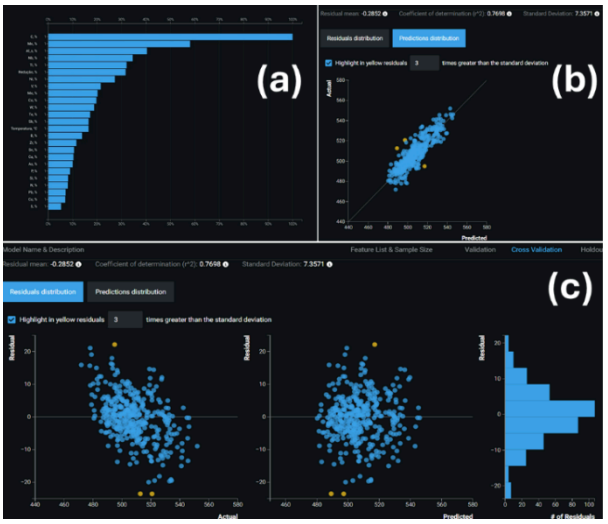

Figure 4: (a) Variable Impact and residual analysis, (b) LR Predicted vs, LR Actual and (c) Residues Analysis for the expanded LR Model.

2.2.4. Influence of the input variables on the output variables (LR)

It was possible to assess the individual impact of input variables, such as each of the chemical elements and process parameters, for LE, LR, and A. This assessment is done through partial dependence plots, which show the average partial relationship between a set of predictors and the predicted output. In these plots, the yellow partial dependence data represent the marginal effect of a feature on the target variable after accounting for the average effects of all other predictive features. It indicates how, while keeping all other variables constant, the value of this feature affects the prediction. In the figure below, it is possible to assess those plots:

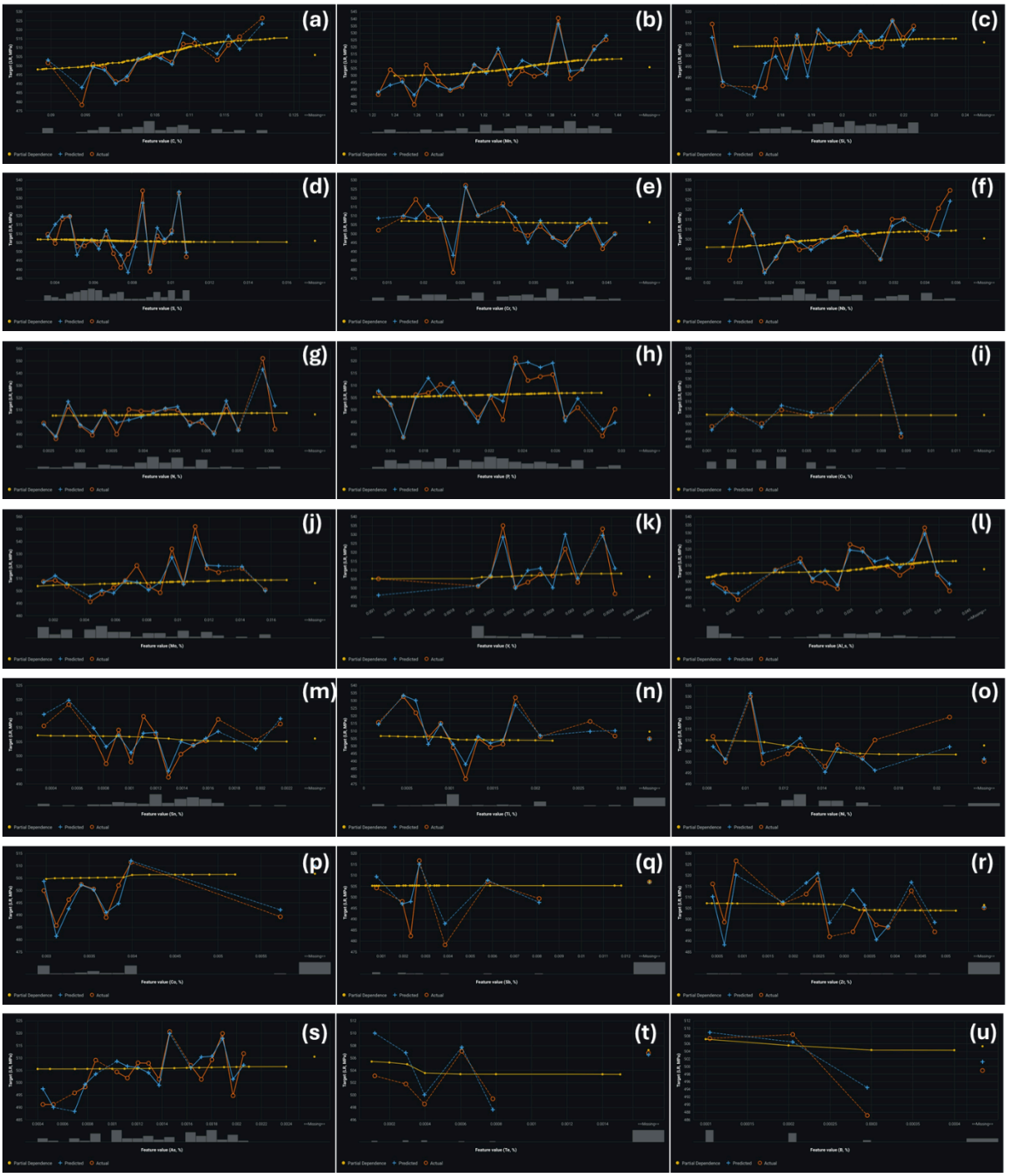

Figure 5 (a) to (z): Partial dependence plots for all the input variables and LR as an output: (a) LR x C, %; (b) LR x Mn, %; (c) LR x Si, %; (d) LR x S, %; (e) LR x Cr, %; (f) LR x Nb, %; (g) LR x N, %; (h) LR x P, %; (i) LR x Cu, %; (j) LR x Mo, %; (k) LR x V, %; (l) LR x Al, %; (m) LR x Sn, %; (n) LR x Ti, %; (o) LR x Ni, %; (p) LR x Co, %; (q) LR x Sb, %; (r) LR x Zr, %; (s) LR x As, %; (t) LR x Te, %; (u) LR x B, %, (v) LR x Ca, %; (x) LR x W, %; (y) LR x Reduction, % and (z) LR x Temperature, °C.

According to partial dependence plots it is possible to obtain some interesting information about how each input variable impacts the output variable, considering LE, LR and A. The interpretation of these plots are exhibited in table 2.

Table 2: (a) Impact of the input variables on the LE, LR and A. (b) Alignment of the impact with technical literature for LR.

Input Variables	LE, MPa		LR, MPa		A, %	
	Simplified	Expanded	Simplified	Expanded	Simplified	Expanded
C, %	↑	↑	↑	↑	→	→
Mn, %	↑	↑	↑	↑	↘	→
Si, %	↗	↗	↑	↗	→	→
P, %	-	↗	-	↗	-	→
S, %	↘	↘	→	↘	↘	→
Cu, %	-	→	-	→	-	→
Ti, %	-	→	-	↘	-	↗
Cr, %	↘	↘	↘	↘	→	↗
Ni, %	-	↘	-	↓	-	→
Nb, %	↑	↑	↗	↑	→	↗
Mo, %	-	→	-	↗	-	→
V, %	-	→	-	↗	-	→
B, %	-	→	-	↘	-	↘
Al, %	-	↗	-	↑	-	→
Sn, %	-	→	-	↓	-	→
W, %	-	→	-	→	-	→
Zr, %	-	↘	-	↘	-	→
As, %	-	→	-	→	-	↘
Ca, %	-	→	-	→	-	→
Co, %	-	→	-	↗	-	↘
Sb, %	-	→	-	-	-	-
N, %	→	→	↗	↗	→	→
Te, %	-	→	-	↘	-	→
TFL, °C	↑	↗	↑	↑	→	→
Redução, %	↗	↗	↑	↑	→	→

Subtitle				
	↑	Substantial increase		High impact
	↗	Slight increase		Medium impact
	→	Stability		Low impact
	↘	Slight reduction		Medium impact
	↓	Substantial reduction		High impact

Input Variables	LR Analysis for the Expanded Model		
	Alloy-Design Range, %	LR/Alloy-Design, %	Literature alligned?
C, %	0.01	5.00	Yes!
Mn, %	0.01	0.60	Yes!
Si, %	0.01	0.62	Yes!
P, %	0.001	0.17	Yes!
S, %	0.001	-0.17	Yes!
Cu, %	0.001	0.00	No (Residual)!
Ti, %	0.0001	-0.16	No (Residual)!
Cr, %	0.001	-0.07	Yes (Residual)!
Ni, %	0.001	-0.17	No (Residual)!
Nb, %	0.001	0.56	Yes!
Mo, %	0.001	0.21	Yes (Residual)!
V, %	0.0001	0.21	Yes (Residual)!
B, %	0.00001	-0.10	No (Residual)!
Al, %	0.001	0.24	Yes!
Sn, %	0.0001	-0.11	Yes (Residual)!
W, %	0.001	0.67	Yes (Residual)!
Zr, %	0.0001	-0.09	No (Residual)!
As, %	0.0001	0.05	No (Residual)!
Ca, %	0.0001	0.14	Yes (Residual)!
Co, %	0.001	0.45	Yes (Residual)!
Sb, %	0.001	0.00	?
N, %	0.001	0.29	Yes (Residual)!
Te, %	0.0001	-0.17	?
TFL, °C	1	0.04	Yes!
Redução, %	0.1	0.58	Yes!

Table 2 (a) shows the impact of the input variables on the output variables, LE, LR and A. This impact is divided into 5 different categories, from a substantial increase in mechanical properties to a substantial reduction in these properties. Some of the input variables, such as C, Mn, Nb, Final Rolling Temperature [TFL] and Reduction, have a great influence in the output variables. Table 2 (b) shows the analysis result of the input variables influence (or impact) in the LR for the expanded model and whether this influence is in accordance with the available technical literature. It is

possible to see that for most of the input variables there is an alignment, including for some residual chemical elements, with the literature and for those cases in which there was no alignment, all elements are residual.

For example, in microalloyed steels, microalloying elements such as niobium can interact with the movement of austenite grain boundaries in two different ways. Firstly, the dispersion of Nb-containing precipitates exerts a retarding pressure on the grain boundaries, which ends up having a pronounced effect on the growth of the austenitic grain. The magnitude of this effect depends on the size, shape, and volume fraction of these precipitates, typically Nb(C,N), or Nb carbonitrides. Secondly, a moving interface drags an atmosphere of solute elements, which exerts a retarding force on this interface. Thus, elements in solution can significantly reduce the mobility of austenitic grain boundaries. This phenomenon is generally known as the solute drag effect **[19]**. From the LR results for the expanded model, it can be seen that Nb contributes 0.56MPa for every 0.001% addition, being in line with the known resistance increase mechanism. In other words, the addition of Nb contributes to the increase in mechanical resistance and the expanded model, as well as the simplified model, demonstrate this trend. Zirconium **[20]**, being a resource little used currently as a microalloying element, presents antagonistic results in quenched and tempered samples of HSLA steels, depending on its content. The increase in Zr content causes the refinement of carbon-containing inclusions and the precipitation of the inclusions consumes the amount of carbon in solid solution present in these HSLA steels, thus reducing the effect of increasing strength by solid solution. Being a strong carbide former, when its content is higher, zirconium can easily combine with carbon to form ZrC. Since the hardness of ZrC is high compared to other carbides, the hardness of these tempered HSLA steels increases with the increase in Zr content. Other studies **[21]** showed that particles such as ZrC, ZrN, as well as other complex particles, in austenitic steels had coarse characteristics and caused a reduction in yield and resistance limits at high temperatures. It is concluded, then, that the addition of Zr plays a dubious role in improving the mechanical resistance of steels. The constructed model indicates a reduction in LR of 0.09MPa for each addition of 0.0001% of Zr, remembering that in this case the levels are residual, that is, there was no intentional addition of Zr to this HSLA steel.

2.2.5. Comparison between the models

The table below illustrates a comparison between some variables of the modeling.

Table 3: Statistical comparison between the models.

Model	Parameter	Residuals Average	Correlation Coefficient, R^2	Standard Deviation
Simplified	LE	-0.1702	0.6268	9.3919
	LR	-0.1255	0.7085	8.2845
	A	-0.4354	0.2516	2.6153
Expanded	LE	0.0492	0.678	8.7251
	LR	-0.2852	0.7698	7.3571
	A	-0.079	0.3944	2.3838

2.2.6. Comparison with previous projects

The same dataset was subjected to modeling aiming at the construction of predictive models **[22]**, similar to the simplified model named in this paper. Artificial Neural Networks and Multiple Linear Regression were used. The results can be viewed in Table 4.

Table 4: Average prediction error according to the model and algorithm and output variable.

Model	Info	Output Parameters		
		LE	LR	A
Expanded Model	Average error, %	1.65	1.11	6.36
	Algorithm	eXtreme Gradient Boost	ARN, 3 Layers	Adaboost A
Simplified Model	Average error, %	1.89	1.33	6.41
	Algorithm	eXtreme Gradient Boost	eXtreme Gradient Boost	eXtreme Gradient Boost
ARN (MatLab)	Average error, %	2.26	1.57	4.76
	Algorithm	ARN, 3 Layers	ARN, 3 Layers	ARN, 3 Layers
Multiple Linear Regression	Average error, %	2.27	1.73	4.86
	Algorithm	MLR (Excel)	MLR (Excel)	MLR (Excel)

It is observed that the simplified model shows superior results compared to the ANN model, except for variable A, where the ANN result was even superior. The expanded model had an even better result than the simplified model.

3. CONCLUSION

- A tool based on a modelling like Auto-ML allows alloy design to be carried out virtually, i.e., before actual production. This implies in a reduction of product development cycle and total cost. Possible heat disqualifications can be minimized because of design errors that may occur during the research. All learned lessons, from past productions, can be used to boost the model efficiency.
- The models proved to be consistent with the metallurgical trends established in most of the analyzed variables and reflected the previously observed trends in the conducted statistical analyses. The models can be used to study the effect of each variable in isolation.
- The use of AutoML showed enormous potential for building predictive models for use in the steel industry. The expanded model, based in the majority of the available data, was able to make predictions with lower mean error and higher correlations. Comparing with simulations using other algorithms, it is clear that the AutoML models are more accurate, presenting lower error and higher correlation (expanded model).
- It was possible to verify the influence of all input variables on the final results. The tool used is capable of recognizing new patterns in a wide range of input variables. This feature shows the potential to investigate unknown relationships between variables and enables the use of new and innovative approaches for the development and research of new products.
- Important limitations encountered:
    - Quality of the database: The industrial origin of the data leads to high data dispersion within its range of variation with low correlations between input and output variables. The model developed for the Elongation [A, %], had significantly poorer performance (Lower correlation coefficient, R^2, see table 3, and higher average error, %, see table 4) compared to the other two models.

- The currently available database, although storing a large number of variables, lacked variables of great importance such as, for example: reductions according to the pass schedule, time inter passes, temperatures for each rolling pass and cooling rate, after the end of rolling.

The Automated Machine Learning (Auto-ML) tool appears as a strategic asset for aiding in the formulation of alloys destined for structural purposes and other steel-based products. This tool eases the research and development (R&D) phase by initially simulating various alloy design scenarios in a virtual environment, thus perfecting the process before actual production begins. By conducting virtual production trials prior to the physical implementation, significant reductions in development expenses and project lead times are achieved. Notably, the empirical models prove commendable predictive accuracy, showed by minimal prediction errors (less than 2% for all LE and LR). Consequently, these empirical models play a pivotal role in data-informed decision-making processes, fostering innovation and substantial cost savings within the steel industry.

Acknowledgments

- **Érico Felipe Aleixo, Technology and Business Intelligence Solutions Director, A10 Now Vertical Company:** For supporting this project, providing all the necessary resources.
- **A10 Consulting, IN Group Now Vertical, São Paulo, SP, Brazil**. For supporting this project, providing the tools, and related discussions.

REFERENCES

1. Pan, G., Wang, F., Shang, C. et al. Advances in machine learning- and artificial intelligence-assisted material design of steels. Int J Miner Metall Mater, 2023, 30, 1003–1024. https://doi.org/10.1007/s12613-022-2595-0.
2. Duan, H., He, H., Yue, S. et al. Analysis of High-Cycle Fatigue Life Prediction of 304 Stainless Steel Based on Deep Learning. JOM, 2023, 75, 4586–4595. https://doi.org/10.1007/s11837-023-06042-8.
3. Srivastava, A., Patra, P., Jha, R. AHSS applications in Industry 4.0: Determination of optimum processing parameters during coiling process through unsupervised machine learning approach, Materials Today Communications, 2022, 31, 103625, ISSN 2352-4928, https://doi.org/10.1016/j.mtcomm.2022.103625.
4. Hamdi, A., Merghache, S.M. Application of artificial neural networks (ANN) and gray relational analysis (GRA) to modeling and optimization of the material ratio curve parameters when turning hard steel. Int J Adv Manuf Technol, 2023, 124, 3657–3670. https://doi.org/10.1007/s00170-023-10833-3.
5. Czinege, I., Harangozó, D. Application of artificial neural networks for characterisation of formability properties of sheet metals, International Journal of Lightweight Materials and Manufacture, 2024, volume 7, Issue 1 pages 37-44, ISSN 2588-8404, https://doi.org/10.1016/j.ijlmm.2023.08.003.
6. Monroe, R. Data-Driven Design Properties for Cast Carbon Steels. Inter Metalcast, 2023. https://doi.org/10.1007/s40962-023-01195-3.
7. Shang, C., Wang, C., Wu, H. et al. Improved data-driven performance of Charpy impact toughness via literature-assisted production data in pipeline steel. Sci. China Technol. Sci., 2023, 66, 2069–2079. https://doi.org/10.1007/s11431-023-2372-x.

8. Durodola, J.F. Machine learning for design, phase transformation and mechanical properties of alloys, Progress in Materials Science, 2022, volume 123 100797, ISSN 0079-6425, https://doi.org/10.1016/j.pmatsci.2021.100797.
9. Chen, M., Zhang, T., Gong, Z., Zuo, W., Wang, Z., Zong, L., Zhao, O., Hu, L. Mechanical properties and microstructure characteristics of wire arc additively manufactured high-strength steels, Engineering Structures, 2024, volume 300, 117092, ISSN 0141-0296, https://doi.org/10.1016/j.engstruct.2023.117092.
10. Shi, Z., Du, L., He, X. et al. Prediction Model of Yield Strength of V–N Steel Hot-rolled Plate Based on Machine Learning Algorithm. JOM, 2023, 75, 1750–1762. https://doi.org/10.1007/s11837-023-05773-y.
11. Altuğ, M., Söyler, H. Optimization with artificial intelligence of the machinability of Hardox steel, which is exposed to different processes. Sci Rep, 2023, 13, 14100. https://doi.org/10.1038/s41598-023-40710-8.
12. Kong, B.O., Kim, M.S., Kim, B.H. et al. Prediction of Creep Life Using an Explainable Artificial Intelligence Technique and Alloy Design Based on the Genetic Algorithm in Creep-Strength-Enhanced Ferritic 9% Cr Steel. Met. Mater. Int., 2023, 29, 1334–1345. https://doi.org/10.1007/s12540-022-01312-7.
13. Chakraborty. S., Chattopadhyay, P., Ghosh. S., Datta, S. Incorporation of prior knowledge in neural network model for continuous cooling of steel using genetic algorithm, Applied Soft Computing, 2017, volume 58, Pages 297-306, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2017.05.001.
14. Jozaghi, T., Wang, C., Arroyave, R., Karaman, I. Design of alumina-forming austenitic stainless steel using genetic algorithms, Materials & Design, 2020, volume 186, 108198, ISSN 0264-1275, https://doi.org/10.1016/j.matdes.2019.108198.
15. Tsutsui, K., Terasaki, H., Maemura, T., Hayashi, K., Moriguchi, K., Morito. Microstructural diagram for steel based on crystallography with machine learning, Computational Materials Science, 2019, volume 159, pages 403-411, ISSN 0927-0256, https://doi.org/10.1016/j.commatsci.2018.12.003.
16. Wang, C., Shen, C., Huo, X., Zhang, C., Xu, W., Design of comprehensive mechanical properties by machine learning and high-throughput optimization algorithm in RAFM steels, Nuclear Engineering and Technology, 2020, volume 52, Issue 5, pages 1008-1012, ISSN 1738-5733, https://doi.org/10.1016/j.net.2019.10.014.
17. Herbst, C. O que é AutoML e quais são as suas vantagens? [Internet]; 2024. Avaliable from: https://www.eldorado.org.br/en/blog/o-que-e-automl-e-quais-sao-as-suas-vantagens/.
18. Chadha, B., Juwe, S. Agile Machine Learning with DataRobot: Automate Each Step of the Machine Learning Life Cycle, from Understanding Problems to Delivering Value, Packt Publishing, Limited, 2021, ISBN 1801076804.
19. Graux, A., Cazottes, S., De Castro, D., et. al. Precipitation and grain growth modelling in Ti-Nb microalloyed steels, Materialia, 2019, volume 5, 100233, ISSN 2589-1529, https://doi.org/10.1016/j.mtla.2019.100233.
20. Wang, F., Zheng, X., Long, J., Zheng, K. and Zheng, Z. Effects of Zirconium on the Structure and Mechanical Properties of High-Strength Low-Alloy Steels under Quenched or Tempered Conditions. steel research int., 2022, 93: 2200352. https://doi.org/10.1002/srin.202200352.
21. Moon J., Jang M.-H., Kang J.-Y., Lee T.-H. The negative effect of Zr addition on the high temperature strength in alumina-forming austenitic stainless steels. Materials Characterization. 2014, 87:12–18. https://doi.org/10.1016/j.matchar.2013.10.029.

22. Oliveira, A. P. Prediction Model of Mechanical Properties of Hot-Rolled Structural Beams: An Approach in Artificial Neural Networks. (Dissertation, master's in metallurgical and mining engineering). Digital Library of Universidade Federal de Minas Gerais, UFMG, 2008, http://hdl.handle.net/1843/MAPO-7RLKBJ.