XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGENEERING

Automatic Sequence Classification in HSM: Enhancing Process with Artificial Intelligence

José Fabiano M Assuncao [1], Leonardo Scardua [2]

[1]*Master's Student - Professional Master's Program in Control and Automation Engineering (Propecaut), Federal Institute of Espirito Santo (IFES) - Serra Campus*
assuncao.fabiano@gmail.com
[2]*Faculty Advisor - Professional Master's Program in Control and Automation Engineering (Propecaut), Federal Institute of Espirito santo (IFES) - Serra Campus*
330 Sabiás Avenue, Morada de Laranjeiras - Serra – ES, ZIP Code: 29166-630, Brazil
lascardua@ifes.edu.br

Abstract. The selection of rolling sequence types in the Hot Strip Mill (HSM) in ArcelorMittal Tubarao (AMT), Largest steel producer in Latin America and global market leader, is traditionally based on empirical evaluation by operational experts. However, the complexity that arises from order variability, slab yard dynamics, and operational constraints limits the efficiency of manual decision-making. This study proposes a methodology to automate the selection of sequence types based on the feature analysis of approximately 330 empirically assessed sequences. Preprocessing and feature engineering techniques were employed to support supervised learning models trained to categorize six main sequence types. The results show consistent performance, with high levels of accuracy, sensitivity, and specificity, highlighting the potential of the approach to improve in 10% process efficiency and compliance with key production KPIs such as decarbonization and customer satisfaction.

Keywords: Machine Learning, Sequence classification, process optimization, Hot strip mill.

1 INTRODUCTION

In steel manufacturing, HSM is where large solid blocks of metal (called slabs) are reheated and flattened into long, thin sheets. These sheets are used in everyday products such as cars, appliances, and construction materials. The process is fast, hot, and precise, and the way slabs are grouped and rolled is called a hot rolling sequence. A rolling sequence refers to the planned order in which metal slabs are processed in the rolling mill. It defines how slabs are grouped and directed for rolling, considering both technical and operational criteria, such as dimensions (thickness, width, length); equipment capabilities and limitations; environmental targets, energy efficiency, and customer requirements.

It is essentially like organizing a list of items to be processed with maximum efficiency while respecting all technical and commercial requirements. In this study, six main types of sequences were identified. The types are as follows: Standard Materials (SM), Specialized materials(SML), Restrictive Materials(RM), Application-specific Materials (MI), Opportunity Materials (MG) and Special Materials (SMT), each with different production goals like prioritizing energy efficiency, reducing carbon emissions, or staying aligned with incoming orders. These types reflect how operators traditionally balance technical constraints with business needs.

The choice of a rolling sequence is a strategic decision that involves multiple input variables, carefully evaluated by operators to meet the technical and commercial requirements of each production. The main factors considered include: steel grade, desired final thickness, temperature profiles throughout the process, required mechanical properties, and specific production constraints. These elements guide the definition of the best arrangement of materials and operational parameters, balancing energy efficiency, metallurgical performance, and the achievement of industrial goals.

This paper presents a machine learning–based decision support system designed to optimize hot rolling sequences in coil production. The proposed approach involves constructing a structured historical database containing process parameters, input conditions, and rolling sequences. These data are used to train a neural network classifier capable of predicting the most suitable sequence for new production scenarios. This strategy reflects a growing trend in industrial process optimization, where data-driven models are enhanced by domain expertise.

According to Latham and Giannetti [1], "The integration of expert knowledge and machine learning algorithms has proven effective in identifying failures within complex industrial processes, in the context of HSM." Such findings reinforce the relevance of combining historical data with intelligent systems to improve decision-making in hot strip mill operations.

The proposed approach leverages machine learning techniques to generate accurate, efficient, and production-feasible sequencing recommendations for hot strip mill operations. Rather than detailing a decision support architecture, the focus remains on predictive modeling grounded in historical process data. This aligns with broader trends in digital transformation, where emerging technologies have shown potential to enhance operational efficiency. As G. Gepitulan and Jackson [2] highlight, such innovations can significantly improve performance in HSM environments, reinforcing the relevance of model-based AI solutions in industrial contexts.

2 LITERATURE REVIEW

To date, no publications have been identified that specifically address the same industrial classification problem presented in this study. However, one works are closely related in scope and methodology:

The study conducted by Del Gallo et al. addresses the challenge of applying artificial intelligence to production scheduling in real industrial environments. By analyzing existing literature, the authors highlight how AI techniques have been successfully employed to improve operational efficiency, reduce costs, and enhance decision-making in complex manufacturing contexts. These findings are directly related to the present work, which also explores the use of machine learning to optimize production sequences in hot strip mill operations.

3 DESCRIPTION OF THE EXPERIMENTAL PROCEDURE (DATA ACQUISITION)

The experiments utilized datasets generated by the HSM production sequencing system (PSIMetals), integrated directly into the company's production planning operations. A dataset comprising 330 records was created through this procedure.

During the experiments, datasets containing process parameters, input conditions, and rolling sequences were generated specifically to support the development and evaluation of the production sequencing system.

The dataset used in this study comprises 330 rolling sequences executed over a one-month period (March 2025), totaling approximately 400,000 tons of steel processed. Each record includes detailed information on process parameters, input conditions, production priorities, and operational constraints, all derived from real industrial operations in a hot strip mill. These data form the basis for training and evaluating the predictive model.

Figure 1 presents an overview of the rolling process, since from the reheating furnace to coil production.

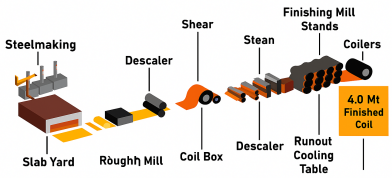


Figure 1. Overview Process

4 DATA PREPARATION and MODEL VALIDATION METHOD

Data preprocessing involved Python tools (Pandas, NumPy, Seaborn), with One-Hot Encoding applied to the target variable. SMOTE was used to correct class imbalance, and missing or irrelevant data were removed. Normalization via MinMaxScaler ensured better performance for scale-sensitive algorithms. Finally, the dataset was split into training (75%) and test (25%) sets to support robust model evaluation. ?????

The validation method used was Holdout Validation: The dataset is divided into two subsets—training and testing. The model is trained on the training set and evaluated on the test set.

5 MODEL CONSTRUCTION METHOD

Artificial Neural Networks: The parameters listed below can be adjusted and combined in various ways to achieve the optimal configuration:

- **Architecture Parameters**: In this multiclass classification task, the best configuration found consisted of three hidden layers with 128 neurons each and ReLU activation function, along with one output layer with six neurons—corresponding to the six target classes—using the SOFTMAX activation function. The number of input features was ten.
- **Training Parameters**: After extensive testing, the following were defined 1000 epochs, batch size of 10, and a validation split of 0.2.
- **Optimization Parameters** The ADAM optimizer was used along with the Categorical Crossentropy loss function.

5.1 EVALUATION METRICS DESCRIPTION

The following key metrics were used to assess model performance is show in Table 1:

Table 1. Performance metrics by class. Overall accuracy: 86,27% (204 samples)

Class	Precision	Recall	F1-score	Support
0	0,75	0,75	0,75	36
1	0,90	0.90	0.90	40
2	0,97	1,00	0,99	33
3	1,00	1,00	1,00	36
4	0,72	0,58	0,65	36
5	0,79	1,00	0,88	23
Weighted average	0,86	0,86	0,86	

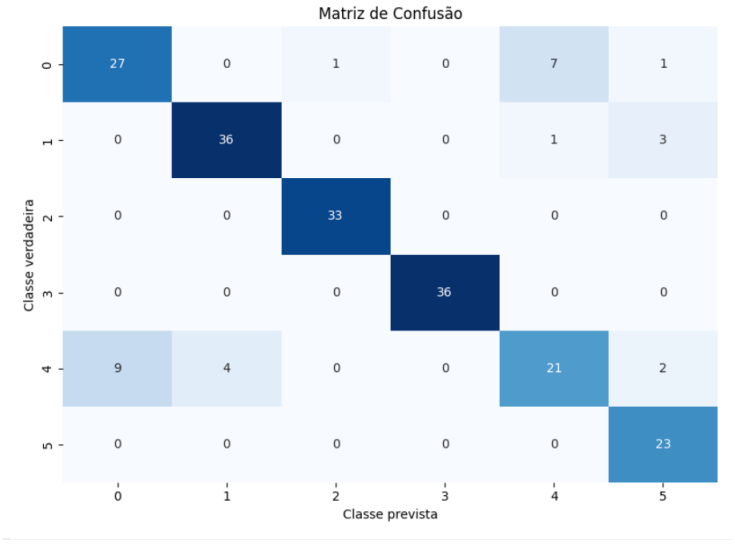


Figure 2. Confusion Matrix

6 CONCLUSIONS

The proposed approach demonstrates that Machine Learning can significantly enhance operational intelligence within the Hot Strip Mill (HSM) process. By enabling the automatic classification of production sequences,

the developed model provides a data-driven tool capable of identifying patterns that directly influence product quality and process stability.

The numerical results confirm that the model achieves **excellent precision and recall for most sequence categories**, particularly for Classes 2 and 3, which reached near-perfect or perfect identification ($\textit{Precision} \approx 1.00$, $\textit{Recall} = 1.00$). Classes 1 and 5 also exhibited strong performance, while moderate results for Classes 0 and 4 suggest opportunities for improvement through refined feature engineering or data balancing. Overall, an accuracy of $\mathbf{86.27}\%$ represents a robust outcome for a multiclass industrial classification problem involving complex and imbalanced data.

Beyond predictive metrics, this study highlights the practical value of integrating AI-based sequence classification into HSM operations. Such models can support real-time decision-making by providing early insights into process deviations, allowing operators to take corrective actions that improve product uniformity, reduce rework, and enhance throughput.

In summary, the proposed methodology establishes a solid foundation for intelligent process monitoring and optimization in the hot rolling industry. Future work will focus on expanding the dataset, improving class separability, and benchmarking against alternative learning architectures to further strengthen its applicability and contribution to overall HSM performance.

Acknowledgements.

The authors would like to express their sincere thanks to the Universidade Aberta Capixaba, a program of the Government of the State of Espírito Santo, linked to the Secretaria de Estado da Ciência, Tecnologia, Inovação e Educação Profissional, for the institutional and financial support (T.O. 1064/2023) that enabled the realization of this study. We also thank IFES for the infrastructure provided and AMT (IT and Production Planning Management teams) for the valuable support provided throughout the development of this study.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] S. Latham and C. Giannetti. A tool to combine expert knowledge and machine learning for defect detection and root cause analysis in a hot strip mill. SN Computer Science, vol. 4, n. Article 628, pp. SP, 2023.

[2] J. M. G. Gepitulan and P. Jackson. Digital solutions for the hot strip mill: Leveraging industry 4.0. AIST Digital Transformation Series, vol. 1, n. 1, pp. 42–49, 2024.