Optimizing Steelmaking with Models, AI, and Federated and Continual Learning

State-of-art automation systems apply optimization models to help controlling industrial processes. Such tools perform even better with Federated Learning and Continual Learning, which enable resilient models with network-wide benefits but without compromising privacy.

TEKSTI **PETRI KANNISTO, BERND KLEIMT, BFI (VDEH-BETRIEBSFORSCHUNGSINSTITUT), STEFANO DETTORI, LAURA LAID, LORENZO VANNINI, SILVIA CATENI, TERESA ANNUNZIATA BRANCA, VALENTINA COLLA, SCUOLA SUPERIORE SANT'ANNA, AITOR ALCAZAR FERNANDEZ, EVIDEN BDS R&D SPAIN**

Knowledge management methods are the core of today's data-driven production systems. The related optimization models can receive power and resilience from external supportive services, such as Federated Learning (FL) and Continual Learning (CL).

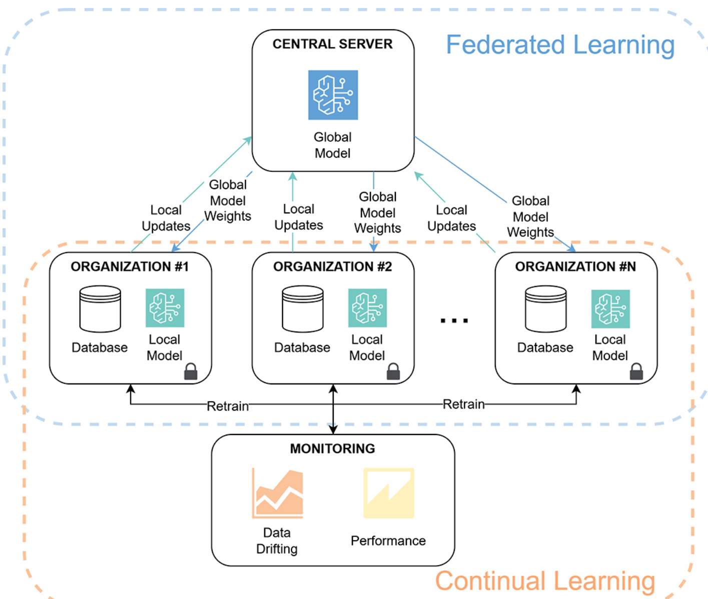

Figure 1. FL ensures privacy by training a global model from local data, while CL addresses model drift, ensuring adaptability and resilience.

Federated Learning and Continual Learning

How can you contribute to collective knowledge, receive feedback, and still preserve your privacy? This is the challenge that Federated Learning (FL) addresses [1].

In FL, each participant or organization retains its local AI model and shares only the model updates with a central service. This service aggregates updates from all participants to create a global model, which is then shared across the network in terms of weights. As a result, the network benefits from shared insights without exposing original data, which is often restricted by privacy regulations.

Complementing FL, Continual Learning (CL) provides ongoing support for maintaining these models throughout their lifecycle [2]. CL monitors for model drift, identifying when performance declines and retraining becomes necessary. This capability enhances resilience, a key objective of Industry 5.0 [3], which emphasizes human-centric and sustainable innovation.

Figure 1 demonstrates how FL and CL services facilitate training and maintaining local models in production environments. While FL safeguards the privacy of each production plant, CL addresses the challenges of model drift, ensuring robust, adaptive performance.

Figure 2. The route from scrap to steel products. Figure copyright © Celsa Group.

Benefits in steel production

FL can help controlling industrial processes, as shown in ALCHIMIA research project, which develops control methods for electric steelmaking. The theme is paramount for modern society for two reasons: first, steel is an essential for the economy, and second, electric steelmaking reuses steel scrap, which is an essential step towards circularity and green transformation. Scrap recycling allows lower CO2 emissions compared to the integrated steelmaking route from iron ores.

Figure 2 illustrates a typical path in electric steelmaking from raw materials to the final products. As we see, this path is long and complex. This text explores FL possibilities related to two processes: the Electric Arc Furnace (EAF) and the subsequent Ladle Furnace (LF).

FL especially fits the very common case when several steel plants of a company group use the same processes, and the data of one of these processes is incomplete for some reason. This can apply, for example, if the data lacks a large variety of operating conditions, covering only a few exceptional situations. Furthermore, different stages in digitalization can mean that a process or area lacks suitable historical data to train models. In such cases, the FL-generated global model can help creating a well-performing local model even without local experiences and without the need to share data from the original company, which might imply complex IPR issues and potential exposure to cyber-security threats.

Model Parametrization for Electric Arc Furnace

The EAF melts steel scrap with the help of electrical energy.

EAF control can be facilitated based on physical models that monitor and predict the behavior of the process. A physical model calculates an estimate for the properties of the produced liquid steel, such as temperature and chemical composition, based on known inputs, such as electrical energy, oxygen injection and charged scrap amounts [4]. Figure 3 illustrates the structure of such a physical model with input and output values.

Although a physical model is an untypical target for FL, FL is being developed for the parametrization of the model. In this approach, multiple processes share their model parameters with the central FL server, which will generate a global model parameter set to help in local model parametrization.

Furthermore, CL will help detecting drifts in the prediction accuracy of the physical models due to changes in process conditions. Additionally, it can reveal when an increase in data coverage can improve model performance. In these cases, the CL scheme suggests retraining to maintain the model performance.

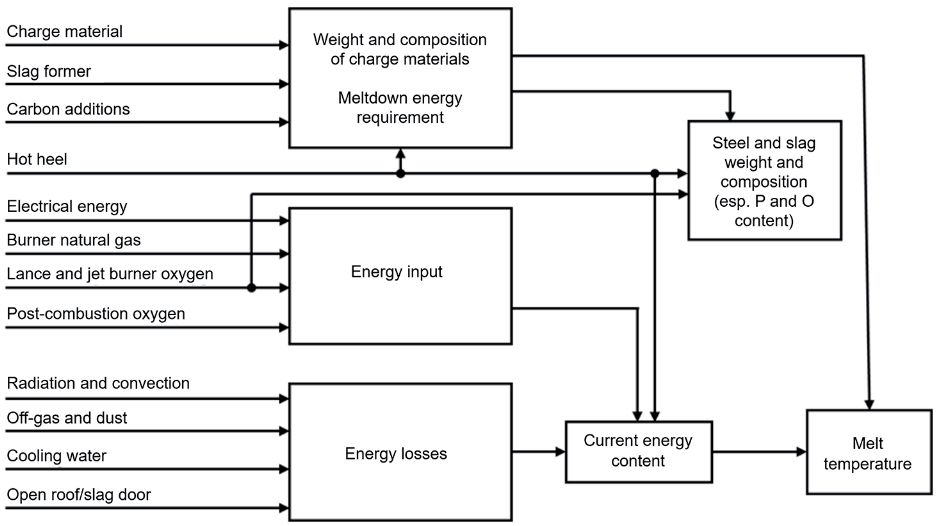

Figure 3. The structure, inputs and outputs of the dynamic EAF process model.

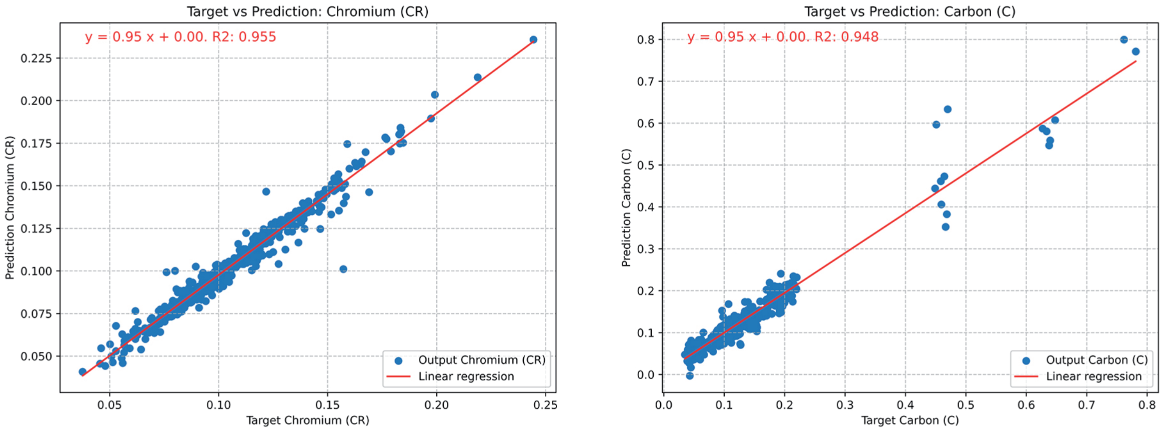

Figure 4. Performance of the LF model (for Chromium and Carbon contents).

More Effective Training for Ladle Furnace

LF process uses the crude liquid steel tapped from the EAF as its input. The operation principle is to receive the crude steel and iteratively add ferroalloys and electrical energy to reach the target conditions of liquid steel for continuous casting.

In this work, the LF model is a Feed-Forward Neural Network. Like many AI models, this neural network is trained with data, and a good data quality ensures proper model training. Therefore, we preprocess the data to eliminate unreliable measurements [5] and, using this consolidated database, thanks to learning algorithms, the model can automatically tune its internal parameters. The model calculates the final steel chemical composition and temperature from the data it receives, in terms of the initial steel chemical composition and temperature, electrical energy and argon usage and the ferro-alloys additions. Figure 4 illustrates, as an example, the performance of the models related to two outputs (Chromium and Carbon content), depicting the relation between the measured target value and the simulated one.

In the FL scheme, each local LF model from the plants will be shared with the central server. In return, the global model will help in adapting the local models to the common, network-wide knowledge for the optimal performance.

On the other hand, CL will reveal degraded model performance and suggest retraining for the local models as needed.

Outlook

The ideas and models presented in this text result from the Horizon-Europe-funded research project ALCHIMIA, which focuses on advanced modelling with techniques, such as AI, FL, and CL in steel production, as well as human factors to consider the stakeholders involved. The results will include not only optimization models but also an optimization framework for the decision support of operators. The project will deliver its final results during 2025.

Funding

This work has been supported by the project "Data and decentralized Artificial intelligence for a competitive and green European metallurgy industry" (ALCHIMIA), which has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101070046. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Bibliography

[1] J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, "A survey on federated learning: challenges and applications," International Journal of Machine Learning and Cybernetics, vol. 14, no. 2, pp. 513-535, 2023. https://doi.org/10.1007/s13042-022-01647-y

[2] L. Wang, X. Zhang, H. Su, and J. Zhu, "A comprehensive survey of continual learning: Theory, method and application," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 8, pp. 5362-5383, 2024. https://doi.org/10.1109/TPAMI.2024.3367329

[3] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang, "Industry 4.0 and Industry 5.0—inception, conception and perception," Journal of Manufacturing Systems, vol. 61, pp. 530-535, 2021. https://doi.org/10.1016/j.jmsy.2021.10.006

[4] B. Kleimt, W. Krieger, D. Mier Vasallo, A. Arteaga Ayarza, and I. Unamuno Iriondo, "Model-based decision support system for electric arc furnace (EAF) online monitoring and control," Metals, vol. 13, no. 8, 2023. https://doi.org/10.3390/met13081332

[5] M. Vannucci, V. Colla, A. Zaccara, S. Dettori, L. Laid: Unsupervised Anomaly Detection Combining PCA and Neural Gases, Communications in Computer and Information Science, 2141 CCIS, 419-432, 2024. https://doi.org/10.1007/978-3-031-62495-7_32