Artificial neural networks in skin-pass mill modeling: an empirical analysis of architecture and hyperparameters

Redes neurais artificiais na modelagem de laminadores de encruamento: uma análise empírica de arquitetura e hiperparâmetros

[1] Italo Pinto Rodrigues

[2] Gabriel Alberto Rodrigues

[3] Bruno Lima Dos Santos

1   Doutorando no curso de Engenharia e Tecnologia Espaciais do Instituto Nacional de Pesquisas Espaciais - INPE
2   Discente do curso de Engenharia do Centro Universitário de Volta Redonda - UniFOA
3   Discente do curso de Engenharia do Centro Universitário de Volta Redonda - UniFOA

ABSTRACT

AI To ensure industrial competitiveness, especially within the steel sector, it is imperative to optimize expenditures and enhance productivity. Against this backdrop, preventing incidents associated with automation and equipment maintenance emerges as a key element to maintain uninterrupted production. Furthermore, minimizing material loss is paramount. Within this framework, a unique opportunity is identified to improve the control of a skin-pass mill, ensuring that the mechanical tension exerted on the steel sheet complies with established specifications. To establish effective control, it is vital that the controller transmits an appropriate signal to the plant. For the validation of control signals, a suitable model of the mill plant is essential. This paper investigates various hyperparameters in MLP-type artificial neural networks using the Levenberg-Marquardt backpropagation algorithm for modeling this plant. These findings are subsequently juxtaposed with a technique widely recognized in the literature, the N4SID. The data suggests that the neural network-based model showcases an improvement of 14% compared to the outcomes derived from N4SID and, moreover, provides the capability for continuous training for better alignment with the real phenomenon.

Keywords:

Artificial Neural Network. System Identification. Data-driven Modelling. Skin-pass Mill. Steel Making

RESUMO

Para garantir a competitividade industrial, especialmente no setor siderúrgico, é imperativo otimizar despesas e aumentar a produtividade. Contra esse pano de fundo, prevenir incidentes associados à automação e manutenção de equipamentos surge como um elemento-chave para manter a produção ininterrupta. Além disso, minimizar a perda de material é de suma importância. Neste contexto, identifica-se uma oportunidade única de melhorar o controle de um laminador de encruamento, garantindo que a tensão mecânica exercida sobre a chapa de aço esteja em conformidade com as especificações estabelecidas. O principal objetivo deste artigo é verificar a viabilidade das redes neurais para modelar o stress mecânico em um laminador de encruamento. Para estabelecer um controle eficaz, é vital que o controlador transmita um sinal adequado para a planta. Para a validação dos sinais de controle, é essencial um modelo adequado da planta do laminador. Este artigo investiga vários hiperparâmetros em redes neurais artificiais do tipo MLP, utilizando o algoritmo de retropropagação Levenberg-Marquardt para modelar esta planta. Essas descobertas são posteriormente comparadas com uma técnica amplamente reconhecida na literatura, o N4SID. Os dados sugerem que o modelo baseado em rede neural mostra uma melhoria de 14% em comparação com os resultados obtidos a partir do N4SID e, além disso, oferece a capacidade de treinamento contínuo para melhor alinhamento com o fenômeno real..

Palavras-chave:

Rede Neural Artificial. Identificação de Sistemas. Modelagem orientada a dados. Laminador de encruamento. Siderurgia.

1   INTRODUCTION

The rolling process requires a properly adjusted controller so that the skin-pass mill can finish the steel sheet. If the controller is not well adjusted, the quality of the steel sheet can be affected (SEO *et al.*, 2020).

Shen *et al.* (2022) assert that physical models overlook the history of mills after successive material passes, specifically their natural degradation. In order to achieve proper controller performance, it is crucial for the system plant to accurately represent the real-world conditions. While modeling techniques can provide assistance, capturing the complexities of a complex system can present significant challenges. Alves *et al.* (2012) highlight key difficulties, including the incorporation of interactions between multiple process variables, nonlinearities, and time delays within the model. Moreover, the system behavior is heavily influenced by operating conditions and the mill's speed.

As time goes by, the controller ends up operating on a plant that is different from its initial state in terms of parameters, which can result in a loss of quality in the steel sheet. So, in order to overcome the complexity of modeling and obtain an accurate model of the mill, data-driven modeling, also known as identification or empirical modeling, can be a useful solution. This approach helps ensure the quality of the steel sheet, preventing material waste and financial losses. To achieve this, operational system data is required for the identification process (AGUIRRE, 2015; SHI *et al., 2022b; SHEN et al., 2022).*

Previously, Rodrigues *et al*. (2013) proposed the use of the N4SID (Numerical Algorithm for Subspace State Space System IDentification) identification method for the skin-pass mill. However, the data sampling considered is not sufficient to adequately map the aging process of the mill.

One way to perform this identification is by using artificial intelligence, particularly Artificial Neural Networks (ANNs), due to their ability to map the model that relates inputs and outputs. The application of such concepts, also present in this project, contributes to avoiding unnecessary losses in the processes (COLLA, 2022; SHI *et al*., 2022).

Escribano *et al*. (2012) applied various Artificial Intelligence Algorithms, such as artificial neural networks and regression trees, to model the behavior of the cold rolling mill, and they obtained satisfactory results.

According to Shen *et al*. (2022), ANN-based models can demonstrate an improvement of up to 50% in predicting the rolling force to be applied to the material, when compared to the physical model. This improvement was observed because the model identified by the ANN considered the steel chemistry, physical rolling parameters, and rolling histories for all rolling stands in an industrial hot strip mill (SHEN *et al., 2022).*

Even though hyperparameter adjustments were made empirically in the studies by Pican *et al.* (1996), Wiklund and Sandberg (2002), He and Liu (2005), and *Ren et al*. (2021), the authors did not clearly present the results of the different topologies, namely, the intermediate architectures in terms of numbers of layers and neurons per layer.

This paper, considering the advantages of artificial neural networks, presents the influence of the number of neurons in performing the identification of a skin-pass mill using MLP networks. The model predicts the mechanical stress applied to the steel sheet.

In this context, the main objective of this paper is to present an exploratory analysis of modeling using artificial neural networks with data from a rolling mill. The neural network architecture will be empirically adjusted, focusing on hyperparameters such as the number of layers and the number of neurons in each layer, to demonstrate the feasibility of using ANNs to model this type of system.

2   BACKGROUND ON FEEDFORWARD NEURAL NETWORK

Artificial Neural Networks (ANN) are mathematically expressed by Equations 1 and 2 (BISHOP, 2006; HAYKIN, 2009):

$$a_j^{[l]} = \sigma\left(z_j^{[l]}\right) \tag{1}$$

$$z_j^{[l]} = \sum_{k=1}^{m} w_{jk}^{[l]} x_j^{[i]} + b_j^{[l]} \tag{2}$$

where a is the output of the neuron, σ is an activation function, w is the weight, x is the input vector of the input layer, and b is the bias. m is the number of inputs, the superscript $[l]$ denotes the lth layer, the superscript (i) denotes the ith training example, the subscript j denotes the input unit, and the subscript k denotes the neuron unit.

Artificial neurons are processing units characterized by inputs, outputs, weights, biases, and a non-linear transfer function, as shown in Figure 1. During ANN training, the weights and biases are updated to minimize the error between the target value and the ANN output (BISHOP, 2006; HAYKIN, 2009; MAIER et al., 2023).

Figure 1 - Neural network architecture with emphasis on the artificial neuron.
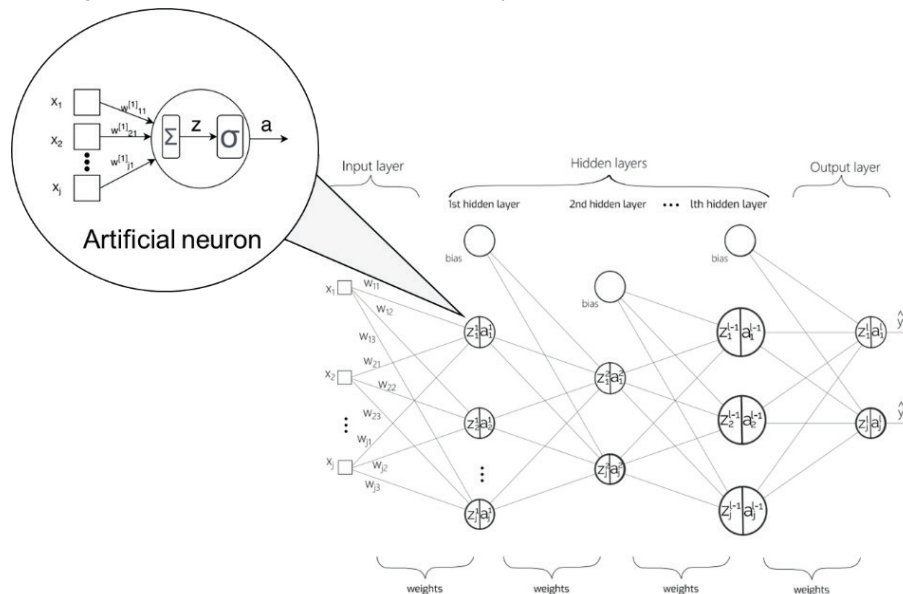
Source: Adapted from Haykin (2009) and Rodrigues *et al.* (2021).

Artificial Neural Networks (ANN) can be structured in various architectures, incorporating activation functions, and learning algorithms. The architecture of an ANN consists of distinct layers - input, hidden, and output - each with a specific size. The size of the input layer is determined by the number of inputs, while the number of hidden layers and neurons within them may vary based on the specific

application. Lastly, the number of neurons in the output layer is determined by the desired number of outputs. (HAYKIN, 2009; RODRIGUES et al., 2021).

A Multi-layer Perceptron (MLP) is a type of ANN that is widely used to solve various problems in different fields (ROSENBLATT, 1958; WERBOS, 1974). The backpropagation algorithm has played a crucial role in enhancing the flexibility of MLPs. With backpropagation learning, the network can effectively learn and encode the mapping between inputs and outputs (RUMELHART; MCCLELLAND, 1987; HAYKIN, 2009).

In the learning process, the network receives an input and passes it through each neuron until it reaches the output layer, producing an output value. The first hidden layer adjusts the input to have values between -1 and 1, which helps ensure all inputs are on the same scale. This adjustment, called mean normalization, allows the network to work effectively with the data, as shown in Equation 3 (RUMELHART; MCCLELLAND, 1987; HAYKIN, 2009).

$$x_{norm} = \frac{x_i - \frac{1}{N}\sum_{i=1}^{N} x_i}{\max(x) - \min(x)} \tag{3}$$

where N is the number of samples available, and x is the ith input sample. The computation of $z_{[l](i)}$ in the first hidden layer is given by Equation 4.

$$Z^{[l]} = W^{[l]}X_{norm} + B^{[l]} \tag{4}$$

where the symbols represented by uppercase letters are matrices. The output of this first layer is calculated with an activation function as in Equation 5.

$$A^{[l]} = \sigma\left(Z^{[l]}\right) \tag{5}$$

In the subsequent steps of ANN computations, the output from the previous layer becomes the input for the current layer. This allows for the flow of information from one layer to the next in a sequential manner, given by Equation 6.

$$z_j^{[l+1]} = \sum_{k=1}^{m} w_{jk}^{[l+1]} a_j^{(i)[l]} + b_j^{[l+1]} \tag{6}$$

In the output layer, the estimated output $\hat{y}_j$ is obtained as in Equation 7:

$$\hat{y}_j = a_j^{[l+1]} = z_j^{[l+1]}$$

$$\tag{7}$$

When dealing with a linear function, the output from the neural network is compared to the desired target value, resulting in an error signal (), as shown in Figure 2. This error signal is then propagated backwards through the network, layer by layer, to adjust the weights and biases. This process continues iteratively until certain stopping criteria are met. These stopping criteria can be defined by the number of epochs (iterations), a specific performance index, or a combination of both. The learning process halts when these criteria are satisfied, indicating that the network has reached an acceptable level of accuracy or performance (HAYKIN, 2009).

Figure 2 - Diagram of artificial neural network identification.
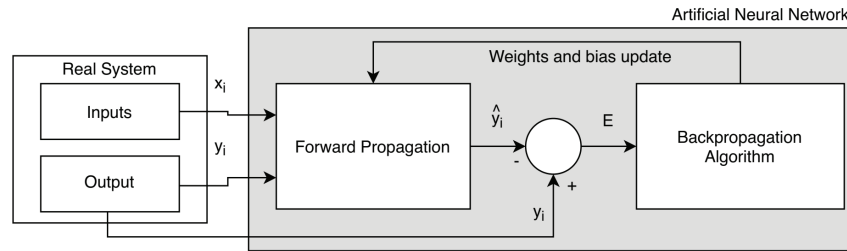
Source: Adapted from Haykin (2009) and Rodrigues *et al.* (2021).

In this study, the training algorithm employed was the Levenberg-Marquardt (LM) method. The LM method combines elements from the methods developed by Levenberg (1944) and Marquardt (1963) aiming to strike a balance between them. The evaluation of the Artificial Neural Network's performance is done using the mean-squared error (MSE) metric, shown in Equation 8, which quantifies the average difference between the predicted and actual outputs (HAYKIN, 2009).

$$E_{MSE} = \frac{1}{N} \sum_{i=1}^{N} [(y_i - \hat{y}_i)^2] \tag{8}$$

where N is the number of samples, y_i is the measured sample and $\hat{y}_j$ is the estimated value.

During the ANN learning process, the data set is divided into the following sets (HAYKIN, 2009; RODRIGUES et al., 2021): i) training set: the training set used to compute errors and to determine the weights and biases update at each iteration during the learning/training process; ii) validation set: the validation set used to minimize over-fitting; iii) testing set: the test set used to estimate the generalization error.

3   SKIN-PASS MILL

The data presented in this study are obtained from a skin-pass mill from Sumitomo. The "mill region", illustrated in Figure 3 refers to the area where the identification will be performed. In this region, the following equipment can be found (RODRIGUES *et al*., 2013):

- Tensiometers: composed of three cylinders attached to a load cell. This assembly measures the force that will be used in the calculation of mechanical stress.

- Backup rolls: are rollers that are in direct contact with the material to be rolled.

- • Central tensioner: it is a set composed of three tensioning rollers, driven by a single motor. The central tensioner is responsible for ensuring the mechanical stress at the output of the rolling mill.

•

Figure 3 - Skin-pass mill diagram.
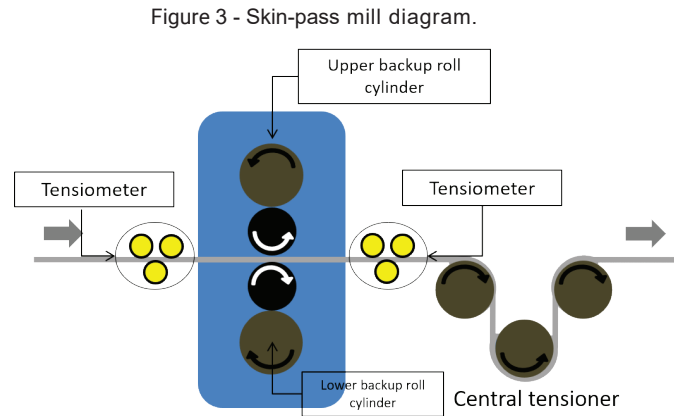
Source: Adapted from Rodrigues *et al*. (2013).

Through hydraulic actuation the lower backup roll cylinder moves approaching the upper backup roll cylinder. The force ($\vec{F}$) applied to the steel strip causes the elongation of the sheet. In order to achieve the desired elongation, it is necessary for the steel strip to be tensioned at the entry and exit of the rolling mill, ensuring that the sheet remains fully stretched (RODRIGUES *et al*., 2013).

The variables that will be used in the identification are presented in Table 1, which also shows the equipment to which these variables are related, the measurement method, the unit of measurement, and their classification in the control system (input or output).

The thickness and width of the steel sheet determine the mechanical stress reference that the system needs to develop in order to process the material without experiencing any defects, loosening, or sheet breakage. It is crucial for the system to generate an appropriate level of mechanical stress to ensure that the material is processed effectively and without any detrimental issues (RODRIGUES *et al*., 2013).

The input mechanical stress information is sent to the motor control of the lower backup roll cylinder, and from there the motor controls the mechanical stress at the entry by adjusting its speed. Similarly, the output mechanical stress information is sent to the control of the central tensioner motor, which then controls the stress at the exit of the rolling mill (RODRIGUES *et al*., 2013).

Table 1 - Variables to be used in the identification.

Variable	Equipment	Sensor	Unit	Classification
Mechanical stress at the input	Tensiometer	Load Cell	kgf/mm^2	Output
Backup roll cylinder speed	Lower backup roll cylinder	Encoder	mpm	Input
Mechanical stress at the output	Tensiometer	Load Cell	kgf/mm^2	Output
Tensioner speed	Central tensioner	Encoder	mpm	Input

Source: Rodrigues et al. (2013).

4   METHODOLOGY

Given the inherent variability in network outcomes based on architecture, experiments delineated in Table 2 were conducted. For this study, the activation function for the hidden layers was chosen as the hyperbolic tangent sigmoid, while the output layer employed a linear activation function. The primary goal was to iteratively adjust the parameters, evaluating the influence of each on the identification process.

Table 2 - Configurations of the ANN architectures.

Case	Layer	Neurons per layer	Total number of neurons
1	1	2	2
2	1	4	4
3	1	8	8
4	1	10	10
5	2	[2 2]	4
6	2	[4 4]	8
7	2	[8 8]	16
8	2	[10 10]	20
9	3	[2 2 2]	6
10	3	[4 4 4]	12
11	3	[8 8 8]	24
12	3	[10 10 10]	30

Source: The authors (2023).

In Table 2, different configurations of neural networks are laid out specifically highlighting the combinations of hidden layers and their respective neuron counts. For clarity, in case 4, there's a single hidden layer with 10 neurons. Meanwhile, case 12 showcases a design with three hidden layers, each containing 10 neurons. The last column in the table provides the total number of neurons present in the neural network's architecture.

It is noteworthy that the dataset encompasses around 8000 samples. The data has been sourced from an authentic apparatus, although it remains inaccessible for public usage. This dataset was divided as follows: 70% for training, 15% for validation, and 15% for testing.

The toolbox used in this study is the feedforwardnet, part of Matlab. The tool allowed for the swift simulation of various architectures, facilitating data control and management.

By combining the number of hidden layers and the number of neurons, the estimated values are obtained and compared with the actual values. The quality of the identification is given by the MRSE (Mean Relative Squared Error), given by Equation 9.

$$MRSE(\%) = \frac{1}{n_o} \sum_{o=1}^{n_o} \left[\sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(\hat{y}_i)^2}} \right] \qquad (9)$$

where n_o is the total number of outputs, and N is the amount of data used to measure the ANN's performance. The result of MRSE indicates that the closer it is to zero the better the identified model is, which means that the model fits the data well.

5   RESULTS AND DISCUSSIONS

In our intensive exploration into the nuanced applications of Artificial Neural Networks (ANNs) in modeling the processes of a skin-pass mill, we unearthed some notable findings. Most significantly, the optimal neural configuration emerged from our experiment 12, which comprised three hidden layers with 10 neurons each. This model recorded an MRSE of a mere 0.7288%. In stark contrast, the simplest model we examined with a single hidden layer and just two neurons registered an MRSE of 0.9577%. The Figure 6 displays all the results obtained from the experiments. The MRSE is highlighted by the color bar.

It becomes clear that as the number of neurons increases, the results improve. However, the optimal definition of the number of neurons, or even the number of layers, is not well-documented in the literature, as pointed out by Svozil *et al*. (1997), Zhang *et al*. (1998), and Maier *et al*. (2023). Although the empirical parameter adjustment process revealed a satisfactory result, the hyperparameter calibration can be driven by a heuristic optimization process, as indicated by Maier *et al*. (2023), potentially utilizing techniques such as genetic algorithms.

Figure 6 - Heat map showing the influence of the number of layers and neurons on MRSE (%).
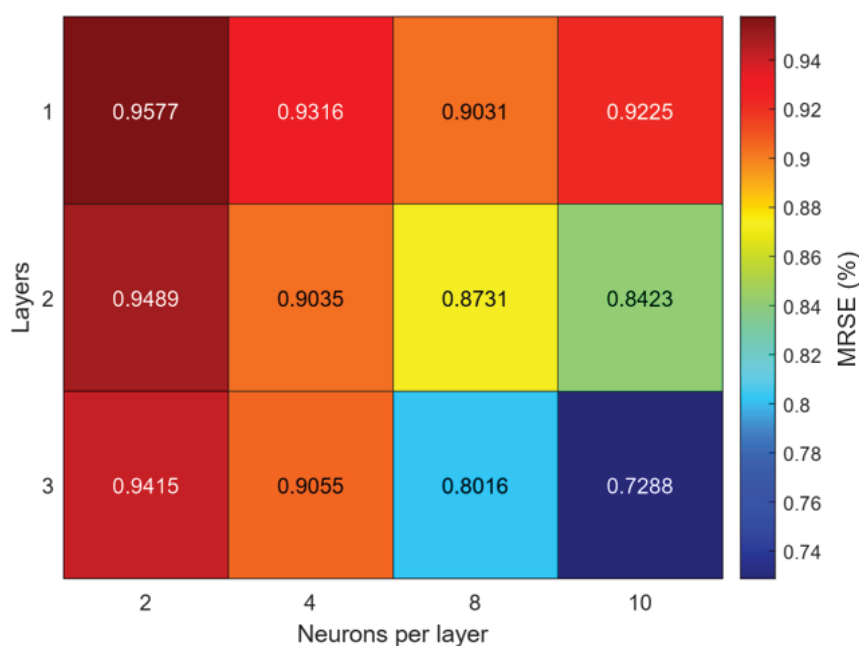
Source: The authors (2023).

Drawing a comparative analysis, prior methodologies, notably the N4SID technique from existing literature (Rodrigues et al., 2013) yielded an MRSE of 0.84%. In this light, our results using neural networks not only surpassed this benchmark but also accentuated the potential advantages of ANNs over more traditional modeling techniques, especially for skin-pass mill operations.

Achieving a 14% improvement in the accuracy of a neural model for a temper mill holds significant implications for the steel manufacturing industry. Firstly, the financial aspect cannot be understated. Precise control in the temper mill process directly translates to cost efficiency. Even seemingly minor inaccuracies can have substantial economic ramifications, leading to material wastage and increased operational costs. Beyond the economic dimension, the quality of the final steel product is paramount. The temper mill process critically influences attributes such as surface finish, flatness, and tensile properties. By improving the accuracy of the modeling process, industries can ensure consistent quality, which is

especially vital in sectors with stringent standards like automotive or construction. Furthermore, operational efficiency sees a boost with accurate modeling. Predicting outcomes with higher precision means fewer real-time operational errors, which in turn results in decreased downtime and recalibrations. This streamlining of the production process can lead to significant time savings and improved throughput. Finally, the safety implications of accurate modeling in such a high-stakes environment as steel production are profound. Reducing the potential for equipment malfunctions through better predictive modeling not only safeguards expensive machinery but, more importantly, minimizes risks to the operators and staff.

Throughout our experiments, a consistent trend became evident: as the number of neurons increased, the model's performance similarly enhanced. Further augmentations in neurons per layer amplified this positive outcome. This observation is not just academically interesting, but also bears significant implications for the future of ANN-based modeling in various industrial settings. It is also noticeable that the best result, observed in Figure 7 was obtained with the architecture that had the highest number of neurons, specifically 30 neurons equally divided into 3 layers.

Figure 7 - Mechanical Stress Estimates Produced by the Identified Architecture.
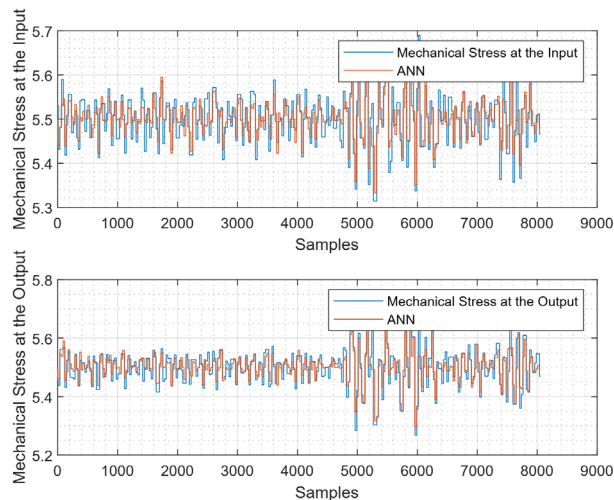
Source: The authors (2023).

While MRSE served as our primary yardstick for evaluating how well the neural model replicated real-world processes, during our iterative training process, we predominantly relied on the MSE for backpropagation and optimization. This distinction is an important one, offering a nuanced understanding of the metrics behind our results. It's also worth noting that our choice for the hyperbolic tangent sigmoid activation function for the hidden layers was deliberate, suggesting a potential area of study regarding the effects of alternative activation functions in comparable scenarios.

Although our ANN provided an impressive representation of the skin-pass mill on a simulation level, its transition to real-world application remains a step we are yet to take. This next phase, while promising, brings with it its own set of challenges and complexities. One of the primary hurdles we faced in our research was the logistics and intricacies associated with obtaining genuine mill data. Navigating through manual data transfers compounded by access restrictions and other operational constraints added another layer of challenge to our endeavor.

On a concluding note, our research offers a tantalizing hint that even more intricate ANN architectures might hold the potential to further reduce error margins. Yet, the practical implementation as

presented by Santos; Barcelos (2020) of these advanced models in real-time industrial operations, such as a skin-pass mill, demands a balanced examination of computational demands against the anticipated accuracy gains. As we reflect on our findings, our primary aspiration is that this study catalyzes further inquiry and innovation, marrying neural theory with industrial pragmatism in ever more effective ways.

6   CONCLUSIONS

Artificial neural networks can be fine-tuned through adjustments like neuron count per layer, layer numbers, and activation functions. Increased neurons, when evenly distributed across layers, lowered the MRSE (%), showing a moderate enhancement compared to earlier research. However, these findings, while consistent within this study, may not apply universally due to dataset-specific attributes influencing neural network performance.

A drawback of neural networks is their potential decreased generalization when faced with variable ranges outside their known boundaries. Addressing this issue should be a focus for subsequent research.

Future endeavors will explore if introducing additional input parameters, such as rolling force, enhances model accuracy. The priority remains understanding variable interrelations, ensuring the model encompasses all relevant physical parameters in the system dynamics. A streamlined network could also facilitate integration into the plant's PLC controller.

ACKNOWLEDGMENT

REFERENCES

AGUIRRE, L. A. **Introdução à identificação de sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais**. 3 ed ed. Belo Horizonte: UFMG, 2007.

ALVES, P. G.; CASTRO, J. A. DE; MOREIRA, L. P.; HEMERLY, E. M. Modeling, simulation and identification for control of tandem cold metal rolling. **Materials Research**, v. 15, n. 6, p. 928-936, 2012.

BISHOP, CHRISTOPHER M. **Pattern Recognition and Machine Learning**. 1. ed. New York: Springer-Verlag, 2006.

COLLA, V. A big step ahead in Metal Science and Technology through the application of Artificial Intelligence. **IFAC-PapersOnLine**, v. 55, n. 21, p. 1-6, 2022. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S2405896322014641>.

ESCRIBANO, R.; LOSTADO, R.; MARTÍNEZ-DE-PISÓN, F. J.; PERNÍA, A.; VERGARA, E. Modelling a Skin-Pass Rolling Process by Means of Data Mining Techniques and Finite Element Method. **Journal of Iron and Steel Research International**, v. 19, n. 5, p. 43-49, 2012. Disponível em: <http://link.springer.com/10.1016/S1006-706X(12)60098-3>.

HAYKIN, S. **Neural Networks and Learning Machines**. 3. ed. Bookman, 2009.

HE, HAI-TAO; LIU, HONG-MIN. The research on integrated neural networks in rolling load prediction system for temper mill. 2005 International Conference on Machine Learning and Cybernetics. **Anais...** . p.4089-4093. v. 7, 2005. IEEE. Disponível em: <http://ieeexplore.ieee.org/document/1527653/>.

LEVENBERG, K. A METHOD FOR THE SOLUTION OF CERTAIN NON-LINEAR PROBLEMS IN LEAST SQUARES. **Quarterly of Applied Mathematics**, p. 164-168, 1944. Disponível em: <https://www.jstor.org/stable/43633451>. Acesso em: 26/8/2023.

MAIER, H. R.; GALELLI, S.; RAZAVI, S.; *et al*. Exploding the myths: An introduction to artificial neural networks for prediction and forecasting. **Environmental Modelling & Software**, v. 167, p. 105776, 2023. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1364815223001627>.

MARQUARDT, D. W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. **Journal of the Society for Industrial and Applied Mathematics**, v. 11, n. 2, p. 431-441, 1963. Disponível em: <https://www.jstor.org/stable/2098941>. Acesso em: 26/8/2023.

PICAN, N.; ALEXANDRE, F.; BRESSON, P. Artificial neural networks for the presetting of a steel temper mill. **IEEE Expert**, v. 11, n. 1, p. 22–27, 1996. Disponível em: <http://ieeexplore.ieee.org/document/482953/>.

REN, X.-Y.; GAO, H.-M.; XU, H.-W.; HUANG, H.-G.; SUN, J.-N. Identification and Control of Elongation System of Skin Passing Mill Based on Intelligent Algorithm. **Journal of Physics: Conference Series**, v. 1820, n. 1, p. 012154, 2021. Disponível em: <https://iopscience.iop.org/article/10.1088/1742-6596/1820/1/012154>.

RODRIGUES, I. P.; JORGE, J. M.; OLIVEIRA, K. F. DE. **Identificação e controle de um laminador de encruamento em malha fechada através de métodos de subespaços**, dez. 2013. Trabalho de Conclusão de Curso, Volta Redonda: Centro Universitário de Volta Redonda.

RODRIGUES, I. P.; OLIVEIRA, P. A. S.; AMBROSIO, A. M.; CHAGAS, R. A. J. Modeling satellite battery aging for an operational satellite simulator. **Advances in Space Research**, v. 67, n. 6, p. 1981-1999, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0273117720309042>.

ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychological Review**, v. 65, n. 6, p. 386-408, 1958.

RUMELHART, D. E.; MCCLELLAND, J. L. Learning Internal Representations by Error Propagation. **Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations**. p.318-362, 1986. MITP. Disponível em: <https://ieeexplore.ieee.org/document/6302929>. Acesso em: 26/8/2023.

SANTOS, B. C. DOS; BARCELOS, A. F. Aplicação da rede neural artificial como ferramenta de diagnóstico e controle do sistema de tensão de um laminador de tiras a frio. XVII - Simpósio de Excelência em Gestão e Tecnologia. **Anais...** , 2020. Resende: Centro Universitário Dom Bosco do Rio de Janeiro.

SEO, M.; BAN, J.; KOO, B. Y.; KIM, S. W. Static Model Identification for Sendzimir Rolling Mill Using Noise Corrupted Operation Data. **IEEE Access**, v. 8, p. 150685-150695, 2020. Disponível em: <https://ieeexplore.ieee.org/document/9169645/>.

SHEN, S.; GUYE, D.; MA, X.; YUE, S.; ARMANFARD, N. Multistep networks for roll force prediction in hot strip rolling mill. **Machine Learning with Applications**, v. 7, p. 100245, 2022. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S2666827021001237>.

SHI, P.; GAO, H.; YU, Y.; XU, X.; HAN, D. Intelligent fault diagnosis of rolling mills based on dual attention-guided deep learning method under imbalanced data conditions. **Measurement**, v. 204, p. 111993, 2022.

SHI, P.; YU, Y.; GAO, H.; HUA, C. A novel multi-source sensing data fusion driven method for detecting rolling mill health states under imbalanced and limited datasets. **Mechanical Systems and Signal Processing**, v. 171, p. 108903, 2022. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0888327022000905>.

SVOZIL, D.; KVASNICKA, V.; POSPICHAL, J. Introduction to multi-layer feed-forward neural networks. **Chemometrics and Intelligent Laboratory Systems**, v. 39, n. 1, p. 43-62, 1997. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0169743997000610>.

WERBOS, P. **Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science**, 1974. Harvard University.

WIKLUND, O.; SANDBERG, F. Modelling and Control of Temper Rolling and Skin Pass Rolling. **Metal Forming Science and Practice**. p.313-343, 2002. Elsevier. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/B9780080440248500151>.

ZHANG, G.; EDDY PATUWO, B.; Y. HU, M. Forecasting with artificial neural networks: **International Journal of Forecasting**, v. 14, n. 1, p. 35-62, 1998. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0169207097000447>.